Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?

Michael Kremer[1] and Alaka Holla[2]

[1]Department of Economics, Harvard University, Cambridge, MA 02138;
email: mkremer@fas.harvard.edu

[2]Innovations for Poverty Action, New Haven, Connecticut 06511;
email: alaka.holla@gmail.com

Key Words

school attendance, peer effects, school quality, teacher incentives

Abstract

Across a range of contexts, reductions in education costs and provision of subsidies can boost school participation, often dramatically. Decisions to attend school seem subject to peer effects and time-inconsistent preferences. Merit scholarships, school health programs, and information about returns to education can all cost-effectively spur school participation. However, distortions in education systems, such as weak teacher incentives and elite-oriented curricula, undermine learning in school and much of the impact of increasing existing educational spending. Pedagogical innovations designed to address these distortions (such as technology-assisted instruction, remedial education, and tracking by achievement) can raise test scores at a low cost. Merely informing parents about school conditions seems insufficient to improve teacher incentives, and evidence on merit pay is mixed, but hiring teachers locally on short-term contracts can save money and improve educational outcomes. School vouchers can cost-effectively increase both school participation and learning.

1. INTRODUCTION

Two billion children, 85% of the world's total, live in the developing world. Their futures and those of their children depend on whether they go to school and how much they learn there. There is fairly strong evidence that education promotes individual earnings (Duflo 2001, Psacharopoulos & Patrinos 2004), technology adoption (Foster & Rosenzweig 1996), and health improvements and fertility reduction (Schultz 1997, 2002; Strauss & Thomas 1995). Although there is little evidence for Lucas (1988) style production externalities or for aggregate returns to education beyond those which would be expected based on the micro returns (see Acemoglu & Angrist 2000, Bils & Klenow 2000, Krueger & Lindahl 2001, Pritchett 2001, Duflo 2004), some have argued that education promotes democracy and better institutions (Barro 1997, 1999; Glaeser et al. 2004), and Hanushek & Woessmann (2008) find a strong relationship between scores on international achievement tests and economic growth.

Even in the absence of production externalities, improving the productivity of education investment is important because education accounts for a substantial share of total investment in developing countries. Standard National Income and Product Accounts suggest that education accounts for 18% of total investment (World Bank 2006), and a much higher share of public investment, but Jorgenson & Fraumeni (1989) argue that National Income Accounts do not properly account for the value of students' time and that, when properly measured, human capital investment constitutes over 80% of total investment in the United States.

Economists have long been interested in the determinants of human capital investment decisions by households, the determinants of learning, and the functioning of educational systems. Historically, randomized evaluations were conducted rarely, usually when governments decided to evaluate a particular large-scale program. Over the past decade and a half, economists studying education in the developing world have increasingly used randomized evaluations to shed light on ways to help more children attend school and improve learning. One strand of the literature examines conditional cash transfer programs following the PROGRESA evaluation in Mexico (Gertler & Boyce 2001, Gertler 2004, Schultz 2004). Another strand of the literature, following work in Kenya (Kremer et al. 2003, Glewwe et al. 2009a), is based on academics working with nongovernmental organizations (NGOs), often using cross-cutting designs and iteratively examining alternative approaches to compare a variety of strategies to address a particular problem. It now seems opportune to review what we have learned, both about education and about randomized evaluations. This article discusses findings from interventions that aimed to increase the quantity or quality of education in a developing-country context and draws on several earlier reviews (Banerjee & Duflo 2006; Kremer & Holla 2008, 2009; Glewwe et al. 2009b).

User fees have been a controversial issue in development policy, but across a variety of contexts, randomized evaluations consistently find that reducing the out-of-pocket cost of education or instituting subsidies consistently increases school participation, often dramatically. Providing merit scholarships to cover costs for subsequent levels of education can also induce greater effort by students.

At first blush, the finding that investments in schooling are sensitive to price might seem a vindication of standard models of human capital investment, but a more careful examination of the results suggests the need to supplement these models. Children are

more likely to go to school if their peers do, contrary to some simple models in which schooling decisions depend on a tradeoff with child labor, but consistent with models in which children have considerable agency over whether they attend school. The magnitude of the behavioral response to out-of-pocket costs often seems larger than expected from a standard human capital investment model unless a large mass of households are just on the margin of attending school; provision of free school uniforms, for example, leads to 10%–15% reductions in teen pregnancy and dropout rates. Deferring payment of conditional cash transfers to coincide with the time fees are required for the next level of education has a larger impact on subsequent enrollment than evenly spaced transfers throughout the year, suggesting that time-inconsistent preferences or savings constraints may be as big a barrier to education as credit constraints.

School participation can be increased without large increases in public spending through the provision of school health programs (in particular, mass deworming) and information about earnings differences between people with different levels of education. Finally, improvements in school quality generally do not lead more children to attend school.

Some common lessons are also emerging on ways to improve the quality of education, although generalization is somewhat more difficult in this area. Student learning in developing countries is often abysmal (Hanushek & Woessmann 2008). Supplying more of existing inputs, such as additional teachers or textbooks, often has a limited impact on student achievement because of distortions in developing-country education systems, such as elite-oriented curricula and weak teacher incentives, as manifested by the absence from school of one out of five teachers during unannounced visits (Chaudhury et al. 2006).

Pedagogical innovations that work around these distortions can improve student achievement at low cost. Technology-assisted learning or standardized lessons can mitigate weaknesses in teaching and substantially improve test scores; remedial education can help students catch up with the curriculum; and tracking students by initial achievement benefits not only students with high initial achievement, but also students with lower initial achievement.

Although linking teachers' pay to attendance can increase student learning, improving teacher incentives may require going outside civil-service structures. There is mixed evidence on whether linking teacher pay to students' test scores promotes students' long-term learning. Simply providing information to communities on school quality without actually changing authority over teachers has little impact. Reforms that go outside the civil-service system by giving local committees resources to hire teachers on short-term contracts or by providing students with vouchers that can be used to attend private schools can increase learning dramatically.

The rest of this article is structured as follows. Section 2 reviews evidence on alternative ways of increasing the quantity of education. Section 3 argues that increasing existing educational inputs often has limited impact owing to distortions in education systems, whereas Section 4 argues that providing inputs to facilitate changes in pedagogy can improve learning. Sections 5 and 6 address strengthening teacher incentives through merit pay, local oversight, and hiring of teachers under short-term contracts. Section 7 reviews evidence on school vouchers, and Section 8 concludes with a discussion of implications for policy and for randomized evaluations.

2. HELPING MORE CHILDREN GO TO SCHOOL

Educational attainment has risen dramatically in the past 40 years. In 1960 in low-income countries, 14% of secondary school–age children were in secondary school, and the working-age population had an average of 1.6 years of education. By 2000, 54% of secondary school–age children were in secondary school, and the average education in these countries was 5.2 years (Barro & Lee 2001). Nonetheless, 100 million children of primary school age—15% of the worldwide total—are not in school (UNESCO 2006). Of these, 42 million are in Sub-Saharan Africa, and 37 million are in South Asia, while 55 million are girls.

This section reviews evidence from randomized interventions on ways to increase access to education, with subsections addressing costs and subsidies (for a more extensive discussion, see Kremer & Holla 2009), merit scholarships, information on returns to education, school-based health interventions, and the responsiveness of school attendance to improvements in school quality.

2.1. Education Costs and Subsidies

Three randomized studies in Kenya measure the responsiveness of school participation to reducing out-of-pocket costs of education. Historically, parents of school children had to provide uniforms, which cost approximately $6, slightly under 2% of Kenya's per capita GDP. Kremer et al. (2003) find that students in primary schools selected for ICS-Africa's Child Sponsorship Program, which paid for these uniforms, remained enrolled an average of 0.5 years longer after five years and advanced an average of 0.3 grades further than their counterparts in comparison schools. This program had other components, although the authors argue that they did not affect enrollment.

Two other studies, however, can isolate the impact of the uniforms. Providing free uniforms to young primary school students reduced their absence by one-third, or 6 percentage points, with larger effects (13 percentage points, or 64%) for students who lacked uniforms prior to the program (Evans et al. 2008). Providing free uniforms to sixth-grade girls reduced dropout rates by 2.5 percentage points from a baseline rate of 18.5% and reduced their childbearing by 1.5 percentage points (from a baseline rate of 15%). It reduced boys' dropout rates by 2 percentage points from a baseline rate of 12% (Duflo et al. 2006).

Subsidizing education can further increase school attendance. The pioneering PRO-GRESA conditional cash transfer program in Mexico provided up to three years of monthly cash grants equivalent to one-fourth of average family income for poor mothers whose children attended school at least 85% of the time, with premia for older children and girls in junior secondary school. The program increased enrollment reported in household surveys by 3.4–3.6 percentage points for all students in grades 1 through 8 and increased the transition rate from elementary school to junior secondary school by 11.1 percentage points from a base of 58%. Among these older children, girls' enrollment increased by 14.8 percentage points, significantly more than boys' enrollment, which grew by 6.5 percentage points (Schultz 2004). The greater impact at older grades suggests that education attainment could perhaps be increased even more if the subsidies were targeted more toward older students. School re-entry after dropout increased among older children, and younger children repeated grades less often, suggesting improvements in effort

among the treatment group, as grade progression typically depends on grades during the school year (Behrman et al. 2005). Interestingly, repetition decreased even for children in grades 1 and 2 who were not yet eligible for program benefits, consistent with spillovers from older siblings, anticipation effects, or a household income effect from the subsidy itself.

Attanasio et al. (2005) and Todd & Wolpin (2006) use the experimental data generated by the PROGRESA experiment to test structural models of parental decisions about schooling and fertility and make out-of-sample predictions using their models to compare the existing PROGRESA subsidy schedule with several alternatives that were not experimented with in the actual program. They argue that eliminating the subsidy in lower grades, where attendance is almost universal, and increasing it in upper grades would leave overall program costs unchanged but increase average completed schooling.

Based in part on the convincing results from the randomized evaluation of Mexico's conditional cash transfer program, similar programs have been established in 25 other countries, mostly middle-income countries. Several have been subject to randomized evaluations, which found similar effects (see Glewwe & Olinto 2004, Maluccio & Flores 2005, Schady & Araujo 2006, Fiszbein & Schady 2009 for a review).

Although it is unsurprising that prices and subsidies affect schooling, a more detailed look at the data suggests a need to go beyond simple models of human capital investment in explaining schooling choices. First, out-of-pocket costs seem more important than models emphasizing opportunity cost would suggest, except in knife-edge cases. Even in a poor country, it is surprising that a $6 uniform could reduce absence rates among primary school students by 64%, as suggested by Evans et al. (2008).

Second, peer effects seem to affect attendance decisions. Schooling gains in the PROGRESA program spilled over to children above the poverty cutoffs for program eligibility. Their primary school attendance increased 2.1 percentage points from a base of 76% (Lalive & Cattaneo 2006), and their secondary school attendance increased 5 percentage points from a base of 68% (Bobonis & Finan 2009). Barrera-Osorio et al. (2007) find that a conditional cash transfer program in Bogota, Colombia, also generated spillovers, reducing school attendance and increasing hours worked for untreated children within treated students' households, but they also provide some evidence of positive treatment spillovers among friends.

Under standard models in which households trade off the value of education against the value of children's time in agricultural labor and household chores, if some children go to school, then labor supply among close substitutes should increase, rather than decrease. Evidence of positive spillovers seems more consistent with an alternative model in which children choose between schooling and social activity with other out-of-school peers. The evidence presented by Barrera-Osorio et al. (2007) is consistent with the standard model operating within the household and a friends model operating outside it.

