



Progress and challenges in achieving an evidence-based education policy in Latin America and the Caribbean

Miguel Urquiola¹

Published online: 11 November 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract This paper reviews the progress economists have made towards achieving an evidence-based educational policy, and highlights challenges that remain. The five main findings are: (1) over the past two decades, much effort has gone into identifying the causal effects of school inputs; this work has produced results and should be expanded. (2) That said, particularly when it originates in small-scale experiments, such work is unlikely to produce a full road map for an evidence-based education policy. (3) There has been less work on the effects of educational inputs at the pre-school stage, and this may be a greater constraint to sound policy than at first appears. (4) There has also been progress on understanding the effects of incentives within education. (5) There is relatively less work surrounding the effects of inputs provided by parents and students, both in terms of resources and effort.

Keywords Education policy · Developing countries

JEL Classification I2

1 Introduction

Recent years have seen increasing calls for evidence-based education policies. In fact, academic researchers have been on a quest for such policies for decades—at least since the publication of the Coleman Report on the equality of educational

✉ Miguel Urquiola
msu2101@columbia.edu

¹ Columbia University and the National Bureau of Economic Research,
New York, NY 10027, USA

opportunity (Coleman et al. 1966).¹ This paper reviews some of the significant contributions economists have made in the area of evidence-based education policies. It also argues that this quest is incomplete, and identifies substantial knowledge gaps, some of which are unlikely to be filled based on the current trends.

Two notes on the focus of the paper are relevant. First, in keeping with this volume, the discussion emphasizes Latin America, although the evidence comes from, and the conclusions are meant to apply to, countries in other regions and with diverse income levels. Second, the focus is on assessing the skills that individuals gain in educational systems before they enter the labor market; less attention is placed on the performance of systems in terms of increasing enrollment. This focus—which Sect. 2 justifies in more detail—is based on the fact that in Latin America, as in other developing regions, much greater strides have been made in increasing enrollment. In addition, the region performs at or above what its income levels might suggest in terms of enrollment, but lags in terms of measures of skills such as test scores.²

The paper begins by setting out a conceptual framework that serves to organize past work and identify remaining gaps. It then considers the evidence, with five main conclusions:

1. Over the past two decades, much effort has gone into identifying the causal effects of school inputs. As Banerjee (2007) suggests, this trend reflects a substantial effort by economists to “get into the machine”—that is, to study schools in detail and understand the educational production function. The growth in this type of research is a salutary trend, as in many countries the implementation of many if not most educational interventions still goes unevaluated. It follows that this type of research should be expanded, even if it means that some education economists must go further into the machine. For example, work on the effectiveness of different types of curricula—a topic far afield from most economists’ training—could not only be quite productive but also seem to be a logical extension of current work.
2. That said, work that is primarily focused on identifying the causal effects of different inputs and interventions—particularly when it originates in small-scale experiments—is unlikely to produce a full road map for an evidence-based education policy. Specifically, the sometimes explicit promise of this work has been to generate a ranking of inputs and interventions in terms of cost-effectiveness, with the rationale that such rankings will one day guide policy throughout the developing world (Dhaliwal et al. 2011). For reasons discussed below, however, such rankings are unlikely to be stable across or even within countries—they may vary with the setting and/or scale of implementation, for instance. This also suggests some caution in implementation.
3. There has been less work on the effects of educational inputs during the preschool stage. The relative scarcity of such work may be a greater constraint

¹ See Rockoff (2009) for a review of systematic, if isolated, research efforts from the 1920s and 1930s.

² There is consensus around this diagnosis. For a thorough review see Vegas and Petrow (2008). See also Behrman and Birdsall (1983).

