18

The Primacy of Education

Anne Case

EDUCATION, HUMAN CAPITAL, AND HUMAN CAPABILITIES

Many economists, pressed to list the keys to economic development, would turn first to education. Beliefs regarding the primacy of education in the development process stem both from the fundamental role of education in income generation and from the many other ways in which education is thought to promote and sustain development and, in turn, to enhance quality of life.

The first section of this essay highlights what we know about the role of education in three interrelated areas: income generation, health status, and fertility. Perhaps surprisingly, given the great amount of attention that the subject has enjoyed, estimates of the extent to which education causes earnings to rise, or promotes better health and longer life, remain the subject of lively debate in economics. We discuss reasons why measuring the impact of education is difficult, and present the best current estimates we have for the effects of education in these important areas.

In bringing education to people in the developing world, an understanding of which school inputs are most effective, and at what levels of education, is paramount. The second half of the essay discusses why we know very little about which inputs are most effective, and contrasts different approaches to assessing the impact of school inputs. We end with a discussion of programs that promise to be more successful in helping governments and policy makers decide how best to invest in human capital.

Education and Income Generation

In both developing and developed countries, better-educated workers earn higher wages on average than do less well-educated ones. This may be true for a host of reasons, which complicates attempts to quantify the causal effects of schooling on earnings. In developing countries, wealthier families can afford to educate their children, and can aid them in finding superior jobs. Untangling which part of earnings is due to education, and which part to (say) the quality of family connections, is often far from straightforward. Years of completed schooling may also reflect a person's abilities, and those who are more able would be expected to earn higher incomes, regardless of schooling. Further complications arise because schooling is of uneven quality, so that, among children in South Africa, for example, having completed six years of schooling in rural KwaZulu-Natal may be very different from completing the same number of years in Durban. More broadly, this last point highlights the problems caused by the rather noisy measures of schooling we generally have available for analysis: not only does schooling vary in quality, but people misremember the number of years they have obtained, and (perhaps more often) their schooling is misremembered by the "knowledgeable household member" chosen to act as the family's informant in many household surveys.

Since the 1970s, much has been learned about the relationship between education and earnings, largely due to marked improvements in the quantity and quality of data collected, and to the attention paid to the measurement issues raised above by both economic theorists and applied economists.[1] Several strategies have been suggested to quantify the impact of education on earnings. For example, some researchers have looked at differences in earnings between siblings and have correlated these with differences in their educations. To the extent that siblings have the same access to family resources and contacts, the differences in earnings between them may be attributable to the differences in their educational attainment. While such a strategy may succeed in neutralizing the role of families, it doesn't explain why siblings differ in their years of schooling. If these differences are due to differences in ability, then this sort of strategy could end up magnifying the bias caused by unobserved ability.

A second tack in analyzing returns to schooling has been to focus on those differences in years of completed schooling that can be attributed to institutional differences, between places or over time. Differences in proximity to a school, enforced minimum school leaving ages, or restrictions on child labor may lead some children to attend school longer. Differences in completed schooling that can be attributed to such institutional arrangements may meet the statistical requirements necessary for them to be useful in estimating returns to education. For example, in recent work Duflo (2001) analyzes the impact of a massive school-building program in Indonesia where, between 1973 and 1978, 61,000 primary schools were built, targeted educated workers earn ›nes. This may be true to quantify the causal tries, wealthier families em in finding superior ıcation, and which part r from straightforward. on's abilities, and those · incomes, regardless of ng is of uneven quality, :, having completed six ry different from com- broadly, this last point easures of schooling we hooling vary in quality, ave obtained, and (per- by the "knowledgeable ›rmant in many house-

ıe relationship between vements in the quantity iid to the measurement ›plied economists.[1] Sev- mpact of education on l at differences in earn- differences in their ed- cess to family resources em may be attributable hile such a strategy may ı't explain why siblings s are due to differences gnifying the bias caused

; has been to focus on ıat can be attributed to ıe. Differences in prox- ages, or restrictions on l longer. Differences in stitutional arrangements r them to be useful in cent work Duflo (2001) program in Indonesia ›ols were built, targeted

in those areas in which children were least likely to have been enrolled prior to the building program. Children young enough to benefit from the new schools, who were living in areas targeted for school building, completed more years of schooling on average. Observing these children as working adults, Duflo estimates an economic return to an extra year of schooling of roughly 10%.