Third, behavioral issues may limit educational investment. Under one variant of the Colombian conditional cash transfer program, part of the monthly payment for regular school attendance was withheld and saved until school fees had to be paid for the subsequent school year. If families were credit constrained, this forced savings should have reduced the value of the subsidy and hence decreased contemporaneous attendance. Conversely, if saving were difficult because of time-inconsistent preferences, then forced savings could raise enrollment in the subsequent year without deterring contemporaneous attendance. In fact, this program variant increased contemporaneous school attendance

rates by roughly the same amount as a more basic PROGRESA-like treatment—2.8 percentage points (from a base of 79.4%)—but unlike the basic treatment, it also increased enrollment in secondary and tertiary institutions in the subsequent school year by 3.6 and 8.8 percentage points (from bases of 69.8% and 22.7%), respectively (Barrera-Osorio et al. 2007).

Although conditional cash transfers have proven popular in middle-income countries, they require someone to monitor students' attendance, and monitors may not always do this accurately (see Shastry & Linden 2007 for nonexperimental evidence that suggests teachers exaggerate students' attendance in an Indian grain-distribution program that conditions on school attendance). School meals potentially automatically condition transfers on students' school participation. Kremer & Vermeersch (2005) find that a school feeding program in informal, community-run preschools in Kenya increased participation of children in treatment schools by 8.5 percentage points (from a baseline rate of 27%). For children in grades 2 through 5 in Jamaica, receiving a school breakfast everyday increased school attendance by 2.3 percentage points, or 3% (Powell et al. 1998).

2.2. Merit Scholarships

Although many countries may be able to eliminate fees or even institute subsidies for primary education, currently available resources may be inadequate to pay for higher levels of education for all. Evidence from Kenya and Colombia suggests that merit scholarships can induce more effort from students working to qualify. The Girls' Scholarship Program in Western Kenya provided sixth-grade girls scoring in the top 15% in their district exam a two-year award consisting of a yearly grant to cover school fees for the remaining two years of primary school, a yearly grant for school supplies, and public recognition at an awards assembly held for students, parents, teachers, and local government officials. Kremer et al. (2009) find that girls eligible to compete increased their test scores by 0.19 standard deviations.

Estimated program effects were different in the two districts in which the program was administered. In the smaller, less prosperous of the two districts, there was substantial suspicion of the NGO implementing the program and attrition from the program. This makes it difficult to reliably estimate program impact, but it is impossible to reject the hypothesis of no program impact. Test-score gains were concentrated in the larger, slightly more prosperous district, where the estimated gain was 0.27 standard deviations. In this district, student absence also decreased by 3.2 percentage points (from a baseline of 13%), and gains spilled over even to those ineligible for the program. Boys' test scores increased by 0.15 standard deviations in the first cohort affected by the program, and girls with little or no chance of winning the awards also benefited from the program, with treatment effects for girls statistically indistinguishable across all quartiles of the baseline test-score distribution.

These spillover effects may have resulted from interactions among students, but they may also have been mediated by changes in teacher behavior induced by the program. Teacher absence in program schools declined by 4.8 percentage points (from a baseline rate of 16%).

Further evidence of students' increasing effort in response to merit scholarships comes from another variant of the Colombian conditional cash transfer program, which offered students who graduated from secondary school and enrolled in a tertiary institution a

transfer equivalent to 73% of the average cost of the first year in a vocational school. The program increased contemporaneous attendance in secondary school by 5 percentage points from a comparison group base of 79.3% (Barrera-Osorio et al. 2007). In the subsequent school year, enrollment in a tertiary institution increased by a dramatic 49.7 percentage points (from a baseline of 19.3%).

Finally, Berry (2008) provides evidence from India that students with lower initial test scores, compared with their high performing classmates, are more likely to attend an after-school reading tutorial when promised toys upon reaching a literacy goal than when their parents are promised monetary incentives.

2.3. Providing Information

Two studies suggest that informing students about the extent to which earnings vary with schooling can increase school participation with minimal fiscal expenditure. In a baseline survey of eighth-grade boys from 150 schools in the Dominican Republic, Jensen (2007) finds that students underestimated the earnings difference between primary and secondary school graduates by 75%. He conjectures that residential segregation by income leads children to underestimate returns to education because poor children disproportionately observe relatively low earners among those they know with high education, and rich children disproportionately observe high earners among the less educated people they know. It is impossible, however, to rule out the hypothesis that students are reporting an accurate estimate of the causal impact of education on earnings for students with their own characteristics and that this is much less than average differences in income by educational attainment.

In any case, Jensen (2007) finds that providing students with information on earnings differences by education led to upward revisions in perceived returns, reduced dropout by 3.9 percentage points, or 7%, in the subsequent year, and increased school completion by 0.20 years four years later. The reduction in dropout rates was concentrated among households above the median level of income, despite evidence that poorer households did revise their perceived returns to education. Jensen (2007) argues that stimulating the demand for education might not be enough if barriers such as credit constraints still limit the investment that households can make.

Nguyen (2008) finds that informing fourth-grade students in Madagascar and their parents about earnings differences by education levels increased average attendance by 3.5 percentage points (from a baseline of 85.6%), although results are only statistically significant when various information interventions are pooled. Test scores increased by 0.20 standard deviations after only 3 months in response to the information about earnings differences among different education groups. A program variant that brought in local role models with high education from different initial income backgrounds to talk about their life stories raised test scores by 0.17 standard deviations, but only when the role model came from a poor income background similar to background of most students, suggesting that students respond to information when they think it applies to them. The role models, however, had no impact on attendance.

2.4. School-Based Health Programs

School health programs are another extremely cost-effective way to increase school participation. Hookworm, roundworm, whipworm, and schistosomiasis affect 2 billion people

worldwide (WHO 2005) and are particularly concentrated among school-age children. Heavy infections can cause iron-deficiency anemia, protein-energy malnutrition, abdominal pain, and listlessness. Because diagnosis is expensive, but medicine to treat the infections costs pennies per dose and has no serious side effects, the WHO recommends mass school-based treatment by teachers in areas where worms are common.

Miguel & Kremer (2004) evaluate a randomized school-based mass deworming intervention in Kenyan primary schools that provided twice-yearly treatment. The intervention reduced absence rates by 7 percentage points from a baseline rate of 30%. The program also generated health and education spillovers, most likely owing to the interruption of disease transmission both within and across schools. School participation for untreated students in treatment schools was 8 percentage points higher than for their counterparts in comparison schools. Pupils were 4.4 percentage points more likely to be in school for each additional thousand pupils attending treatment schools within 3 km of their school.

Similarly, a deworming and iron supplementation program in urban Indian preschools reduced absence rates by 5.8 percentage points (or 20%) among 4–6 year olds, with the highest gains among groups with high baseline anemia rates—girls and children from low socioeconomic status areas (Bobonis et al. 2004).[1]

2.5. School Quality

Some argue that poor school quality can explain low school attendance (PROBE Team 1999). Indeed, standard human capital theory suggests that people might be more willing to invest in education when its quality is higher. However, in general there is little evidence that programs aiming to increase school quality—for example, by providing more inputs, reforming pedagogy, or improving teacher incentives—increase school participation, even when they do improve learning. Sections 3 through 6 present results from a number of programs that improved school quality but failed to increase student participation.

There is some evidence, however, that switching from one-teacher schools to two-teacher schools can improve attendance and that girls are more likely to attend schools with female teachers. Banerjee et al. (2005) examine the impact of hiring an extra teacher (female when possible) in one-teacher nonformal education centers in rural India. These schools, designed to teach basic numeracy and literacy skills to children who do not attend formal schools, are plagued by high rates of teacher and child absenteeism. Teacher absence in one-teacher schools most often means that these schools close for the day. These schools are also typically staffed by men. Only 19% of teachers in the comparison schools were female. Of the new hires in treatment schools, however, 63% were female.

Hiring an extra teacher increased girls' attendance by 50% (to six on average, from a baseline attendance of four female students per school), as measured from surprise monitoring visits, but the program did not significantly change boys' attendance. There is some evidence that the program had a smaller impact on girls' enrollment when the original teacher in the school was female, suggesting that households may be more likely to send a girl to school if at least one of the teachers is female.

---

[1]It is difficult to track the long-run effects of deworming on education using randomized evaluations, but based on nonexperimental data, Bleakley (2007a,b) concluded that an early twentieth century–deworming campaign in the U.S. South substantially increased education.

2.6. Summary

School attendance has risen dramatically in the developing world, and the randomized evaluations summarized here suggest that it could be raised further with little public expenditure by providing information on earnings differences by education levels and by school health programs. **Figure 1** (see color insert) presents the Abdul Latif Jameel Poverty Action Lab's (2005) summary of the cost-effectiveness of various programs in increasing school participation. At only $3.50 per year of additional education generated, mass deworming is the lowest cost method of increasing school participation examined. (This study did not examine the provision of information about wages for people with different education levels.)

Reducing out-of-pocket costs for parents and providing school meals, conditional cash transfers, and merit scholarships require greater public expenditure, but because these constitute transfers and because they are typically highly progressive, these expenditures may generate important benefits beyond their impact on school participation.

Evidence from randomized evaluations suggesting that peer effects and time-inconsistent preferences may play an important role in education decisions casts some doubt on the presumption that households will choose optimal levels of education in the absence of credit-market constraints. This may help explain the disconnect between economists' models of education, which only suggest narrowly circumscribed reasons for governments to provide education, and the presumption of policy makers that individuals will under-invest in education without government intervention.

Getting children into school, however, is only half the battle. The other half is making sure that they learn more once in school. The rest of this review focuses on the key challenge for the future: improving school quality.

3. INCREASING EXISTING INPUTS

The remaining sections discuss evidence from randomized evaluations on the effectiveness of different options for improving school quality: increases in existing inputs, pedagogical changes, merit pay for teachers, teachers hired outside the civil-service system, and school vouchers. First, however, some background on quality in developing countries may be useful before discussing the effectiveness of increasing various inputs. Many children in developing-country schools learn remarkably little (Lockheed & Verspoor 1991, Glewwe 1999, Hanushek & Woessmann 2008). In Bangladesh, for example, Greaney et al. (1999) found that 58% of rural children age 11 and older could not identify seven of eight presented letters. Internationally, comparable achievement data are available for a number of middle-income countries and a few low-income countries, and with the exception of a few countries in East Asia, scores are much lower in the developing world. In the Trends in International Mathematics and Science Study, for example, the average eighth-grade test-taker in South Africa answered only 18% of the questions correctly on the math portion in 2003, compared with 51% in the United States (Gonzales et al. 2004). Filmer et al. (2006) note that on the Program for International Student Assessment of 2000, the average science score among students in Peru was equivalent to that of the lowest scoring 5% of U.S. students.

What explains these differences in test scores? Governments in developing and developed countries spend similar fractions of GDP on education (World Bank 2006), but

because a much higher fraction of the population in developing countries is of school-age, expenditure per pupil as a fraction of GDP is lower, particularly in primary education. In 2000, governments in high-income countries spent 18.8% of GDP per capita per student on primary education; low-income countries spent just 7%.

Teachers in developing countries are paid salaries that are an average of 3.7 times per capita GDP in low-income countries (UNESCO 2005), and their salaries typically account for three-quarters of education budgets (Bruns et al. 2003). Developing countries respond to the high costs of teachers by maintaining high pupil-teacher ratios—an average of 28:1 in primary school, compared with 14:1 in the developed world (UNESCO 2006).

Nonteacher resources are also often scarce. In a 1995 rural Kenyan sample of schools, for example, 80% of students were in classrooms that had less than one English textbook for every 20 students (Glewwe et al. 2009a). More than 40% of pupils in Sri Lanka did not have a place to sit in 2005 (Zhang et al. 2008).