- to sound policy recommendations than at first appears. For example, if educational investments made early in life affect the productivity of investments made later (Cunha and Heckman 2007)—and this remains to be fully shown—then an evidence-based educational policy will look more complicated. In other words, it will have to take into account that part of the return to preschool educational investments that may come in the form of increasing the returns to educational investments during a later period (e.g., high school). In short, further work on the impact of preschool inputs is desirable and may also help to improve policy beyond the preschool sphere.
4. There has been progress on understanding the effects of incentives within education. For example, there has been significant work on the effects of introducing competition in school systems. In general, the effects of voucher and similar initiatives have not been nearly as robust or positive as most economists expected. That said, recent research suggests progress toward understanding how the design of policies introducing competition (e.g., school vouchers) might be improved. In short, while caution in implementation is desirable, further experimentation also seems warranted.
 5. There is relatively less work surrounding the effects of inputs provided by parents and students, both in terms of resources and effort. The key issue here is to what extent parental and student effort matter, and how their level responds to incentives. Parents are even harder for policymakers to control than school administrators or teachers, so understanding incentives is critical. For example, the return to skill in the labor market, or in admission to universities, may be crucial in terms of determining parents' and students' attitudes toward skill accumulation. The corollary is that the organization of the educational market, and its relation with the labor market, may involve a broad set of institutions that set up an educational system for either success or failure. If this is the case, then understanding these incentives may be essential not only to school system design but also to correctly exploiting all the knowledge gained from experimentation. In short, while education economists might want to invest energy "getting into the machine", they should not lose sight of the fact that broad institutions/incentives may matter (Acemoglu and Robinson 2012).

The next section lays out a basic conceptual framework. Sections 3–6, then review research and derive the implications discussed above, and Sect. 7 puts forth conclusions.

2 A focus on skills

This section provides some brief background on why this paper focuses on interventions aimed at raising skills rather than interventions aimed at raising enrollment.

Along an enrollment dimension, the educational systems of Latin American countries are not underperformers, at least in a relative sense. Enrollment rates in the region are relatively high when one controls for income levels. Figure 1 illustrates this by plotting primary enrollment rates by GDP in 2005. Enrollment

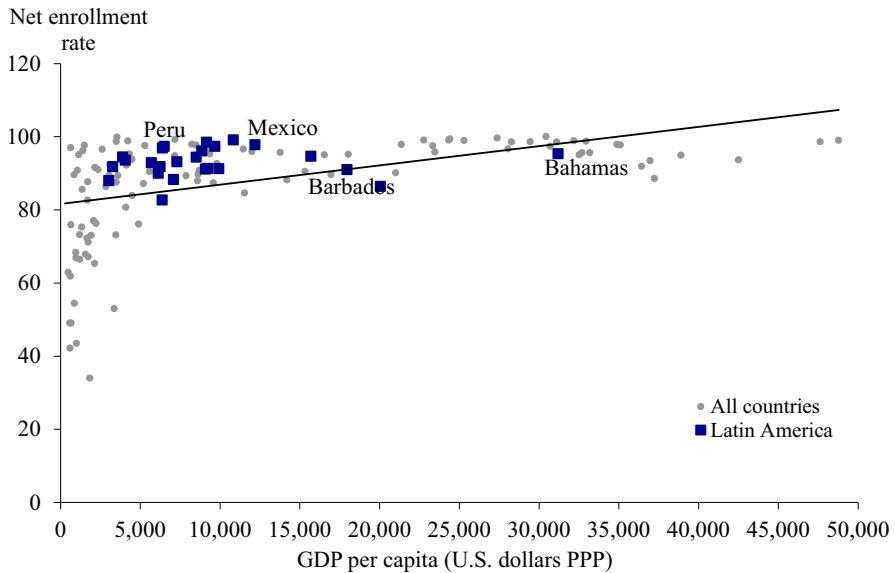


Fig. 1 Primary enrollment and per capita income (PPP), 2005. Source: based on data from the UNESCO Institute for statistics, <http://stats.uis.unesco.org>. Accessed on 1 May 2015. *PPP* purchasing power parity

rates for Latin American countries are at or above levels one would expect given incomes.