This work broadly confirms the findings of Psacharopoulos (1994), who provides a comprehensive set of estimates on the profitability of investments in education around the world. Psacharopoulos concludes that primary schooling remains "the number one investment priority" for developing countries. He also finds the return to an additional year of education is marginally higher for girls (increasing earnings by 12.4% on average) than for boys (11.1%), and that the returns to education follow the same rules as other sorts of investments, declining as the investment is expanded.

In regions in which there are payoffs to learning—due to the introduction of new technologies or to changes in market conditions, for example—investment in education can yield a large return. However, in order for returns to education to be positive, either there must be economic opportunities that take advantage of the skills embodied in education, or investment in education must induce innovation. Rosenzweig (1995) makes this point through an example of differences in the returns to primary education between regions of India during the green revolution. In those areas agroclimatically suited to the use of the new higher-yielding variety seeds, returns to primary schooling rose: those farmers who had been to school were apparently more skillful in adopting the new seeds. However, those areas unsuited to the new seeds saw no change in the returns to primary schooling.

Recognition that better-educated workers need opportunities, if incomes are to be enhanced by schooling, is especially important when large changes in education policy are under consideration—changes, for example, such as the enforcement of compulsory schooling laws or extensive school-building programs. Duflo (2004), in a follow-up to her original work on the Indonesian school-building program, finds that while the program led to large increases in the proportion of primary school graduates in the labor force in those parts of the country that undertook the largest building efforts, this resulted in slower wage growth in the earnings of older workers—a result that would be expected if (for whatever reason) physical capital wasn't increased in response to the increase in human capital. It may be the need for different vintages of human capital, rather than physical capital, that holds down returns to education. Kremer and Thomson (1998) posit that older and younger workers are imperfect substitutes, having comparative advantages in different, complementary tasks, which may explain why many African countries witnessed marked increases in educational attainment but little improvement in economic growth.

Countries have also been observed reaping the benefits of earlier investments in education, but only after economic reforms are introduced. Drèze and Sen (2002) contrast the differences in economic performance in China and India before and after market reforms enacted in China in 1979. China had invested much more heavily in education than had India through the 1970s, with the result that literacy rates among adults in China (51% for women, and 79% for men) were markedly higher than those in India (26% and 55% percent, respectively) by the early 1980s. However, Drèze and Sen note that until the Chinese market reforms of 1979, India and China had similar economic growth rates. It took market reform, in combination with higher literacy rates in China, to sustain rapid economic expansion there in the 1980s and 1990s.

Moreover, once there is tangible evidence that human capital investments yield a handsome return, couples may choose to limit the number of children they raise, in order to offer each child more education. Lucas (2003) argues that the phenomenon of children leaving the family farm, where all necessary skills are acquirable through on-the-job training, for work that requires additional skills learned at school, where there is a substantial return to this investment, may result in fertility declines. In this way, improved opportunities—brought on by technological change or the opening of markets—blazes a trail from fertility reduction and increased educational attainment to sustained economic growth.

Education and Health Status

In both developed and developing countries, a strong correlation exists between schooling and good health, whether measured using mortality rates, morbidity rates, or self-reported health status. Each additional year of schooling for men in the United States is associated with an 8% reduction in mortality, a result consistent with those found in many European countries.[2] In surveys run in both the developed and the developing world, people with greater levels of schooling report themselves to be significantly healthier (Case 2002).