Teachers often lack skills and face weak incentives. Harbison & Hanushek (1992) find that a sample of teachers in Northeast Brazil scored less than 80% on a Portuguese language test meant for fourth graders. In a cross-country survey, Chaudhury et al. (2006) note that the absence rate for primary school teachers was 27% in Uganda, 25% in India, 19% in Indonesia, 14% in Ecuador, and 11% in Peru. Even when teachers are in school, they are not necessarily teaching. In India, 75% of teachers were present in the school, but only about half of all teachers were actually teaching in the classroom when enumerators arrived. In general, higher-income regions had lower absence rates.

Under efficiency wage models, high salaries spur workers to exert more effort to avoid risking their jobs. However, because only 1 out of 3000 public-school headmasters surveyed in India reported a case in which a teacher was fired for absence, it is unsurprising that Chaudhury et al. (2006) report little evidence that absence rates are correlated with teacher salaries and instead find that absence is more correlated with working conditions, such as school infrastructure.

Evidence from randomized evaluations suggests that reducing pupil-teacher ratios (Section 3.1) and increasing existing nonteacher inputs (Section 3.2) often have limited impact on test scores owing to systemic distortions, such as weak teacher incentives and orientation of curricula toward elites (Section 3.3).[2]

3.1. Increasing Teacher Inputs

Some experimental and nonexperimental studies in developed countries suggest that reducing pupil-teacher ratios can increase test scores, although others challenge this view (Krueger 1999, Angrist & Lavy 1999, Krueger & Whitmore 2001, Hanushek 2002, Krueger 2002, Rice 2002). There is no single definitive study from the developing world, but evidence from two randomized studies in Kenya and two in India suggests that reducing pupil-teacher ratios often has little or no impact on test scores.

In Kenya's Extra Teacher Program, ICS-Africa provided school committees with funds to hire an additional teacher on a contract basis for grades 1 and 2, bringing the average pupil-teacher ratio down to 46 from 84. Students assigned to civil-service teachers in treatment schools, however, scored no better than their counterparts in comparison schools (Duflo et al. 2007a).

---

[2]See Hanushek (1995) for a discussion of nonexperimental evidence.

As discussed in Section 2.1, when the Child Sponsorship Program in Kenya paid for uniforms that parents would ordinarily have had to pay for, treatment schools, which had a baseline class size of 27 students, received an average of 9 additional students, partly because of transfers (Kremer et al. 2003). Despite these extra students, students attending the treatment schools prior to the program did not suffer significant test-score losses.[3] Thus, it is possible to accommodate large increases in class size after reducing out-of-pocket costs without reductions in learning.

The evidence from India is similar. The randomized introduction of an extra teacher in rural Indian nonformal education centers (discussed in Section 2.5) also did not significantly improve test scores, although this effect was measured with considerable noise (Banerjee et al. 2005). Finally, the Balsakhi Program in two urban areas removed 15–20 low-performing students from the classroom and provided tutoring by women from the local community for two hours everyday. The roughly 20 students who remained in the classroom therefore had a lower pupil-to-teacher ratio for these two hours. After two years, however, these students did not score any better than students in comparison classrooms (Banerjee et al. 2007a).

Why does this evidence from Kenya and India differ from some of the experimental evidence from the United States that shows achievement gains from smaller class sizes? This could perhaps be because in developing countries, initial pupil-teacher ratios are much larger, and even in the context of small classes, teachers rarely tailor instruction to individual children's needs, possibly because they are not trained or incentivized to do this. As discussed in Section 3.3, pupils did learn more, at least in the context of the Kenyan Extra Teacher Program, when hiring an extra teacher was accompanied by splitting classes by initial achievement.

3.2. Increasing Nonteacher Inputs

A widespread view, articulated by Filmer & Pritchett (1999) in their review of retrospective studies, is that political economy problems distort education expenditure toward teachers rather than nonteacher inputs. Filmer & Pritchett estimated that nonteacher inputs have a marginal product per dollar that is 10–100 times higher than teacher inputs. Even education-spending skeptics often believe that textbook availability promotes learning (Lockheed & Hanushek 1988).

Glewwe et al. (2009a), however, find no significant impact of textbooks on average test scores in schools in rural Kenya and reject the hypothesis that textbook provision raised average student test scores by as little as 0.07 standard deviations. The program also failed to reduce grade repetition, dropout rates, and student absence. Similarly, Glewwe et al. (2004) note that flipcharts presenting material from the Kenyan science, math, or geography curriculum failed to improve test scores. In both cases, however, nonexperimental estimates suggest a strong positive impact of the inputs.

3.3. Systems Distortions

Closer examination of the data in the randomized evaluations above suggests that the failures to increase scores likely resulted in part from broader distortions in education

---

[3]The program paid for textbooks, but (as discussed below) textbooks did not improve average scores in this context.

systems, such as weak teacher incentives and inappropriate curricula. Teacher absence might explain why the Kenyan Extra Teacher Program did not increase test scores for students assigned to civil-service teachers in program schools. Having contract teachers in the treatment schools reduced the likelihood that civil-service teachers were in class and teaching by 12.9 percentage points from a base of 58.2%.

Again, a deeper examination of the textbooks data shows how systems distortions can undermine the effectiveness even of apparently essential nonteacher inputs. Although textbooks did not increase average test scores, they did benefit the strongest students. Students with baseline test scores in the top quintile scored 0.22 standard deviations higher on the endline test if they were in textbook schools (Glewwe et al. 2009a). Moreover, eighth-grade students in treatment schools were more likely to enter secondary school than students in comparison schools, a finding consistent with textbooks' being most helpful to initially high-achieving students because only academically strong students progress to secondary school.

Textbook provision most likely created this heterogeneous impact because textbooks are designed for a curriculum and education system that are oriented toward the strongest students. For example, the language of instruction in the Kenyan education system after the first few years of school is English, most students' third language. The median students in lower grades had difficulty even reading the textbooks.

Glewwe et al. (2009a) argue that a confluence of three factors common in many developing countries leads to this mismatch between the curriculum and the needs of the majority of the population: (*a*) a centralized education system with a single national curriculum; (*b*) heterogeneity among students, associated with rapid expansion of education; and (*c*) the political dominance of the elite who prefer an education system and instructional materials targeted toward their children.

4. REFORMING PEDAGOGY TO CORRECT SYSTEM DISTORTIONS

Although test scores are often unaffected by additional existing resources, they can be increased, sometimes dramatically, by inputs that allow shifts in pedagogy that can overcome weaknesses in teachers' training and incentives and that address curricular weaknesses. Experiments in the United States suggest that direct instruction, a pedagogical approach in which teachers' activities are prescribed in detail, can raise test scores, especially for poor-performing students (see Kirschner et al. 2006 for a review).

Three randomized programs in developing countries suggest that technology-assisted programs that help impose an appropriate curriculum can improve learning.[4] Jamison et al. (1981) evaluate a Nicaraguan program in which some first-grade classrooms were randomly assigned 150 daily radio mathematics classes of 20–30 min in length and a postbroadcast lesson taught by the teacher. A second group received mathematics workbooks. After one year, students exposed to radio instruction scored 1.5 standard deviations higher on mathematics tests than students in a comparison group, and students assigned workbooks scored about one-third of a standard deviation higher.

Similarly, a computer-assisted learning program in urban India increased test scores in mathematics (Banerjee et al. 2007a). Trained instructors from the local community gave primary school children two hours of computer access per week to play educational games

---

[4]Barrow et al. (2008) find similar results in the United States.

emphasizing basic math skills from the government curriculum. Test scores increased by 0.35 standard deviations after one year and by 0.46 standard deviations after two years. One year after the program ended, some of the gains no longer persisted, but a 0.1 standard-deviation advantage did remain.

In another program in Indian schools, He et al. (2007) find that provision of either electronic-machine- or flash-card-based activities designed to help teach English increased test scores in English by 0.3 standard deviations. In schools randomly chosen to have the material introduced by teachers rather than NGO workers, there were also spillovers to test scores in math.

These results certainly do not imply that technology is the key to improving learning in the developing world. Schools had computers prior to the computer-assisted learning program evaluated by Banerjee et al (2007a); they were just rarely used. It is worth noting that in all these programs, technology was used to fundamentally change the type of instruction given to students rather than simply integrated into the existing curriculum and implemented with teachers' discretion. Kouskalis (2008) argues this is key to their success and provides some nonexperimental evidence suggesting that the introduction of computers into schools in Southern Africa had little impact on learning. Some randomized evaluations of computer-assisted learning in developed countries are also consistent with this view (Angrist & Lavy 2002, Campuzano et al. 2009).

In two separate remedial education programs in India, semivolunteers with low levels of education helped students catch up with the curriculum when given carefully scripted lessons and formulas for teaching children who have fallen behind. The Balsakhi Program in two different urban areas in India (discussed in Section 3.1) paid young women from the community less than one-tenth of a regular teacher's salary to tutor children who had failed to master basic literacy and numeracy skills for two hours per day outside the classroom. Test scores for all children increased in treatment schools by 0.14 standard deviations after one year and by 0.28 standard deviations after two years, with most of this increase due to large gains among children at the bottom of the test-score distribution and among the children who received the remedial instruction (Banerjee et al. 2007a). These treatment effects were consistent across the two different cities in which the program was implemented (Mumbai and Vadodara). One year after the program ended, a 0.1 standard-deviation test-score advantage over the comparison schools persisted.

A reading intervention in rural India trained community volunteers who had a tenth- or twelfth-grade education for four days to teach children how to read and significantly improved reading achievement (Banerjee et al. 2009). In 55 of the 65 villages that received the program, volunteers started reading camps, which enrolled roughly 8% of all children in the treatment villages, with initially poor readers more likely to attend the camps. When receipt of the treatment instruments for reading camp attendance, children who attended reading camps were 22.3 percentage points more likely to be able to read at least several letters and 23.2 percentage points more likely to be able to read at least a word or paragraph.

Although existing curricula may serve many students poorly, simply adjusting curricula downward would poorly serve stronger students, with potentially large social costs. Splitting classes could potentially allow instruction to be tailored to students' needs, but one concern is that this might hurt students assigned to the lower tracks.

The Kenyan Extra Teacher Program discussed above sheds light on this issue. In half of the schools randomly selected to receive an extra teacher, students were randomly assigned

to classrooms. In the other half, students above the median in pre-intervention achievement scores were assigned to one classroom and those below the median level were assigned to the other. (These classrooms were then randomly assigned to either the contract teacher or the civil-service teacher.)

Both initially low– and initially high–achieving students learned more under tracking, contrary to simple peer-effects models in which all students benefit from having high-achieving peers. Tracked students gained 0.14 standard deviations more than their counterparts in untracked program schools after 18 months (Duflo et al. 2008). One year after the program ended, the tracked students had a 0.16 standard-deviation advantage. Gains were statistically indistinguishable for the students in the below-the-median and above-the-median classes. In fact, a regression discontinuity analysis shows that in tracked schools, scores of students near the median of the pretest distribution are independent of whether they were assigned to the above-the-median or below-the-median classroom. In contrast, in untracked program schools, students benefit on average from having academically stronger peers. This suggests that tracking was beneficial because it helped teachers focus their teaching to a level appropriate to most students in the class.

Teacher behavior seems to be an important channel through which tracking affects students. Civil-service teachers in the tracked program schools were 11.2 percentage points more likely to be found in class and teaching on a random day than their counterparts in the untracked program schools, who were in class and teaching only 45% of the time, a decrease in absence entirely due to the greater presence of teachers assigned to the above-the-median class.[5] This may reflect easier or more pleasant teaching conditions when students are better prepared and less heterogeneous.