Although the result is less stark, the same holds for secondary (Fig. 2) and pre-primary (Fig. 3) enrollment rates.

Further, enrollment rates have been improving over time, and many if not most countries in the region have devoted substantial resources to achieve further improvements via conditional cash transfer programs. These programs have been credibly shown to increase enrollment. In short, this dimension has seen improvements, is likely to see more, and has not suffered from a lack of fresh resources.

The situation is quite different, however, in terms of skills, at least as measured by test scores. For instance, Fig. 4, drawn from Vegas and Petrow (2008), illustrates that in the 2003 Program for International Student Assessment (PISA) math test, not only was Latin American performance low, it was also lower given what a linear prediction based on GDP per capita would suggest.

To provide a more qualitative sense of the situation, LLECE (2001) considered the percentage of public and private school children who attain different levels of reading readiness.³ Roughly, Level 1 refers to a basic literal understanding of texts—such as being able to identify the actors in a simple plot. Level 2 is the ability to not only understand a text but to express its basic elements in words different from those used in the original. Level 3 explores whether children can “fill in the

³ LLECE stands for “Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación”.

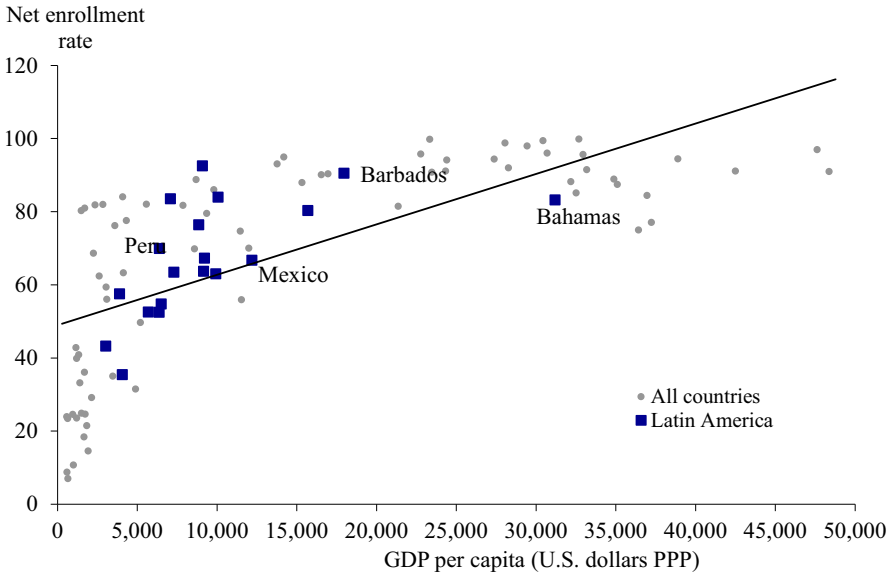


Fig. 2 Secondary enrollment and per capita income (PPP), 2005. Source: based on data from the UNESCO Institute for statistics, <http://stats.uis.unesco.org>. PPP purchasing power parity