There are many ways in which education may affect health. Not only does education lead to higher income and to less risky choices of occupation, but education also increases people's understanding of sanitation and hygiene, improves their ability to read labels of all sorts, encourages their use of health care systems, and, in countless other ways, acts to protect and promote their health (Caldwell 1986).

Health disparities between better- and less well-educated people often increase when a new health technology is introduced. If better-educated people understand the importance of a health innovation more quickly, or are able to change their behaviors more rapidly to take advantage of health advances, we would expect to see differences in health status between better- and less well-educated people widen, at least in the medium run, until those with less education are able to catch up with the new technologies. Preston (1996) presents evidence on this, using data from U.S. censuses conducted around the beginning of the twentieth century. He argues convincingly that until the germ theory of disease was advanced in the late 1800s, the mortality rate for the children of schoolteachers was no better than the national average. However, by the early 1920s, when knowledge of how to protect against germs was spreading, the mortality of teachers' children fell to 40% below the national average.

As was true of income, there are many other reasons why a correlation might exist between education and health, which makes it difficult to quantify the impact of education in this domain. Healthier people may be better able to succeed in the classroom, leading to a channel from better health to additional schooling. Healthier people may anticipate a longer life, and thus greater lifetime returns to education, which may also lead them to choose more years of schooling. (The extent to which the shadow of HIV/AIDS will affect schooling decisions in the developing world is not yet well understood, but it seems likely that the specter of premature death will influence investment choices.) In addition, there may be determinants of health and education that lead to a spurious correlation between them. People with more self-control, for example, may invest more in both their health and their education. Evidence along these lines comes from creative work showing that although there is a strong, negative relationship between smoking and completed schooling at age twenty-four, this can be explained by differences in smoking behavior at age seventeen, when all of the individuals under study were still in the same grade (Farrell and Fuchs 1982). Since future schooling cannot reach back and cause smoking at age seventeen, the association between smoking and education is more likely attributable to third factors that drive them both.

New evidence on the causal effect of education on longevity comes from innovative work that exploits differences in years of schooling that grew out of changes in compulsory schooling laws in the United States. Lleras-Muney (2005) finds that differences between states in these laws in the early part of the twentieth century are strong predictors of completed education for individuals raised in the 1920s and 1930s, and she uses the variation in years of completed education that can be attributed to differences between states in their compulsory schooling laws to identify the impact of schooling on mortality. She finds that, in 1960, an additional year of education increased life expectancy at age thirty-five by as much as 1.7 years. This work is the most convincing work to date on the causal impact of education on health, since compulsory laws are not expected to influence health outcomes except through their impact on years of completed education.

Education and Fertility

There are many reasons why we would expect increased education to have a causal effect on fertility, and many researchers have documented the close articulation between education and fertility decline. Complementary to the arguments made by Lucas—that enhanced opportunities for educated workers act as a catalyst for fertility transition—Caldwell (1982) hypothesizes that, in the developing world, schools serve to advance the values of the Western middle class, leading to a restructuring of family relationships and a reversal in the flow of household resources (in favor of children). Prior to the onset of mass education, children worked inside and outside the house, doing chores and contributing time and money to the household resource base. Once children are in school, not only do they have less time for work, but their status as students also tends to lower the household's expectations about their work. And, given that the education of children has a public good component, society also invests in children's educations, raising expectations generally that children's families will protect that investment. Thus, while Lucas emphasizes the increased return to investment in education as leading to fewer children, Caldwell focuses on the increased costs of raising a child and diminished expectations of the lifetime return to parents from that child—both consequences of schooling—as setting off a fertility transition.