5. TEACHER INCENTIVES

This section reviews evidence from programs that more directly tried to change teacher incentives, either by conditioning teacher compensation on teacher attendance (Section 5.1) or on students' test scores (Section 5.2). It draws on Glewwe et al. (2009b) and on Banerjee & Duflo (2006), who discuss alternative strategies to reduce teacher and health worker absence.

5.1. Compensation Based on Attendance

As noted above, systems for monitoring teacher presence are not functioning in many parts of the developing world. Would providing teachers incentives to attend more frequently improve learning or would teachers start coming to school but not teach or teach so poorly that students would learn little?

A study in rural India suggests that improvements in teacher attendance do translate into higher test scores. Duflo et al. (2007b) evaluate a program that rewarded high presence rates among teachers in NGO-run nonformal education centers, using cameras with tamper-proof time and data functions to measure teacher presence. The teachers were not civil servants but local community members who were paid low wages. Prior to the intervention, teacher absenteeism was high (approximately 44%), despite an NGO policy of dismissal for absence. The camera program decreased teacher absence in treatment

---

[5]Civil-service teachers in tracked program schools were also 5.4 percentage points more likely to be found in school.

schools by 21 percentage points, roughly halving the absence rate over its 30-month duration.

Treatment-school teachers were no less likely to be teaching conditional on being in school than comparison school teachers. Students in treatment schools also scored 0.17 standard deviations higher on tests after one year, and the graduation rate to mainstream government schools increased by 10 percentage points (from a baseline of 16%).

Although this study suggests that a system of automatic monitoring with enforcement by physically remote agents who are prepared to enforce the rules is technically feasible and indeed provides better incentives for teachers,[6] a later effort to introduce this system with higher-skilled, higher-status, and more politically powerful health-care workers ran into strong political obstacles (Banerjee et al. 2007b). Kremer & Chen (2001) also find that teacher absence did not decrease under a program in rural Kenya in which head-masters were supposed to reward teacher presence with bonuses. Headmasters provided bonuses to all teachers regardless of attendance, suggesting it is difficult to change an entrenched culture of high teacher absence through local monitors who have discretion.

5.2. Compensation Based on Students' Test Scores

An alternative approach to providing teachers incentives is to link teacher salaries to student performance. Compared with compensation based on attendance, compensation based on student performance could potentially lead teachers to increase not only school attendance, but also time spent on lesson planning, homework, or teaching conditional on presence in school. It could also lead more talented teachers to enter the profession. Conversely, opponents of test-score-based teacher incentives argue that teachers' tasks are multidimensional and only some aspects are measured by test scores, so linking compensation to test scores could cause teachers to sacrifice the promotion of curiosity and creative thinking to teach the skills tested on standardized exams (Holmstrom & Milgrom 1991, Hannaway 1992).

The extremely weak teacher supervision systems in many developing countries raise the potential for both these benefits and costs. It can be argued that teachers in many developing countries already teach to the test and that the main problem is to get teachers to come to school. Developing countries, however, may also be more prone to attempts by teachers to game any incentive system.

A project in western Kenya provides evidence on these hypotheses. Annual prizes ranging from 21% to 43% of monthly salary were offered to teachers in grades 4–8 in schools that either were top-scoring or had most improved on the annual government (district) exams administered in those grades. The program created incentives not only to raise test scores, but also to reduce dropout rates, as students who did not take the government exams at the end of the year were assigned low scores.

Glewwe et al. (2008) find that students in treatment schools were more likely to take the government exam in both years of the program. This improved scores on the formula, but teachers induced students just to take exams rather than to stay in school. The program did not affect dropout, repetition, or eighth-grade graduation rates.

---

[6]One issue is whether results would generalize to multiteacher schools in which teachers may feel less responsibility to be in the classroom and have more opportunities to socialize with other teachers.

Treatment-school students on average gained 0.14 standard deviations more on the incentivized government exams than their counterparts in comparison schools. However, several patterns in the data suggest that teachers did not change their behavior to improve long-term learning but rather focused on short-run signaling. First, teacher absence rates did not improve. Second, students did not report an increase in homework assignments, nor did observers notice any change in pedagogy (such as the use of blackboards or teaching aids or teachers' levels of energy or caring). Instead, 88% of treatment schools reported an increase in exam-preparation sessions.

Third, the NGO administering the program also tested students with an exam that differed from the government exams in content and format and that was designed to detect performance differences among a wider range of students. No incentives were attached to performance on this exam, and in neither of the intervention years did test scores improve significantly on this exam.

Fourth, even on incentivized exams gains were short-lived, consistent with a model in which the teachers orientated things toward short-run test scores rather than long-run learning. In postintervention years, they had completely dissipated, and there were no differences between the treatment and comparison schools on the government exams. In contrast, gains from some other programs in the same area persisted (e.g., the Extra Teacher Program and the Girls' Scholarship Program).

There is some direct evidence that the program improved test-taking techniques. On the NGO exams, students in treatment schools were less likely to leave answers blank, more likely to answer multiple-choice questions, and less likely to leave answers blank at the end of the test. This may help explain the absence of a significant effect on the NGO exams, particularly on questions with a format differing from that of the government tests because these exams had fewer questions that would have benefited from these test-taking skills.

A program in India that was part of the Andhra Pradesh Randomized Evaluation Study (APRESt) also linked teacher pay to student test scores, paying for every percentage-point improvement in teachers' test scores, with a 10 percentage-point improvement in scores yielding a bonus equivalent to approximately 30% of a monthly salary. In some respects, results from the Indian program were similar to those of the Kenyan program.

As in Kenya, the Indian APRESt incentives program increased test scores by 0.22 standard deviations over two years but did not change teacher absence rates. Direct classroom observation indicates that treatment teachers were not more likely to be teaching at the point of observation. Teachers in treatment schools, however, were 38 percentage points more likely to report that they provided special preparation for the exams (from a comparison base of 25%).

However, there is some evidence that the APRESt program, unlike the Kenyan program, improved learning. Test-score improvements were identical on questions with an unfamiliar format that were designed to be conceptual and on questions with a more familiar format, and pupils in incentive schools scored 0.11 and 0.18 standard deviations higher in science and social studies, respectively, subjects that were not linked to incentives. These results are consistent with increases in human capital rather than simple improvements in test-taking techniques, although it is possible that test-taking techniques that helped on the standard questions also helped on the more conceptual questions and on tests in other subjects.

In one APRESt program variant, individual teachers received the bonus pay based on average test-score gains made by their own students; in the other variant, all teachers in a school received the bonus based on average gains made in the entire school. In the first year of the program, test-score gains were similar in the individual- and school-based incentive groups (Muralidharan & Sundararaman 2008a). In the second year of the program, however, test scores increased by 0.27 standard deviations in the individual-incentives schools, which was significantly higher (at the 10% level) than the 0.16 standard deviation increase in the group-incentives schools.

In summary, both evaluations in Kenya and India suggest linking teacher pay to student performance does not lead to reduced teacher absence but does lead to increased preparatory sessions. In the Kenyan context, there is no evidence that long-run learning increased, but in India the current evidence is much more favorable, although it is not possible to know whether test-score improvements in achievement will be sustained after the incentives have been removed as the APRESt program is ongoing. The success of merit-pay programs may very well be dependent on context.

6. SYSTEMS REFORMS

This section reviews evidence from more fundamental reforms to teacher incentives that involve going beyond civil-service systems. Many argue that providing information to parents can substantially improve the accountability of service providers (World Bank 2004). Among three interventions that provided communities with more information on schools to empower parents (Section 6.1), two had no statistically significant impact on teacher performance or student achievement, whereas one had modest effects. Conversely, two interventions that transferred real control over hiring and firing decisions to parents in Kenya and India did increase test scores (Section 6.2). Section 6.3 argues that decentralizing authority over schools without decentralizing financing responsibility can create serious distortions.

6.1. Parental Involvement

In 2000, the government of Uttar Pradesh in India established Village Education Committees (VECs) that were supposed to monitor the performance of the schools, report problems to higher authorities, hire and fire community-based teachers, and use any additional resources for school improvement from a national education program. Most people, however, knew close to nothing about VECs. In household surveys, nearly 92% of households did not know that a VEC existed, and among the nonheadmaster members of the VECs, 23% did not even know that they were members (Banerjee et al. 2006). These surveys, combined with assessment tests given to children, also indicate that parents considerably overestimated their children's reading and math abilities and were not fully aware of how poorly their schools were functioning.

In a randomized intervention in this area, a team of NGO workers organized a village meeting to inform people about the quality of local schools, state-mandated provisions for schools (pupil-teacher ratios, infrastructure, mid-day meals, and scholarships), local funds available for education, and the responsibilities of VECs. A second intervention did all of this and also gave villagers a specific monitoring tool by encouraging and equipping

communities to participate in testing to see whether children can read simple text and solve simple arithmetic problems.

An average village of approximately 360 households sent about 100 people to the meetings, but in neither treatment group were VECs more likely to perform any of their functions than in the comparison group. Similarly, treatment parents were not more involved with their children's school nor were they more likely to know about the state of education in their village or consider it a major issue. (Banerjee et al. 2009). Both interventions also failed to improve teacher and student absence, which remained high at 25% and 50%, respectively.

Two programs in Kenya provided parents with some influence over civil-service teachers. In Kenya, the teacher service commission hires teachers in public primary schools centrally and assigns them to schools. Promotion is determined by the Ministry of Education, not by parents. However, local school committees, comprising mostly students' parents, have historically raised funds for schools' needs, such as classroom repairs or textbook purchases, and sometimes use these funds to hire lower-paid contract teachers locally to supplement the regular civil-service teachers.

One program aimed to improve incentives for civil-service teachers with prizes awarded by school committee members and to strengthen ties between school committees and local educational authorities through training and joint meetings. Preliminary results show that average teacher attendance did not change (de Laat et al. 2008). In treatment schools, committee members met more often with parents, but teacher behavior in school was rarely discussed and teacher absence was never discussed. There is also little systematic and significant evidence that pedagogy within the classroom changed or that student attendance or achievement improved. There is some evidence that the existence of the program led to changes in who was elected to the school committee, with more educated and older people becoming members of the committee as the committee had increased authority.

Under the Kenyan Extra Teacher Program first described in Section 3.1, half of the school committees in untracked program schools and half in the tracked program schools were also randomly selected for training to help them monitor the contract teachers (e.g., soliciting inputs from parents, checking teacher attendance), and a formal review meeting was arranged for the committees to review the contract teachers' performance and decide on contract renewal. The monitoring program had no impact on the attendance rates of the contract teacher, but contract teachers had very high attendance rates in any case, as discussed below (Duflo et al. 2007a). Civil-service teachers were 7.3 percentage points more likely to be in class and teaching in program schools with monitoring than civil-service teachers in program schools without monitoring, although this result is not quite significant at the 10% level. Students assigned to these civil-service teachers in schools with monitoring were 2.8 percentage points more likely to be in school and scored 0.18 standard deviations more in math than students of civil-service teachers in program schools without monitoring.

In sum, there is mixed evidence on the impact of training school committees to monitor civil-service teachers or giving them a share in authority over these teachers, and it is reasonable to expect that effects might differ by context. However, in none of the cases observed so far has this generated dramatic improvements and outcomes.[7]

---

[7]However, a community mobilization program to monitor health clinics in Uganda substantially improved clinic attendance and health outcomes (Bjorkman & Svensson 2009).