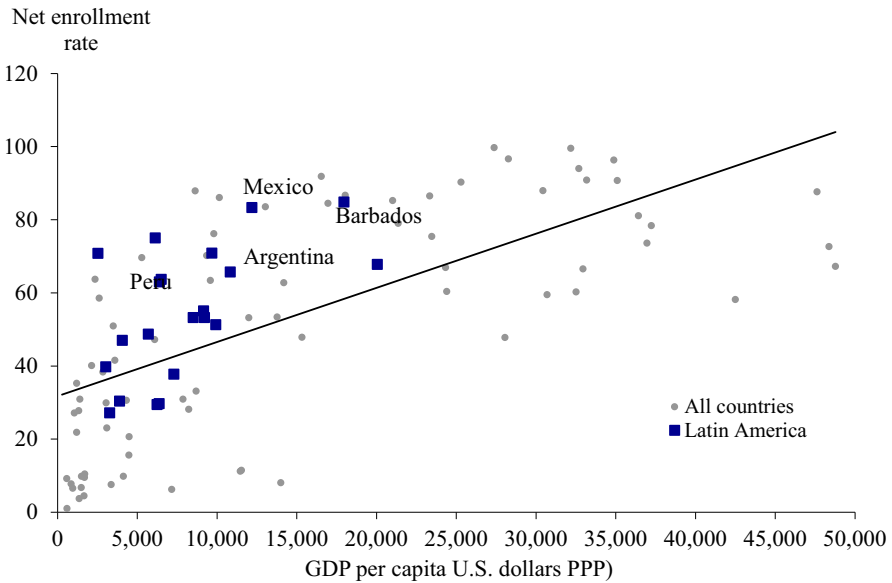


Fig. 3 Pre-primary enrollment and per capita income (PPP), 2005. Source: based on data from the UNESCO Institute for statistics, <http://stats.uis.unesco.org>. PPP purchasing power parity

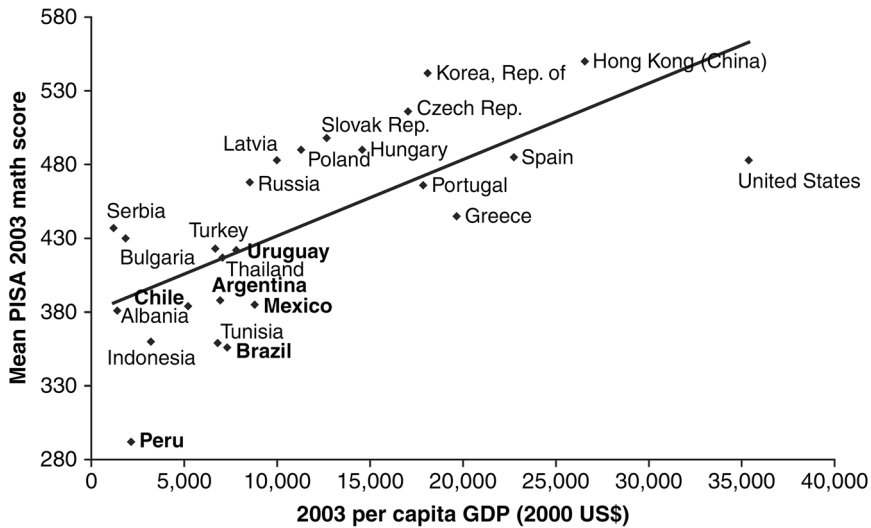


Fig. 4 Latin American performance in international tests. Source: Vegas and Petrow (2008)

blanks” in a text regarding aspects such as assumptions and causation. The majority of 3rd and 4th graders in the region have attained proficiency at Level 1, although more than one in ten children are unable to meet this benchmark in all countries save Argentina, Brazil, and Chile. By Level 3, more than half of children fail to attend proficiency in all countries, except Argentina and Chile. Thus, by this objective standard, skills in Latin American are low.

There is less of a sense of how test scores have evolved historically, as time series data on this dimension are harder to come by (although there is a clear upward trend in self-reported literacy). In international testing, the region has not made progress, with the recent exception of Chile. Beyond this, casual observation suggests that in many countries more fresh public resources have been devoted to expanding enrollment (e.g., conditional cash transfers have proliferated) than in serious undertakings to improve the production of skills. In short, by both relative and absolute measures, Latin American educational systems are producing low levels of skills. This problem—and the serious challenges that addressing it present—accounts for the fact that the remainder of this paper is devoted to this issue.

3 Framework

Suppose individuals accumulate skills over two periods, preschool and school. These can be thought of as corresponding roughly to ages 0-5 and 6-18, and we label them $t = 1$ and 2, respectively.