Other researchers have focused on the relationship between women's education and fertility decline, arguing that a woman's education reduces her desired family size, changes the relationship between her desired number of children and planned number of births, and improves her ability to achieve her desired family size.[3] Education increases the opportunity cost of women's time, because the skills learned at school find a return in the marketplace. Better-educated women may have higher aspirations for their children, which may cause them to weigh "quality" more heavily in a "quality vs. quantity" trade-off with regard to their children. Declines in fertility and infant mortality move hand in hand, and women's education may also have an indirect effect on fertility through the role it plays in reducing infant mortality. Better-educated women are more likely to know about hygiene and nutrition, and are more likely to act on this knowledge.[4] Education is apt to give women more voice in household decisions, allowing them to stand up to men in general. If women are the protectors of the needs of small children, then children are apt to benefit indirectly in this way from mothers' schooling.

These arguments on the impact of mass education in reversing intergenerational resource flows, and of women's education in reducing total fertility, are all sensible. Unfortunately, it is difficult to find evidence that education has a causal effect on fertility, evidence that would allow us to reject the view that the association between increased education and reduced total fertility rates is due to some third factor. Girls in developing countries who are educated beyond primary school may be a highly motivated, very select group, who may have lower total fertility for other reasons.[5] It may not be women's education per se that causes fertility to decline, but that educated women are more likely to marry educated men, and these men may have strong preferences for lower fertility. Young women who have had children may find it difficult to return to school—because of the demands placed on them at home and because many schools discriminate against young mothers returning to school. All of these factors would lead us to find a connection between women's education and fertility but not one that was causal.

Even keeping these third factors in mind, some researchers argue that the evidence supports a causal impact of education on fertility. Caldwell (1982) notes that fertility declines in the countries of nineteenth-century Europe followed immediately after increases in mass education in these countries and that, within a decade of the introduction of compulsory schooling, fertility was declining for all occupational groups. However, this is far from settled territory. It may be that fertility declined as a result of the mortality decline that Europe witnessed in the late nineteenth century, and the mortality declines were responsible for increased schooling. This, then, would put the timing of schooling increases and fertility declines in close proximity.

A second piece of evidence comes from India. Drèze and Murthi (2000) find that women's education is the most important correlate of fertility decline, both across districts and within districts over time. Because they are following districts from decade to decade, these researchers can estimate the impact of education on fertility solely using differences in these variables over time within each district. Doing so allows them to rule out some of the "third factor" explanations for the relationship between education and fertility by eliminating differences between districts that remain fixed over time. Drèze and Murthi find large effects of women's education: a ten percentage point increase in female literacy is associated with an expected decline in the total number of children born to women during their lifetimes of 0.2 child. To understand the magnitude of this estimate, it is interesting to compare it with the impact of religion on fertility: a ten percentage point increase in the proportion of the population in the district that is Muslim is associated with an increase in total fertility of 0.2 child.

EDUCATIONAL PRODUCTION

Taking as given that education plays an important role in development, we are led to a second set of questions: How should countries deploy school resources to increase educational attainment? Is it more important for a school with a fixed budget to reduce class sizes or to increase teachers' salaries? Do student outcomes respond more to the availability of textbooks or to enhanced teacher incentives? Understanding the impact of school inputs is indeed an important goal, one that has spawned a very large literature devoted to measuring the impact of school inputs. Unfortunately, most papers in this literature attribute a causal effect to the association between school inputs and student performance, which is not appropriate if resource allocation responds to students' needs, as will almost always be the case. Schools and parents can and do respond to the academic readiness of their children by moving the levers they have available—setting class sizes and allocating teachers' aides and classroom resources, for example. That schools respond to children's academic readiness makes it difficult to evaluate the relative merits of different inputs, and adds much confusion to the debate over resource effectiveness.

School Inputs as Choice Variables

This point is addressed thoughtfully in a paper by Lazear (1999), who builds on the idea that, at any point during the schoolday, there is some chance that any given student will be "disruptive," initiating behavior that temporarily stops other students in the class from learning. "Disruptive" behavior includes misbehavior, as well as asking questions to which other students in the class already know the answer. A prerequisite for learning to take place at a point in time is that all children in the class are nondisruptive at that moment. If p is the probability that a student is not being disruptive at a given point in time, then the probability that no student in the class of size n is being disruptive at a point in time is p^n. Lazear notes that we would expect children to behave most of the time but, even if $p = 0.98$, in a class of twenty-five students, disruptions would occur 40% of the time: ($1 - 0.98^{25} = 0.40$). Learning per student decreases with increases in class size, and increases with the probability that children are behaving.