6.2. Local Hiring of Contract Teachers

In response to the high cost and low quality of some centralized school systems, some countries have established alternative, locally controlled systems, including a system for hiring contract teachers outside the normal civil-service system. Evidence from the Balsakhi Program in India (discussed in Sections 3 and 4) suggests that hiring outside of the civil-service system can be an effective way to improve student achievement. In the cross-country survey of absence discussed above, Chaudhury et al. (2006) find that teachers from the local area are less likely to be absent in all six countries, although neither contract teachers nor teachers in nonformal education centers in the Indian survey attended significantly more often than teachers in government-run schools.[8]

Some evidence of the effectiveness of contract teachers comes from evaluating the contract teachers themselves in the Kenyan Extra Teacher Program (Duflo et al. 2007a).[9] Unlike the regular civil-service teachers, the contract teachers were directly hired by local school committees, which consisted mostly of school parents, and they were not unionized or subject to civil-service protections. These contract teachers had the same academic qualifications as regular teachers but were paid less than one-fourth as much. In program schools, they were roughly 16 percentage points more likely to be in class and teaching than civil-service teachers in comparison schools, who attended 58.6% of the time, and 29.1 percentage points more likely than civil-service teachers in program schools, corresponding to 39% and 70% lower absence rates, respectively.

Students assigned to contract teachers scored 0.23 standard deviations higher and attended school 1.7 percentage points more often (from a baseline attendance rate of 86.1%) than students who had been randomly assigned to civil-service teachers in program schools.

Similar results arose in another contract teacher program conducted in India alongside the APRESt teacher-incentives program (discussed in Section 5.2). Unlike the Kenyan Extra Teacher Program, this program did not assign the contract teachers to particular classrooms and did not employ contract teachers with similar academic qualifications as their civil-service counterparts. Only 44% of contract teachers had at least a college degree, and only 8% had received a formal teacher training degree or certificate, compared with 85% and 99%, respectively, for civil-service teachers.

Contract teachers in the APRESt program were 10.8 percentage points less likely to be absent over two years than their civil-service counterparts, who had an absence rate of 26.8%; they were also 8.4 percentage points more likely to be engaged in teaching activity during random spot checks at the school, compared with a baseline rate of 39% (Muralidharan & Sundararaman 2008b). Their students' test scores improved by 0.12 standard deviations.

As in the Kenyan Extra Teacher Program, civil-service teachers in schools with contract teachers increased their absence rates by 2.4% and decreased teaching activity by 3.2%.

Several characteristics of contract teachers could potentially account for these successes. They served at the discretion of local school committees rather than having civil-service protection and employment security. There were also more likely to come from the local area. It is also important to recognize that many contract teachers eventually become

---

[8]This latter result also holds with village/town fixed effects.

[9]For nonexperimental studies, see Vegas and de Laat (2005) and Umansky & Vegas (2007).

civil-service teachers, and that may be an important part of their motivation, so the results cited above should perhaps be interpreted as the impact of a system in which teachers first work on a probationary status (as advocated by Gordon et al. 2006) rather than as the effect of a system in which there were only contract teachers and no civil-service teachers.

Figure 2 (see color insert) summarizes the changes in teacher presence induced by some of the reviewed programs. Simply hiring extra contract teachers without increasing their incentives through changes in the classroom environment or through increases in local oversight may actually decrease effort from existing civil-service teachers. Monitoring by headmasters that could use their discretion resulted in no attendance gains, nor did performance pay linked to test scores or provision of information about teacher absence.

External monitoring without discretion (cameras) and contract teachers do generate large impacts on teacher attendance at low cost. The Abdul Latif Jameel Poverty Action Lab (2008) estimates that the cost per additional day of teacher presence was $2.20 in the camera program; costs were actually negative for the Kenyan contract teachers if the alternative is hiring extra civil-service teachers because contract teachers were paid much less.

Although not providing any explicit incentives to teachers, merit scholarships that increase student effort and tracking students by their incoming achievement levels also can motivate teachers to show up for work at very little cost [$2.00 per additional day of teacher presence for the case of merit scholarships (Abdul Latif Jameel Poverty Action Lab 2008), the cost of hiring an extra contract teacher for the case of tracking]. Training school committees also encouraged marginally more attendance in Kenya at a total cost of approximately $75 per school when allowing for a Ministry of Education staff member to conduct the training and less if training were conducted more locally.

6.3. Pitfalls of Mismatches Between Authority and Responsibility Under Partial Decentralization

Kremer et al. (2003) argue that the partially decentralized system of school finance set up after Kenyan independence distorted incentives for education spending and the setting of school fees because of mismatches in authority and responsibility. Under this system, once local communities built a school, the central government was responsible for paying teachers. Because the present discounted value of teachers' salaries greatly exceeds the cost of school construction, local committees had incentives to build too many small schools. In their area of study, Kremer et al. (2003) find that the median distance from a school to the nearest neighboring school was 1.4 km.

They argue that this system of partial local control created incentives for headmasters and school committees to set fees and other attendance requirements (such as uniforms) beyond the level preferred by the median parent, even if this deterred poorer households from participating in school. Because committee members are elected on a class-by-class basis, parents of children attending later grades were over-represented. Moreover, because a major task of school committees historically was fundraising, local elites were also over-represented. They had an incentive to choose a higher level of fees than the marginal parent. Headmasters also had incentives to prefer higher school fees and costs than the median parent because extra teachers were officially assigned only when a grade's enrollment reached 55, which would have been difficult in a context of many nearby schools,

and because average scores on the national exam administered in eighth grade are often used to judge the performance of headmasters and teachers.

Kremer et al. (2003) argue that the evidence of transfers into treatment schools in the Child Sponsorship Program after the provision of free uniforms (discussed in Sections 2.1 and 3.1) suggests that parents preferred the combination of lower costs, more nonteacher inputs, and sharply higher pupil/teacher ratios associated with the program. Further evidence comes from the change in political equilibrium following the advent of multiparty democracy, which led to the abolition of school fees and surges in enrollments, consistent with education policy's move toward the preferences of the median parent rather than the median member of school committees.

Recall that learning did not suffer with average class-size increases of nine students under the Child Sponsorship Program (discussed in Sections 2.1 and 3.1). Kremer et al. (2003) suggest that with the savings from a class-size increase of less than nine, the Kenyan government could finance the textbooks, classrooms, and uniforms that were provided through the program without external funds and thus increase years of schooling by 17%.

7. SCHOOL CHOICE

A more radical option to systems reform would be to allow public and private schools to compete for students under a school-voucher program. In the average developing country in the 2002–2003 school year, 11% of all primary enrollment and 15% of all secondary enrollment were in private schools (UNESCO 2006), and there is evidence that private-school enrolment is growing (Tooley 2004, Andrabi et al. 2007, Kremer & Muralidharan 2008). Incentives may be stronger in these private schools. Private-school teachers in India, for example, are 8 percentage points less likely to be absent than civil-service teachers in the same village and 2 percentage points less likely overall (Chaudhury et al. 2006).[10]

Although findings on the effects of school choice in the United States are mixed (e.g., see Greene et al. 1997, Rouse 1998, Witte 1995, Hoxby 2000, Krueger & Zhu 2003), a study in Colombia suggests that vouchers with a merit-scholarship component can be effective in improving short- and medium-term student achievement for participants in a developing-country setting, where the absolute quality of public schooling is low. Under the Programa de Ampliacion de Cobertura de la Educacion Secundaria (PACES), Colombia awarded nearly 125,000 vouchers between 1991 and 1997 that partly covered the costs of private school for students from poor neighborhoods. Vouchers were renewable every year through grade 11, contingent upon satisfactory academic performance (grade promotion). In Bogotá and other Colombian cities where demand for the vouchers was high, applicants entered a lottery, which allowed for randomized evaluation of the program.

Angrist et al. (2002, 2006) find that lottery winners were 15 percentage points more likely to attend private school (from a baseline of 54%), 10 percentage points more likely to complete eighth grade (from a baseline of 63%), and scored 0.2 standard deviations higher on standardized tests (equivalent to a grade-level difference among seventh to ninth grade Hispanic students in the United States who had taken the same test). They also completed an additional 0.12–0.16 grades. These effects were larger for girls than for

---

[10]Private schools may be more common in areas with high public-school teacher absence because they enter areas with badly performing public schools or because private-school entry triggers the exit of influential families from public schools and thus weakens performance pressure on public schools.

boys. Winners were also 5–7 percentage points more likely to graduate from high school (from a baseline of roughly 30%) and scored higher on high school completion and college entrance exams.

Lottery winners also worked 1.2 fewer labor hours per week (from a baseline of 4.9 hours) than lottery losers. Although this might reflect an income effect for the household from the voucher itself, the renewability conditional on satisfactory performance may have induced winners to devote more time to studying. Angrist et al. (2002) analyze program impact on private expenditure and implied wage benefits and conclude that is it very cost-effective.

Although the evidence from Colombia shows that vouchers can improve both student participation and achievement for participants, there is debate on the effects for nonparticipants. Advocates argue that vouchers are beneficial even for the students that remain in public schools because school choice induces public schools to improve to retain pupils (Hoxby 2000). Most voucher skeptics are concerned that gains for participants may come at the expense of nonparticipants, through changes in sorting that worsen peer composition in public schools (for two nonrandomized studies of Chile's national voucher program that try to examine aggregate effects, see Gallego 2006, Hsieh & Urquiola 2006).

Evidence from another component of the PACES program suggests that the increase in test scores in Colombia did result at least partially from an increase in productivity rather than just increases in peer quality (Bettinger et al. 2007). In a vocational variant of the PACES program, lottery winners that used their vouchers to attend private vocational schools had peers with lower test scores and lower participation rates on college entrance exams because losers in this lottery attended public academic schools. Nevertheless, lottery winners were more likely to stay in private school, finish eighth grade, and take the college entrance exam and were less likely to repeat a grade. Winners also scored between one-third and two-thirds of a standard deviation higher than lottery losers. Moreover, other studies suggest that sorting might not be detrimental to the achievement of initially poor performing students (Duflo et al. 2008).

Another potential concern with a voucher system, particularly in countries with high ethnic diversity, could be ideological and cultural segregation if parents choose to educate their children with an ideology similar to their own (James 1986a,b; Kremer & Sarychev 2008; Pritchett 2009).

8. IMPLICATIONS FOR POLICY AND FUTURE RESEARCH

Since the early PROGRESA evaluation (Gertler & Boyce 2001, Gertler 2004, Schultz 2004, Levy 2006) and collaborations between NGOs and academics in Kenya (Kremer et al. 2003, Glewwe et al. 2009a), randomized evaluations have been used to test a much wider variety of policies than many thought possible. Working with NGOs to conduct a series of related evaluations in a comparable setting, often with cross-cutting treatment designs, generates considerable cost savings in data collection and allows cost-effectiveness comparisons across program variants (Kremer 2003). Moreover, as argued by Banerjee & Duflo (2009) in this volume, this approach forces a confrontation with reality that leads to broader insights and to the development of an iterative process of generating and testing new hypotheses based on the results of earlier work. For example, finding that textbook provision did not increase average test scores in Kenya pointed the way to a broader understanding of mismatches in curriculum and ultimately to the development and testing

of policies such as remedial education (Banerjee et al. 2007a, He et al. 2007) and tracking (Duflo et al. 2008) to address this mismatch. Because randomized evaluations give researchers far fewer degrees of freedom to choose among alternative specifications after seeing the data, one can be relatively confident about the results in a particular setting.