Assume that the skill, θ , that individuals accumulate over these two periods, is a function of:

- Their innate ability upon birth, denoted by α_0 . This is essentially a genetic draw, and as its subscript suggests, it is exogenous in terms of decisions made in periods 1 and 2.
- Their parents' endowments, p_0 . For example, literate parents have an easier time teaching their children how to read.
- Their parents' investments in each period, p_t . These investments can take the form of activities (e.g., playing with a young child or helping a teenager with homework) or expenditure on material inputs (e.g., supplying nutrition or a home computer).
- The school-based investments they receive in each period, s_t . For example, in period 1 some children will attend a preschool. In period 2, most children will attend primary and perhaps secondary school. These investments may be funded by parents (e.g., if the family pays for private schooling) or provided via public subsidies (e.g., public schools or vouchers). Note that, like p_t , s_t should be thought of as a vector that includes many components, ranging from infrastructure to effective teachers or curricular design.
- The individual effort, e_2 , that individuals exert during period 2. We assume that in period 1 children are not conscious of making an effort to learn—this is not essential—but that by primary or secondary school students realize that learning may require trade-offs such as doing homework instead of playing or working.

More formally, skill at the end of period 1 is given by:

$$\theta_1 = f_1(\alpha_0, p_0, s_1, p_1). \quad (1)$$

We will assume that all the arguments in Eq. (1)—except for α_0 and p_0 —are endogenous in the sense that they may respond to the level of the others. For instance, parents may adjust the level of inputs they provide, p_1 , if the investment on the part of schools, s_1 , changes. For a specific example, if their child receives thorough reading instruction at school, parents might be less likely to provide it at home (Das et al. 2013; Pop-Eleches and Urquiola 2013; Fredriksson et al. 2014).

Skill at the end of period 2 is given by:

$$\theta_2 = f_2(\theta_1, \alpha_0, p_0, s_2, p_2, e_2). \quad (2)$$

The presence of θ_1 reflects that skills acquired in the preschool period may affect the production of skills during primary or secondary school (Berlinski et al. 2008, 2009). Cunha and Heckman (2007) refer to this as self-productivity. Further, Eq. (2) is a general expression and does not rule out that the level of any given argument affects the impact that others have on skill. For example, children who have attained a higher level of skill in the preschool period, θ_1 , may be better positioned to benefit from school inputs, s_2 , such as teacher instruction. These are dynamic complementarities, in the terminology of Carneiro and Heckman (2003) and Cunha and Heckman (2007).

For a final ingredient, we suppose that parental investments, p_t , can also depend on returns, r , and the organization of the school system, o . For example, if parents perceive that the accumulation of skills has a high return—say in terms of their children eventually gaining admission into a selective university or securing high-wage employment—they will be more willing to invest their time and money in schooling. Similarly, the organization of the education system may affect parents' investment choices. To illustrate, some European and African countries have school systems structured such that not all children are entitled to the transition from primary into academic secondary schools. Their entry, as the school and class to which they are assigned, is contingent on testing performance. Similarly, some countries (e.g., Brazil, Chile, and Turkey) rely largely on test-based admissions for higher education. Parents may be more willing to pay for tutoring in such settings.

Assume also that at least some actors responsible for setting the level of school investments, s_t , also respond to returns, r , and the organization of the school system, o . For example, the effort teachers and administrators exert might not be independent of the incentives they have (Friedman 1955; McMillan 2005; Reback et al. 2014). Analogously, the organization of the school system itself might affect their decisions. For example, Duflo et al. (2011) show that different degrees of tracking may affect teachers' effort or preferences. Finally, it can similarly be assumed that the effort students exert may be a function of returns and organization—that is, students also respond to incentives.

4 School inputs

The above discussion highlights that an evidence-based education policy requires knowing how skills are determined by many ingredients. To discuss the state of knowledge, we begin by looking at the set of ingredients—in the framework in Sect. 2—that has received the most attention: the impact of school inputs, s_2 . This will illustrate challenges that arise even when one restricts attention to a single argument of Eqs. (1) or (2).