As Lazear notes, by itself this tells us little about optimal class size. Answers to questions about optimal class size depend on several factors. We also need to know the value of education to the students, for which (as discussed above) estimates exist. In addition, we need to know the cost of providing teachers and classrooms. With this information, working under the assumption that any given child is disruptive ($1-p$) fraction of the time, we can calculate the optimal class size.

Lazear makes several related points about the relationship we should observe between class size and student outcomes if class size has been set optimally. He stresses that because class size is a choice variable, we should expect to find only small, or possibly perverse, class size effects in cross-sectional data. We should expect to see fewer disruptive students in larger classes and, if class size varies primarily because of differences in student behavior, then we should expect to find larger classes with better students and better outcomes, leading to a perverse relationship between class size and educational output.

This provides an explanation for the sometimes small and insignificant effects of class size on student outcomes reported in the literature. Surveys by Hanushek (1986, 1995), for example, argue that school facilities have little effect on outcomes, particularly on test scores (although, for developing countries, the results are quite mixed, with some research finding large and significant effects of school inputs).

One reason researchers analyzing developing country data find significant effects of school inputs, while those analyzing industrialized country data do not, may be because schools in developing countries are less responsive in general to the needs of students. Lazear also notes that his results are con-

lt to evaluate the
ion to the debate

1999), who builds
re is some chance
avior that tempo-
sruptive" behavior
other students in
ning to take place
ndisruptive at that
ng disruptive at a
in the class of size
tes that we would
= 0.98, in a class
le time: (1 − 0.98[25]
in class size, and

mal class size. An-
several factors. We
nts, for which (as
know the cost of
on, working under
raction of the time,

ship we should ob-
ze has been set op-
variable, we should
ze effects in cross-
students in larger
ferences in student
with better students
between class size

ll and insignificant
literature. Surveys
l facilities have little
gh, for developing
h finding large and

data find significant
zed country data do
e less responsive in
his results are con-

sistent with the fact that researchers find large and significant effects of school inputs in those cases where variation in inputs is due to some sort of experiment—that is, to some identifiable factor thought to influence input choice but not otherwise influence student outcomes.

Experiments and Quasi Experiments

There have been very few true experiments designed to evaluate the merits of school inputs. Some researchers argue that this is not an accident. Experiments offend the sensibilities of those who were trained in schools of education, and who view schools as complex social organizations that can be better served by management consultants than by social scientists running experiments (Cook 2001).

There are many questions, however, where evidence based on experiments may offer insight that would otherwise be lost. Glewwe et al. (2004), for example, analyze retrospective data on the presence of flip charts in Kenyan primary schools and their impact on children's test scores. When looking just at the association between flip charts and test scores, they find a large and significant effect, with the presence of flip charts increasing student test scores by 20% of a standard deviation. Such results might cause schools to place the small amount of discretionary funds they have available on flip charts. However, when these same authors ran controlled trials, in which flip charts were given to a randomly selected set of schools, whose test scores were later compared with those in a set of control schools that did not receive the flip charts, these authors found no effect of flip charts on test scores—suggesting that their earlier findings were due to some other factor correlated with both the presence of flip charts and students' test results.

Experiments designed to help us better understand the role of different school inputs are currently being run in many developing countries (Kremer 2003). However, until they are well established, our best hope of quantifying the impact of school resource allocation on student achievement comes from *natural experiments*—that is, from identifiable factors that affect school inputs, but whose only connection with the outcomes under study comes via their influence on inputs used.