A key question has been the extent to which the results will generalize across settings. Because similar randomized evaluations have now been conducted in multiple environments, it is now possible to assess this issue. Results on ways to increase schooling are remarkably consistent across settings. Across a wide range of contexts, prices and subsidies have a substantial impact on school participation. Studies in Kenya and Colombia suggest that merit scholarships can provide incentives for students (and households) to increase current investment in education. School health programs in Kenya and India (Bobonis et al. 2004, Miguel & Kremer 2004) have proven cost effective, and programs in the Dominican Republic and Madagascar that informed students about earnings differences among people with different levels of education led to increased school attendance (Jensen 2007, Nguyen 2008).

There are also strong similarities in findings on ways to improve school quality, although here there are also important differences across contexts, perhaps because household behavior is similar across contexts but provider behavior is more subject to the idiosyncrasies of particular incentive systems. **Figure 3** (see color insert) summarizes the test-score gains of primary and secondary school students in some of the programs reviewed in Sections 3 through 7. Four studies in Kenya and India found no test-score impact of changes in pupil-teacher ratios (Banerjee et al. 2005, 2007a; Kremer et al. 2003; Duflo et al. 2007a). Increasing existing nonteacher inputs had limited impact in two studies. A more detailed examination of the results suggests that this may result from distortions in education systems, including curricula tailored to the elite and weak teacher incentives.

The top five evaluations in **Figure 3** show that programs to address these distortions in curriculum and pedagogy (including technology-assisted learning, remedial education, and tracking) can dramatically raise test scores at a low cost, especially for students that typically get left behind (Jamison et al. 1981, Banerjee et al. 2007a, He at al. 2007). Multiple studies suggest that changes to education systems that strengthen teacher incentives can also have a positive impact, although not all the reforms examined were successful, and even the successful ones generated smaller test-score impacts than did pedagogical changes. Evidence from nonformal education centers in India suggests that students learn more when teachers are provided strong incentives to attend school. Kenyan and Indian studies examining linking teacher pay to student test scores both find gains on incentivized tests and no impact on teacher attendance, but an improvement in extra preparatory sessions for students. However, Glewwe et al. (2008) argue that there is little evidence of underlying learning, whereas Muralidharan & Sundararaman (2008a) interpret some of their results as evidence of long-term learning. It is unclear the extent to which this reflects differences in the context, the details of the program, or measurement and interpretation.

Several studies in Kenya and India suggest that programs designed to empower local communities to hold teachers accountable have limited impact if they simply provide information to communities about the performance of students and teachers, but that when school committees are given authority to hire teachers on short-term contracts, outside of civil-service rules, the teachers they hire are cheaper and are present at school more often, and their students learn more.

The only randomized evaluation of a voucher program in the developing world has been of Colombia's PACES program. That program, which operated on both sides of the market, combining school choice with a merit-scholarship component, increased both attainment and learning.

Not only are results across settings fairly similar, but with some exceptions (e.g., Barrera-Osorio et al. 2007, Kremer et al. 2009) there is little evidence of differences in program effects across sites when a program was implemented across different areas.

Although some results from the developing world are consistent with findings from the developed world, there are also important differences. The observed benefits of technology-assisted learning and tracking are consistent with some of the evidence from the United States that shows that strictly prescribed curricula can be effective in improving learning in poorly performing schools (see Kirschner et al. 2006 for a review) and that computerized math instruction tailored to individual students' learning levels can improve test scores in larger, more heterogeneous classrooms (Barrow et al. 2008). Although reducing pupil-teacher ratios seems to have limited impact in the developing world, the one randomized evaluation in the United States finds a positive impact (Krueger 1999). The Colombian evidence suggests a much more positive impact of vouchers on test scores than the U.S. literature. This may be because the public education system is stronger in the United States, but further research would be needed to test this conjecture.

Moving to a higher level of generalization, the evidence of savings constraints, peer effects, and the lack of an attendance response to improvements in school quality suggests it will be important to supplement our models of human capital investment—possibly with insights from behavioral economics to better fit the data (Kremer & Holla 2009).

An important step forward has been the use of randomized evaluations to test structural models (e.g., see Attanasio et al. 2005, Todd & Wolpin 2006). Two important future steps in this literature would be to impose further discipline on the exercises by writing down models and sample selection rules in advance of the data's becoming available and explicitly formulating and testing models that nest standard human capital and behavioral models.

Another message to take away from this review is that the estimates from some randomized evaluations such as the textbooks and flipcharts examples can be quite different from retrospective estimates. Omitted variable bias in developing-country samples is most likely upward rather than downward because compensatory programs that provide more resources to areas with unmeasured disadvantages are less prevalent in developing countries. Recent research suggests that econometric techniques designed to address these issues, such as regression discontinuity and value-added models, can lead to seriously misleading consequences in some contexts (Andrabi et al. 2008, Rothstein 2008, Urquiola & Verhoogen 2008).[11]

It is also worth noting some important caveats for the studies reviewed here and for randomized evaluations in education more generally. First, most of the reviewed studies relate to primary education, and the challenges involved in preschool, secondary, vocational, and tertiary education may be quite different. Higher levels of education are more

---

[11]Buddelmeyer & Skoufias (2004) argue that regression discontinuity design performs well in replicating the experimental results of the PROGRESA program except in the first year of the program. One concern with this, however, is that the regression-discontinuity-design estimates match the experimental ones only when ineligible households in either treatment or comparison villages serve as the control groups in these estimates.

expensive to provide, and it is less likely that local hiring of community members with minimal training can substitute for a civil-service or formal private-sector provision.

Second, randomized evaluations typically measure the full effect of an intervention—both the direct effect on the outcome of interest (such as test scores) and indirect effects through changes in student or teacher behavior (Das et al. 2004, Glewwe & Kremer 2005). For some purposes, knowing the full effect may be desirable, but in assessing the welfare effects of an intervention, it is important to appropriately measure changes in other inputs induced by the policy, as well as to keep track of program costs and redistributive impact.

Another caveat is that randomized evaluations typically measure partial equilibrium effects. Because most randomized evaluations tend to be implemented in areas that are smaller than an entire labor market, it might be difficult to measure general equilibrium effects on wages (see Duflo 2004 for difference-in-difference estimates). However, as discussed above, it is possible to get at peer effects within smaller geographic areas or within the classroom (Lalive & Cattaneo 2006, Barrera-Osorio et al. 2007, Duflo et al. 2008, Bobonis & Finan 2009), and future randomized evaluations should be able to measure local spillovers in school-choice programs as long as markets are geographically segmented.

Third, as seen in the long-run estimates in **Figure 3**, some programs reviewed here generated impacts of different magnitudes in different years of the program or after the program ended, which suggests the importance of knowledge depreciation across time (Andrabi et al. 2008). This issue should be explored further by longer run follow-ups.

Despite these caveats, the accumulation of evidence from randomized evaluations is shedding light on how to cost-effectively help the 100 million out-of-school children of primary school age attend school and how to help those already in schools learn more.

DISCLOSURE STATEMENT

ACKNOWLEDGMENTS

LITERATURE CITED

Acemoglu D, Angrist J. 2000. *How large are human capital externalities? Evidence from compulsory schooling laws*. NBER Macroecon. Man.

Abdul Latif Jameel Poverty Action Lab. 2005. Education: meeting the millenium development goals. *Fight. Poverty: What Works* 2005(1):1–4

Abdul Latif Jameel Poverty Action Lab. 2008. Solving absenteeism, raising test scores. *Policy Brief-case* 6(Sept.):1–4

Andrabi T, Das J, Khwaja AI. 2007. *Students today, teachers tomorrow? Identifying constraints on the provision of education*. Work. Pap., Harvard Univ.

Andrabi T, Das J, Khwaja AI, Zajonc T. 2008. *Do value-added estimates add value?: accounting for learning dynamics*. Work. Pap., Harvard Univ.

Angrist J, Bettinger E, Bloom E, King E, Kremer M. 2002. Vouchers for private schooling in Colombia: evidence from a randomized natural experiment. *Am. Econ. Rev.* 92:1535–58

Angrist J, Bettinger E, Kremer M. 2006. Long-term consequences of secondary school vouchers: evidence from administrative records in Colombia. *Am. Econ. Rev.* 96:847–62

Angrist J, Lavy V. 1999. Using Maimonides' rule to estimate the effect of class size on children's academic achievement. *Q. J. Econ.* 114:533–76

Angrist J, Lavy V. 2002. New evidence on classroom computers and pupil learning. *Econ. J.* 112: 735–65

Attanasio O, Meghir C, Santiago A. 2005. *Education choices in Mexico: using a structural model and a randomized experiment to evaluate PROGRESA*. Work. Pap, London Inst. Fiscal Stud., Centre Eval. Dev. Policies

Banerjee A, Banerji R, Duflo E, Glennerster R, Khemani S. 2006. *Can informational campaigns spark local participation and improve outcomes? A study of primary education in Uttar Pradesh, India.* World Bank Policy Res. Work. Pap. 3967

Banerjee A, Banerji R, Duflo E, Glennerster R, Khemani S. 2009. Pitfalls of participatory programs: evidence from a randomized evaluation in education in India. *Am. Econ. J. Econ. Policy.* In press

Banerjee A, Cole S, Duflo E, Linden L. 2007a. Remedying education: evidence from two randomized experiments in India. *Q. J. Econ.* 122:1235–64

Banerjee A, Duflo E. 2006. Addressing absence. *J. Econ. Perspect.* 20:117–32

Banerjee A, Duflo E. 2009. The experimental approach to development economics. *Annu. Rev. Econ.* 1:In press

Banerjee A, Duflo E, Glennerster R. 2007b. Putting a band-aid on a corpse: incentives for nurses in the Indian public health care system. *J. Eur. Econ. Assoc.* 6:487–500

Banerjee A, Jacob S, Kremer M, Lanjouw J, Lanjouw P. 2005. *Promoting school participation in rural Rajasthan: results from some prospective trials*. Work. Pap., MIT

Barrera-Osorio F, Bertrand M, Linden L, Perez-Calle F. 2007. *Using conditional transfers in education to investigate intra family decisions: evidence from a randomized experiment*. Work. Pap., Univ. Chicago Graduate School Bus.

Barro R. 1997. *Determinants of Economic Growth*. Cambridge, MA: MIT Press

Barro R. 1999. Determinants of democracy. *J. Polit. Econ.* 107:158–83

Barro RJ, Lee J-W. 2001. International data on educational attainment: updates and implications. *Oxf. Econ. Pap.* 53:541–63

Barrow L, Markman L, Rouse C. 2008. *Technology's edge: the educational benefits of computer-aided instruction*. NBER Work. Pap. No. 14240

Behrman JR, Sengupta P, Todd P. 2005. Progressing through PROGRESA: an impact assessment of a school subsidy experiment in rural Mexico. *Econ. Dev. Cult. Change* 54:237–75

Berry J. 2008. *Child control in education decisions: an evaluation of targeted incentives to learn in India*. Work. Pap., MIT

Bettinger E, Kremer M, Saavedra J. 2007. *Are educational vouchers only redistributive?* MIT Abdul Latif Jameel Poverty Action Lab Work. Pap.

Bils M, Klenow P. 2000. Does schooling cause growth? *Am. Econ. Rev.* 90:1160–83

Bjorkman M, Svensson J. 2009. Power to the people: evidence from a randomized experiment on community-based monitoring in Uganda. *Q. J. Econ.* In press

Bleakley H. 2007a. Disease and development: evidence from hookworm eradication in the American South. *Q. J. Econ.* 122:73–117

Bleakley H. 2007b. *Malaria eradication in the Americas: a retrospective analysis of childhood exposure*. Work. Pap., Univ. Chicago

Bobonis G, Finan F. 2009. Neighborhood peer effects in secondary school enrollment decisions. *Rev. Econ. Stat.* In press

Bobonis G, Miguel E, Sharma CP. 2004. Anemia and school participation. *J. Hum. Resour.* 41:692–721

Bruns B, Mingat A, Rakotomalala R. 2003. *Achieving Universal Primary Education by 2015: A Chance for Every Child.* Washington, DC: World Bank

Buddelmeyer H, Skoufias E. 2004. *An evaluation of the performance of regression discontinuity design on PROGRESA.* World Bank Policy Res. Work. Pap. 3386, Washington, DC

Campuzano L, Dynarksi M, Agodini R, Rall K. 2009. Effectiveness of reading and mathematics software products: findings from two student cohorts. *NCEE 2009-4041*, Natl. Center Educ. Eval. Reg. Assist., Inst. Educ. Sci., U.S. Dept. Educ., Washington, D.C.