The interest in school inputs reflects the attention that economists give to productive efficiency. Specifically, a standard policy question is: are resources in the educational sector well allocated in terms of maximizing skills? This is a well-defined question and its analysis can provide direct implications. Addressing it requires researchers to address three challenges:

1. Obtaining data on school inputs and outcomes—in other words, measures of the different elements of s_t and of some aspect of θ_t (e.g., test scores).
2. Ascertaining the causal effect of each element of s_t (i.e., estimating terms like $\partial\theta_2/\partial s_2$)—in other words, this effect is the change in skill ($\partial\theta_2$) induced by the change in the amount of the school input (∂s_2), keeping all other factors constant.
3. Obtaining information on the costs of each element of s_t (e.g., books or teacher training).

With these three elements, one can carry out cost–benefit comparisons and, based on the result, reallocate budgets such that the “bang for the buck” is equalized across inputs.⁴ At the simplest level, this would provide an evidence-based education policy. The next three subsections review how researchers have tackled each of these three challenges in recent decades. The remaining subsections address complications and knowledge gaps.

4.1 Data availability

In its earliest stages, research on the impact of school inputs was held back by the first challenge: data availability. There was little information on what inputs were offered to different schools and children, and there were few available standardized measures of skills. In the United States, this situation began to change noticeably after the Coleman Report (Coleman et al. 1966).

Since then, efforts to compile data have intensified and today even most low-income countries collect at least some data on school resources and student achievement. In addition, initiatives such as the PISA supply data that are comparable across countries.⁵ Nonetheless, the sustained collection of data on skills is one area where governments and multilateral agencies should remain vigilant.⁶

4.2 Causality

The second challenge to devising an evidence-based education policy—that related to ascertain the causal effects of each input—is methodologically more complex, and resources alone do not necessarily solve the problem. The basic problem is that unobserved characteristics might influence both the levels of s_t and θ_t that individuals display. For example, suppose one notices a significant correlation between s_2 and θ_2 (e.g., children in schools with low class sizes test well). If s_2 is correlated with p_2 —for example, parents willing to spend on small classes might generally also be more motivated and willing to help with homework—it will be difficult to determine if the better performance is due to the lower class sizes, as opposed to greater parental involvement. In the econometric terminology, it will be challenging to isolate or identify the causal effect of s_2 , $\partial\theta_2/\partial s_2$. If research cannot achieve such identification, then no evidence-based policy on school inputs is feasible.

It deserves explicit mention that this identification problem is not typically solved by the use of statistical techniques such as multivariate regression, multi-level models, or propensity score matching.⁷ However sophisticated, such techniques can

⁴ See Levin and McEwan (2001) for a thorough discussion on cost-effectiveness analyses.

⁵ The PISA is focused on high-income countries. For greater coverage in Latin America, see UNESCO’s Latin American Laboratory for Assessment of the Quality of Education testing initiative.

⁶ Bolivia is one example where efforts are not always sustained. By the mid-1990s Bolivia had implemented achievement tests in a representative sample of primary schools, but those tests ceased in the mid-2000s.

⁷ As the saying goes, statistical techniques solve statistical problems, they do not solve identification problems.

only work with variables actually observed, leaving open the door for unobserved factors—such as parental motivation in the example above—to bias estimates.

The early literature, while producing numerous papers, rarely dealt squarely with this fundamental identification problem.⁸ Specifically, the 1980s and 1990s saw the publication of early meta-analyses and literature reviews (Hanushek 1986, 1995; Fuller and Clarke 1994) that sought to provide a guide for policy. Although there were debates (Hedges et al. 1994; Hanushek 1994) surrounding the validity of these reviews, criticism rarely focused on what was actually a central constraint to their usefulness: the quality of the available studies they reviewed in terms of isolating causal effects. As emphasized in subsequent reviews (Krueger 2003; Glewwe and Kremer 2006), inferences drawn from numerous biased estimates may of course still be biased.