Variation in school inputs that comes from such factors allows us to avoid the pitfalls discussed above—that schools' input choices reflect deliberate responses to students' needs—that prevent us from estimating the causal effects of inputs on student achievement. There is a second pitfall that natural experiments also often help us avoid. School administrators are not the only actors influencing the quality of a child's education. Parents who care about education may move to be close to good schools and may be willing to pay higher housing prices to do so. They may fight to increase local school funding and quality. Such parents may also instill in their children a strong desire to learn, and they may spend time and effort at home helping children with their studies. In such cases, a positive relationship between school resources and outcomes for children may be due to unobserved parental tastes for education, and it may not be possible to disentangle the effects of such tastes from those of school inputs. Natural experiments may allow us to do so. In the following subsections, we present results of studies on schooling in Israel and South Africa, in which school inputs were allocated in a manner that may allow us to quantify the impact of school resources.

Results from Israel: Maimonides' Rule

Angrist and Lavy (1999) examine the impact of Maimonides' Rule (a twelfth-century biblical dictum governing class size) on student test scores. Maimonides' Rule states that a class size is allowed to rise until there are forty students in a given class. When the forty-first student enters, the class is cut in half, so that instead of one class of forty-one, there are now two classes—one with twenty students and one with twenty-one students. Angrist and Lavy use the nonlinear relationship between the local number of students and the class size predicted by Maimonides' Rule to estimate the impact of class size on student performance, and to evaluate the effect of being just below the number of students for whom an additional teacher would be required, and of being just above that number. Maimonides' Rule yields highly irregular patterns in class size that are precisely mirrored in student test scores, with students in smaller classes scoring significantly higher on tests. Here, the important part of the identification strategy is that otherwise identical children are being treated differently. Their treatment depends on the number of children who are to be served locally. Angrist and Lavy find that a reduction in predicted class size of ten students is associated with a 0.25 standard deviation increase in fifth graders' test scores.

Results from South Africa

Case and Deaton (1999) use variation in school quality between magisterial districts in South Africa to estimate the impact of school quality on children's progress through school. Under apartheid, blacks were severely limited in their residential choice. Black parents were forced to send their children to black schools, whose funding decisions were made in Pretoria, by white-controlled entities on which blacks were not represented and over which they had no control. Over time, large differences evolved in average class sizes, with some districts averaging twenty children per teacher in black schools and others averaging upwards of eighty children per teacher. Controlling for household background variables—which themselves have powerful effects on outcomes, but have no effect on pupil-teacher ratios in South Africa—they find strong and significant effects of pupil-teacher ratios on enrollment, on educational achievement, and on test scores for numeracy. The striking result—that variables such as the mean schooling of parents in a community are uncorrelated with pupil-teacher ratios—provides support for the view that black families in the past had little control over the quality of the schools in their communities.

Interpreting Nonexperimental Evidence (Shoe Leather versus Technical Fixes)

Where school quality is thought to vary either because school administrators allocate resources systematically, according to student need, or because parents sort themselves according to their tastes for education, it is essential to have evidence from experiments or quasi experiments with which to judge the impact of school resources. Some researchers, grappling with these issues, have tried to minimize the estimation bias caused by the behavior of administrators and families, either by controlling directly for all available school and family variables, or by instrumenting school quality variables on variables thought to be correlated with school quality but not otherwise correlated with student achievement.