Chaudhury N, Hammer J, Kremer M, Muralidharan K, Rogers H. 2006. Missing in action: teacher and health worker absence in developing countries. *J. Econ. Perspect.* 20:91–116

Das J, Dercon S, Habyarimana J, Krishnan P. 2004. *When can school inputs improve test scores?* Policy Res. Work. Pap. 3217, World Bank, Washington, D.C.

de Laat J, Kremer M, Vermeersch C. 2008. *Teacher incentives and local participation.* Work. Pap., Harvard Univ.

Duflo E. 2001. Schooling and labor market consequences of school construction in Indonesia: evidence from an unusual policy experiment. *Am. Econ. Rev.* 91:795–814

Duflo E. 2004. The medium run consequences of educational expansion: evidence from a large school construction program in Indonesia. *J. Dev. Econ.* 74:163–97

Duflo E, Dupas P, Kremer M. 2007a. *Peer effects, pupil-teacher ratios, and teacher incentives.* Work. Pap., MIT

Duflo E, Dupas P, Kremer M. 2008. *Peer effects and the impact of tracking: evidence from a randomized evaluation in Kenya.* Work. Pap., MIT

Duflo E, Dupas P, Kremer M, Sinei S. 2006. *Education and HIV/AIDS prevention: evidence from a randomized evaluation in Western Kenya.* Work. Pap., MIT

Duflo E, Hanna R, Ryan S. 2007b. *Monitoring works: getting teachers to come to school.* NBER Work. Pap. No. 11880

Evans D, Kremer M, Ngatia M. 2008. *The impact of distributing school uniforms on children's education in Kenya.* Work. Pap., Harvard Univ.

Filmer D, Hasan A, Pritchett L. 2006. *A millenium learning goal: measuring real progress in education.* Center Glob. Dev. Work. Pap. 97

Filmer D, Pritchett L. 1999. What education production functions really show: a positive theory of education expenditures. *Econ. Educ. Rev.* 18:223–39

Fiszbein A, Schady N. 2009. Conditonal cash transfers: reducing present and future poverty. *World Bank Policy Res. Rep.*, Washington, DC

Foster A, Rosenzweig M. 1996. Technical change and human capital returns and investments: evidence from the green revolution. *Am. Econ. Rev.* 86:931–53

Gallego F. 2006. *Voucher-school competition, incentives, and outcomes: evidence from Chile.* Work. Pap., Pontificia Univ. Catolica Chile

Gertler P. 2004. Do conditional cash transfers improve child health? Evidence from PROGRESA's control randomized experiment. *Am. Econ. Rev. Pap. Proc.* 94:336–41

Gertler P, Boyce S. 2001. *An experiment in incentive-based welfare: the impact of PROGRESA on health in Mexico.* Work. Pap., Univ. Calif., Berkeley

Glaeser E, LaPorta R, Lopez-de-Silanes F, Shleifer A. 2004. Do institutions cause growth? *J. Econ. Growth* 9:271–303

Glewwe P. 1999. *The Economics of School Quality Investments in Developing Countries: An Empirical Study of Ghana.* New York: St. Martin's Press

Glewwe P, Kremer M. 2005. Schools, teachers, and education outcomes in developing countries. *Handbook of Economics of Education 1*, ed. E Hanushek, F Welch, pp. 945–1017. Amsterdam: Elsevier

Glewwe P, Kremer M, Moulin S. 2009a. Many children left behind? Textbooks and test scores in Kenya. *Am. Econ. J. Appl. Econ.* 1:112–35

Glewwe P, Kremer M, Moulin S, Zitzewitz E. 2004. Retrospective vs. prospective analyses of school inputs: the case of flipcharts in Kenya. *J. Dev. Econ.* 74:251–68

Glewwe P, Holla A, Kremer M. 2009b. Teacher incentives in the developing world. In *Performance Incentives: Their Growing Impact on American K-12 Education*, ed. M Springer. Washington, DC: Brookings Inst. Press

Glewwe P, Ilias N, Kremer M. 2008. *Teacher incentives*. Work. Pap., Harvard Univ.

Glewwe P, Olinto P. 2004. Evaluating the impact of conditional cash transfers on schooling: an experimental analysis of Honduras PRAF program. *Final Rep. USAID, Int. Food Policy Res. Inst.*

Gonzales PJCG, Partelow L, Pahlke E, Jocelyn L, Kastberg D, Williams T. 2004. Highlights from the Trends in International Mathematics and Science Study (TIMSS) 2003. *NCES 2005–005, U.S. Dept. Educ., Natl. Center Educ. Stat.*, Washington, D.C.

Gordon R, Kane TJ, Staiger DO. 2006. Identifying effective teachers using performance on the job. *Hamilton Project Discuss. Pap.* 2006-01, Brookings Inst.

Greaney V, Khandker SR, Alam M. 1999. *Bangladesh: Assessing Basic Learning Skills*. Dhaka, Bangladesh: Univ. Press

Greene JP, Peterson PE, Du J. 1997. Effectiveness of school choice: the Milwaukee experiment. *Harvard Univ. Prog. Educ. Policy Gov.*, Cambridge, Mass.

Hannaway J. 1992. Higher order thinking, job design, and incentives: an analysis and proposal. *Am. Educ. Res. J.* 29:3–21

Hanushek EA. 1995. Interpreting recent research on schooling in developing countries. *World Bank Res. Obs.* 10:227–46

Hanushek EA. 2002. Evidence, politics, and the class size debate. See Mishel & Rothstein 2002, pp. 39–66

Hanushek EA, Woessmann L. 2008. The role of cognitive skills in economic development. *J. Econ. Lit.* 46:607–68

Harbison RW, Hanushek EA. 1992. *Educational Performance of the Poor: Lessons from Rural Northeast Brazil*. New York: Oxford Univ. Press

He F, Linden L, MacLeod M. 2007. *Helping teach what teachers don't know: an assessment of the Pratham English language program*. Work. Pap., Columbia Univ.

Holmstrom B, Milgrom P. 1991. Multi-task principal-agent analysis: incentive contracts, asset ownership, and job design. *J. Law Econ. Organ.* 7:24–52

Hoxby C. 2000. Does competition among public schools benefit students and tax payers? *Am. Econ. Rev.* 90:1209–38

Hsieh C-T, Urquiola M. 2006. The effects of generalized school choice on achievement and stratification: evidence from Chile's school voucher program. *J. Public Econ.* 90:1477–503

James E. 1986a. Benefits and costs of privatized public services: lessons from the Dutch educational system. In *Private Education: Studies in Choice and Public Policy*, ed. D Levy, pp. 113–38. New York: Oxford Univ. Press

James E. 1986b. The private nonprofit provision of education: a theoretical model and application to Japan. *J. Comp. Econ.* 10:255–76

Jamison D, Searle B, Galda K, Heyneman S. 1981. Improving elementary mathematics education in Nicaragua: an experimental study of the impact of textbooks and radio on achievement. *J. Educ. Psychol.* 73:556–67

Jensen R. 2007. *The perceived returns to education and the demand for schooling*. Work. Pap., Harvard Univ.

Jorgenson DW, Fraumeni BM. 1989. The accumulation of human and nonhuman capital, 1948–1984. In *The Measurement of Saving, Investment, and Wealth*, ed. RE Lipsey, HS Tice, pp. 227–86. Chicago: Univ. Chicago Press

Kirschner PA, Sweller J, Clark RE. 2006. Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41:75–86

Kouskalis E. 2008. *Does not compute: the introduction of new technologies to South African and Namibian classrooms*. Undergraduate thesis. Harvard Univ.

Kremer M. 2003. Randomized evaluation of educational programs in developing countries: some lessons. *Am. Econ. Rev. Pap. Proc.* 93(2):102–6

Kremer M, Chen D. 2001. *Interim report on a teacher incentive program in Kenya*. Work. Pap., Harvard Univ.

Kremer M, Holla A. 2008. *Increasing the quantity and quality of human capital investment in the developing world*. World Bank Comm. Growth Dev. Work. Pap.

Kremer M, Holla A. 2009. Pricing and access: lessons from randomized evaluations in education and health. In *What Works in Development: Thinking Big and Thinking Small*, ed. W Easterly, J Cohen. Washington, D.C.: Brookings Inst.

Kremer M, Miguel E, Thornton R. 2009. Incentives to learn. *Rev. Econ. Stat.* In press

Kremer M, Moulin S, Namunyu R. 2003. *Decentralization: a cautionary tale*. Work. Pap., Harvard Univ.

Kremer M, Muralidharan K. 2008. Public and private schools in rural India. In *School Choice International*, ed. P Peterson, R Chakrabarti, pp. 91–110. Cambridge, MA: MIT Press

Kremer M, Sarychev A. 2008. *Why do governments operate schools?* Work. Pap., London School Econ.

Kremer M, Vermeersch C. 2005. *School meals, educational achievement and school competition: evidence from a randomized evaluation*. World Bank Policy Res. Work. Pap. 3523

Krueger A. 1999. Experimental estimates of education production functions. *Q. J. Econ.* 114:497–532

Krueger A, Lindahl M. 2001. Education for growth: why and for whom. *J. Econ. Lit.* 39:1101–36

Krueger A, Whitmore DM. 2001. The effect of attending a small class in the early grades on the college-test taking and middle school test results: evidence from Project STAR. *Econ. J.* 111:1–28

Krueger AB. 2002. Understanding the magnitude and effect of class size on student achievement. See Michel & Rothstein 2002, pp. 7–36

Krueger AB. 2002. A response to Eric Hanushek's "Evidence, politics, and the class size debate." See Michel & Rothstein 2002, pp. 73–88

Krueger A, Zhu P. 2003. *Another look at the New York City school voucher experiment*. NBER Work. Pap. 9418

Lalive R, Cattaneo A. 2006. Social interactions and schooling decisions. *IZA Discuss. Pap. 2250*, Inst. Study Labor, IZA

Levy S. 2006. *Progress Against Poverty: Sustaining Mexico's Progresa-Oportunidades Program*. Washington, DC: Brookings Inst.

Lockheed M, Verspoor A. 1991. *Improving Primary Education in Developing Countries*. New York: Oxford Univ. Press

Lockheed ME, Hanushek E. 1988. Improving educational efficiency in developing countries: What do we know? *Compare* 18:21–37

Lucas R. 1988. On the mechanics of economic development. *J. Monetary Econ.* 22:3–42

Maluccio JA, Flores R. 2005. *Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social*. Washington, D.C.: Int. Food Policy Res. Inst.

Miguel E, Kremer M. 2004. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 72:159–217

Mishel L, Rothstein R, eds. 2002. *The Class Size Debate*. Washington, DC: Econ. Policy Res.