Neither of these approaches can solve the fundamental identification problem: such data do not have the information in them, and cannot be made to disgorge it by "technical fixes," such as sample selection corrections or instrumental variables.[6] It is unlikely that the complete set of variables that jointly determine school inputs and children's outcomes will ever be available to researchers. Even if they were, they cannot estimate the causal effect of school inputs: all are as much determined by achievements as by determinants of them. For this reason, calls in the literature for very expensive data collection—based on surveys that would collect information on every aspect of school production (class size, teacher incentives, textbooks, teacher autonomy, and so on) and every child and household characteristic (cognitive skills, attitudes toward school, aspirations, family background)—are misplaced. These variables are all determined jointly: students' attitudes toward school depend on their cognitive skills, which depend in turn on the students' attitudes toward school. Both of these depend upon (among many other things) class size, and class size will likely depend on students' cognitive skills and attitudes toward school. The point is not that researchers in education lack the "tools" that other social science researchers have in their toolboxes for disentangling the causal effects of each on the rest. The point is that such tools do not exist. Researchers could have at hand every variable related to school quality and to children's abilities and their households' characteristics, but they will not be able to use these data to settle disputes on the magnitude or significance of school inputs or operating style.

Estimation in which researchers identify variables that could be used as instruments for school inputs when estimating their impact on student outcomes is equally unlikely to yield meaningful results, unless there is a genuine experiment or quasi experiment where some individuals get treated in a way that affects their education—for example, by being just above or below Maimonides' cutoff for an additional class. On average, parents with more education often work to improve the quality of schools their children attend. But parents' education also has a direct effect on children's achievements, and therefore is not eligible for use as an instrument. The same is true of household income, distance to the better school, and most other variables thought to influence school quality.

Evaluating outcomes between schools that are operating under different rules in order to assess features of school operating systems is also generally not advisable. The phenomenon addressed by Lazear—that schools make choices based on the conditions they face (quality of the student body and faculty, attentiveness of parents, and so on)—comes into play here as well. To take a concrete example, the World Bank and many governments are interested in knowing whether school decentralization improves student outcomes. Decentralized schools give more control to the local decision makers, who may have superior information and may be better equipped to monitor the functioning of the schools. Evaluation of some countries' experiences with decentralization is made difficult, however, when schools are chosen by the government to participate in the decentralization program, or are allowed to volunteer for the program. Schools that choose to participate (or are "volunteered" by the government) differ in observable and unobservable ways from those that do not. The "treatment" and "control" terminology that researchers sometimes use when contrasting outcomes between two groups can obscure the fact that the choice to innovate may be related to features of the school that have their own effects on student achievement.

PROPOSALS FOR FUTURE RESEARCH

In spite of all the papers written on the relationship between school inputs and educational outcomes, almost everything is still unknown. We know that differences in the underlying conditions (pupil readiness, the value of education, the opportunity cost of teachers' time) should affect optimal allocations. We have argued above that our ignorance on the effectiveness of different policy interventions is likely to remain, unless governments or international organizations are prepared to do the hard (and sometimes expensive) work of documenting the impact of different policies. This idea is far from new. Newman et al. (1994) provide a thoughtful discussion of randomized control designs for the evaluation of social programs in developing countries, arguing not only that the results of experimental (randomized control) evaluations are the most robust, but also that they can make a virtue of necessity. Often resource constraints dictate that a program cannot begin everywhere at once. Those are often cases in which randomized control design can be built into a program's introduction at low cost.

We end by contrasting two different large-scale school intervention programs, one in which evaluation has been built in, the other in which evaluation is largely absent. Policy makers can learn lessons from both when forming a working agenda on education research.

India's District Primary Education Program (DPEP)

DPEP began in 1994, as a collaborative effort between the government of India and the World Bank, the European Commission, and the Department for International Development (DFID, U.K.). According to Aggarwal (2000), DPEP was designed to help poor areas, and was targeted to reach areas in which female literacy was especially low. Under DPEP, districts are given a high degree of discretion in developing strategies to provide access to primary education for all children and reduce primary dropout rates, equalize enrollment across genders and social strata, and improve test scores. However, the first districts chosen for treatment were selected "on the basis of their ability to show success in a reasonable time frame" (Pandey 2000, p. 14) and, within districts, the areas with the lowest female literacy were avoided. As a result, it is not possible to answer important questions—such as, on average, how effective is DPEP expected to be? That the program was intended to reach areas where female literacy was low, but program rollout was avoided in such areas, does little to aid our understanding of whether this is a program that will effectively equalize enrollment by gender!