Muralidharan K, Sundararaman V. 2008a. *Teacher performance pay: experimental evidence from India*. Work. Pap., Univ. Calif, San Diego

Muralidharan K, Sundararaman V. 2008b. *Contract teachers: experimental evidence from India*. Work. Pap., Univ. Calif., San Diego

Nguyen T. 2008. *Information, role models, and perceived returns to education: experimental evidence from Madagascar*. Work. Pap., MIT

Powell CA, Walker SP, Chang SM, Grantham-McGregor SM. 1998. Nutrition and education: a randomized trial of the effects of breakfast in rural primary school children. *Am. J. Clin. Nutr.* 68:873–79

Pritchett L. 2001. Where has all the education gone? *World Bank Econ. Rev.* 15(3):367–91

Pritchett L. 2009. The policy irrelevance of the economics of education: Is "normative as positive" just useless, or is it worse? In *What Works in Development: Thinking Big and Thinking Small*, ed. W Easterly, J Cohen, Brookings Inst.

PROBE Team. 1999. *Public Report on Basic Education in India*. New Delhi: Oxford Univ. Press

Psacharopoulos G, Patrinos HA. 2004. Returns to investment in education: a further update. *Educ. Econ.* 12:111–34

Rice JK. 2002. Making the evidence matter: implications of the class size research debate for policy makers. See Mishel & Rothstein 2002, pp. 90–95

Rothstein, Jesse 2008. *Teacher quality in educational production: tracking, decay, and student achievement*. Work. Pap., Princeton Univ.

Rouse CE. 1998. Private school vouchers and student achievement: an evaluation of the Milwaukee parental choice program. *Q. J. Econ.* 13:553–602

Schady N, Araujo MC. 2006. *Cash transfers, conditions, school enrollment, and child work: evidence from a randomized experiment in Ecuador.* World Bank Policy Res. Work. Pap. 3930

Schultz P. 1997. Demand for children in low income countries. In *Handbook of Population and Family Economics*, ed. M Rosenzweig, O. Stark, pp.349–430. Amsterdam: North Holland

Schultz TP. 2002. Why governments should invest more to educate girls. *World Dev.* 30:207–25

Schultz TP. 2004. School subsidies for the poor: evaluating the Mexican PROGRESA poverty program. *J. Dev. Econ.* 74:199–250

Shastry GK, Linden L. 2007. *Identifying agent discretion: exaggerating student attendance in response to a conditional school nutrition program*. Work. Pap., Columbia Univ.

Strauss J, Thomas D. 1995. Human resources. In *Handbook of Development Economics*: Vol. 3, ed. J Behrman, TN Srinivasan, pp. 1883–2023. Amsterdam: North Holland

Todd P, Wolpin KI. 2006. Using experimental data to validate a dynamic behavioral model of child schooling: assessing the impact of a school subsidy program in Mexico. *Am. Econ. Rev.* 96:1384–417

Tooley J. 2004. Private education and education for all. *Econ. Aff.* 24(4):4–7

Umansky I, Vegas E. 2007. Inside decentralization: how three Central American school-based management reforms affect student learning through teacher incentives. *World Bank Res. Obs.* 22:197–215

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

UNESCO. 2006. *Education for All Global Monitoring Report: Literacy for Life*. Paris: UNESCO

Urquiola M, Verhoogen E. 2008. Class-size caps, sorting, and regression: discontinuity design. *Am. Econ. Rev* 99:179–215

Vegas E, de Laat J. 2005. Do differences in teacher contracts affect student performance? *World Dev. Rep. 2004 Backgr. Pap.*

WHO. 2005. *Deworming for Health and Development: Report of the Third Global Meeting of the Partners for Parasite Control*. Geneva: WHO

Witte JF. 1995. Three critical factors in the school choice debate. *Soc. Sci. Q.* 76:502–6

World Bank. 2004. *World Development Report 2004: Making Services Work for Poor People*. Washington, DC: World Bank

World Bank. 2006. *World Development Indicators* 2006. Washington, DC: World Bank

Zhang Y, Postlethwaite TN, Grisay A, eds. 2008. *A View Inside Primary Schools: A World Education Indicators Cross-National Study*. Montreal: UNESCO
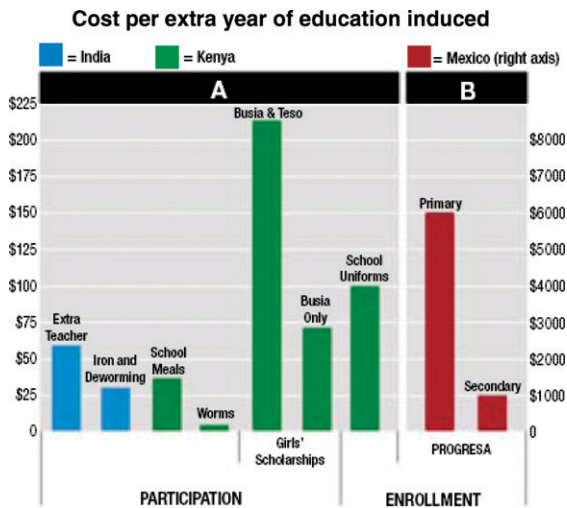
Figure 1

The cost-effectiveness of select interventions. Figure taken from Abdul Latif Jameel Poverty Action Lab (2005), "Education: meeting the millennium development goals," *Fighting Poverty: What Works*, Issue 1.
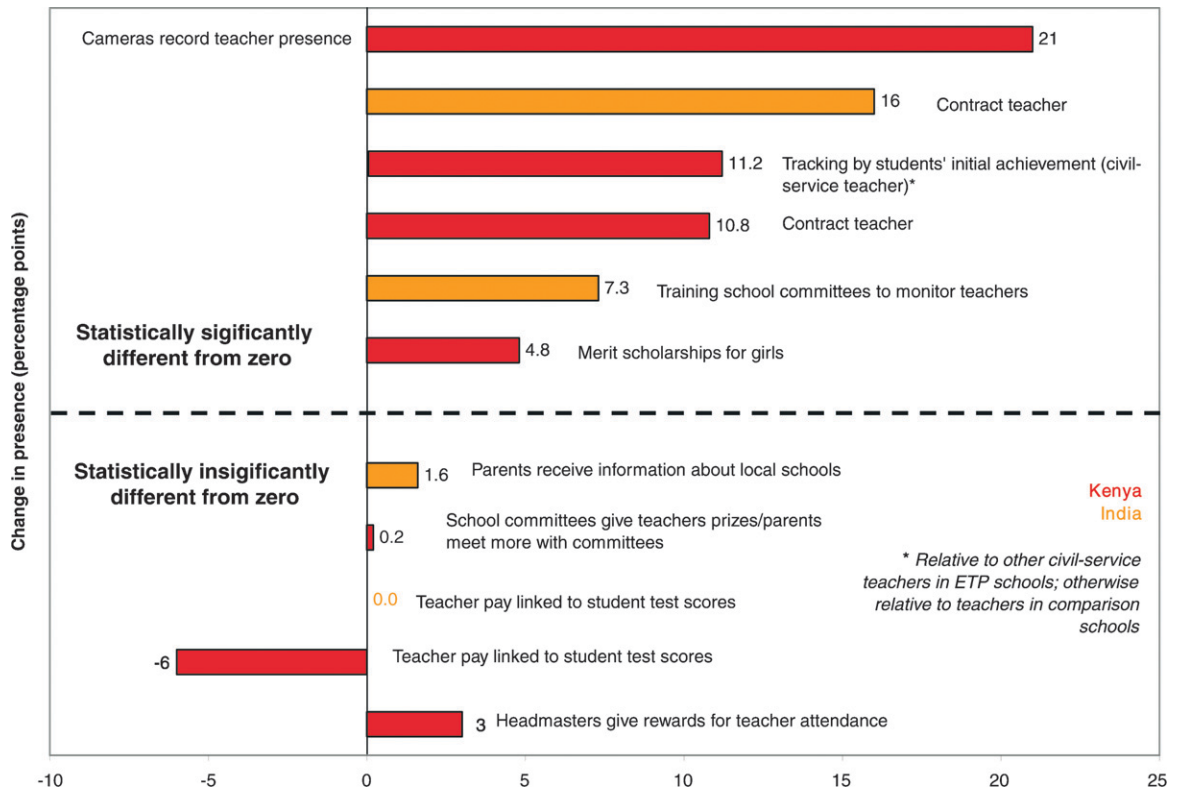
Figure 2

Increases in teacher presence, by program. The largest estimated effect is shown when there is a range of estimates. The figure illustrates the probability of a teacher being in class and teaching a random school day for the Extra Teacher Program (ETP) in Kenya and the probability of a teacher being in school for all other programs.
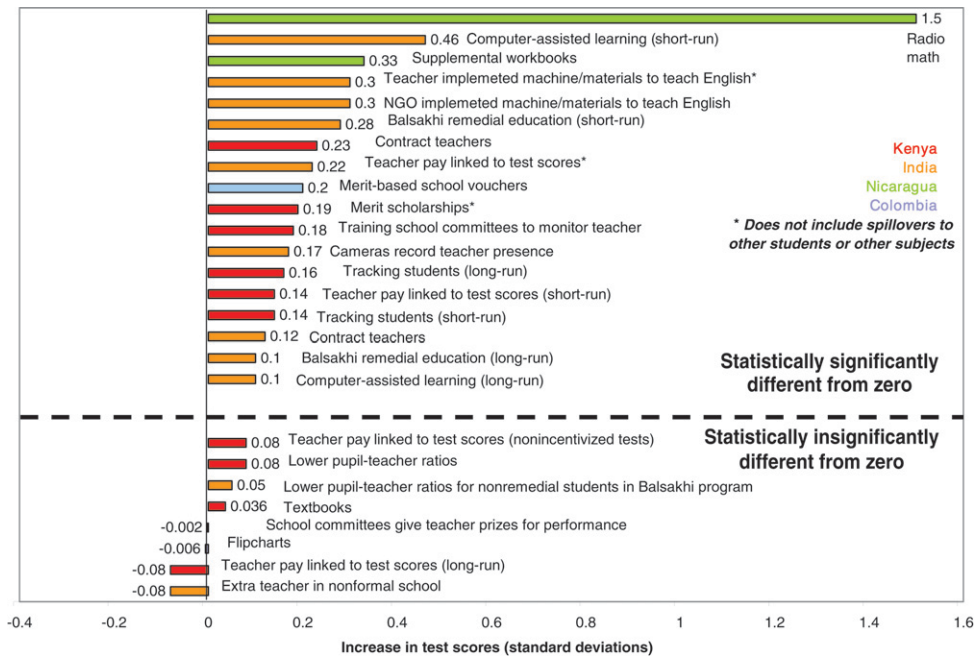
Figure 3

Increases in test scores, by program. The largest estimated effect is shown when there is a range of estimates. Effects measured after a program has been discontinued are labeled as long-run. Otherwise, they have been measured while the program was still in place. Tests were identical in the following program pairings: (1) lower pupil-teacher ratios in Kenya, tracking, training school committees to monitor teachers, and contract teachers in Kenya; (2) lower pupil-teacher ratios and more homogeneous peers in India, Balsakhi remedial education, and computer-assisted learning; (3) teacher pay linked to test scores in India and contract teachers in India. Reported test-score gains have been measured with the following lags after program start: 1 year for radio math, supplemental workbooks, teacher implemented machine/materials to teach English, nongovernmental organization (NGO) implemented machine/materials to teach English, cameras record teacher presence, and merit scholarships; 18 months for group (1) above and extra teacher in nonformal school; 2 years for groups (2) and (3), teacher pay linked to test scores in Kenya, and headmasters give prizes for teacher attendance; 3 years for merit-based school vouchers; 1–2 years for flipcharts; 1–3 years for school committees give prizes for performance, and 1–4 years for textbooks. This figure does not include achievement gains measured by graduation rates or the attainment of certain levels of reading proficiency.