In addition to the problems of evaluation caused by the nonrandom selection of initial sites for intervention, evaluation of DPEP is not based on the differences that develop between the DPEP and non-DPEP districts. Aggarwal notes in passing that "[t]here is a group of professionals who advocate that the progress of DPEP districts should be compared with non-DPEP districts to have more realistic assessment of the DPEP gains. While there is some justification in the argument, this provision does not form part of the proposed [monitoring and evaluation] mechanisms. . . . Comparison between DPEP and non-DPEP districts will not be adequate to measure the differential impact of DPEP since the base conditions in both areas are different in terms of many other inputs" (Aggarwal 2000, p. 36). But without data from non-DPEP districts, it is not possible to evaluate DPEP properly, which deprives educators, governments, parents, and students of a chance to understand clearly which of the many DPEP changes are working, and which ones only reflect changes in the country that are occurring in all (DPEP and non-DPEP) schools.

Mexico's Progresa Program

In contrast, school interventions in Mexico have been evaluated carefully. Mexico's Progresa (now Oportunidades) program is a large-scale poverty alleviation program designed to increase human capital. Under Progresa, parents are given transfers if their children attend school regularly. The program couldn't be initiated everywhere at once, and the decision was made to evaluate the difference in outcomes between groups who were treated and similar groups (randomly chosen) who were not.

Results of the intervention are powerful, with findings suggesting that the program has been successful in reducing the age of school entry, in decreasing the extent of grade repetition, and in reducing dropout rates (Behrman et al. 2001). This sort of evaluation, based on differences between treatments and controls, need not be as rare as is currently the case, and suggests an important way forward.

ACKNOWLEDGMENTS I thank Angus Deaton for helpful discussions and comments on an earlier draft.

NOTES

1. See the discussion provided in Card (1999).
2. See Elo and Preston (1996) for evidence from the United States, and Valkonen (1989) for estimates from Europe.
3. See Murthi et al. (1995).
4. Caldwell (1986) presents evidence that the interaction between mothers' education and access to an adequate health facility is a powerful combination in increasing child survival in Nigeria.
5. See Bledsoe et al. (1999) for a discussion on this point.
6. See Freedman (1991) for the seminal contribution here.

BIBLIOGRAPHY

Aggarwal, Y. "Monitoring and Evaluation Under DPEP, Measurement of Social Impact." Mimeo, National Institute of Educational Planning and Administration, India, 2000.

Angrist, Joshua D., and Victor Lavy. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114 (1999): 533–575.

Behrman, Jere, Piyali Sengupta, and Petra Todd. "Progressing Through PROGRESA: An Impact Assessment of a School Subsidy Experiment." Mimeo, University of Pennsylvania, 2001.

Bledsoe, Caroline H., J. A. Johnson-Kuhn, and John G. Haaga. "Introduction." In *Critical Perspectives on Schooling and Fertility in the Developing World*, edited by Caroline H. Bledsoe, J. B Casterline, J. A. Johnson-Kuhn, and John G. Haaga1–22. Washington, D.C.: National Academy Press, 1999.

Caldwell, John C. *Theory of Fertility Decline*. London: Academic Press, 1982.

Caldwell, John C. "Routes to Low Mortality in Poor Countries." *Population and Development Review* 12 (1986): 171–220.

Card, David. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier Science, 1999.

Case, Anne. "Health, Income and Economic Development." In *Annual World Bank Conference on Development Economics 2001/2002*, edited by B. Pleskovic and N. Stern, 221–241. Washington, D.C.: World Bank, 2002.

Case, Anne, and Angus Deaton. "School Inputs and Educational Outcomes in South Africa." *Quarterly Journal of Economics* 114 (1999): 1047–1084.

Cook, Thomas D. "Reappraising the Arguments Against Randomized Experiments in