Since the early 1990s, research has made significant progress in dealing with the challenge of causality. The two most common approaches used for this purpose have been randomized evaluations and regression discontinuity designs.⁹ Randomized control trials work by randomly assigning individuals or schools to treatment and control groups. The treatment group receives an educational input, s_t , while the control group does not. Random assignment generally ensures that the two groups are similar along all dimensions, including unobservable characteristics such as parental motivation. Because the only difference between the treatment and control group is that the former receives an input while the latter does not, then any difference in their outcomes can be attributed to this input.¹⁰

The regression discontinuity design aims for an analogous result. In this case, the treatment is not randomly assigned but rather depends on a “running variable”. For example, schools with average student incomes below a certain threshold might receive a school input—say, a school library—whereas schools with incomes above this threshold do not. The intuition is that while wealthier schools on average are different from low-income schools, very close to the threshold that determines treatment, they should be similar. For example, if the threshold were the 50th percentile of income, then one might compare schools with income at the 49th percentile with those at the 50th percentile. By construction, these two groups of schools are very similar in terms of income. Under the assumption that they also are similar along dimensions including unobservable characteristics such as parental

⁸ However, there certainly were exceptions. For example, Rockoff (2009) reviews rigorous studies on class size from the 1920s and 1930s. Although this research was not always on a large scale or sophisticated by modern standards, it did aim to produce causal estimates.

⁹ The renewed emphasis on these techniques was part of a broader effort by applied economists, often working in areas related to labor and education, to produce causal evidence (Card and Krueger 1992; Meyer 2005; Angrist and Lavy 1999; Hoxby 2000; Kremer and Miguel 2004; Duflo 2001). Both randomization and regression discontinuity had been applied to education topics since much earlier in the 20th century. Rockoff (2009) reviews work from the 1920s applying randomization to class size, and the regression discontinuity method dates to Thistlewaite and Campbell (1960), who analyzed the effect of scholarships.

¹⁰ For further background see Banerjee and Duflo (2008).

motivation, then any difference in their outcomes can be attributed to the school input under study.¹¹

These approaches have permitted arguably causal estimates of the impact of numerous educational inputs, including:

- Class size (Angrist and Lavy 1999; Krueger 1999; Urquiola 2006; Banerjee et al. 2007; Duflo et al. 2011; Fredriksson et al. 2012)
- Classroom libraries (Abeberese et al. 2012)
- Computers and computer-aided instruction (Linden 2008; Barrera and Linden 2009; He et al. 2009; Cristia et al. 2012; Malamud and Pop-Eleches 2011; Mo et al. 2012)
- Flashcards (He et al. 2009)
- Flipcharts (Glewwe et al. 2004)
- Lump-sum grants to schools (Pradhan et al. 2013; Das et al. 2013; Blimpo et al. 2011)
- Textbooks (Glewwe et al. 2004; Jamison et al. 1981)
- Tutoring or remedial instruction (Banerjee et al. 2007; Chay et al. 2005)
- Tutoring software (Linden 2008).

The above list is not meant to be exhaustive. For example, Glewwe (2002) provides more detail and McEwan (2013) presents an update on randomized evaluations and features a meta-analysis.¹² In particular, McEwan (2013) classifies interventions that have been evaluated via randomized control trials according to the magnitude of their impact on test scores. His summary covers not just school inputs but other types of interventions covered below. We include all of them in the following list for completeness, and then refer back to them:

- Close to zero and statistically insignificant effects: monetary grants and deworming.
- Small mean effect sizes that are not always robust to controls: providing information to parents and improving school management and supervision.
- Larger effect sizes (in ascending order of estimated impact): instructional materials, computers or instructional technology, teacher training, smaller classes, smaller learning groups within classes, ability grouping (tracking), student and teacher performance incentives, and contract or volunteer teachers.

4.3 Costs and cost-effectiveness

A third challenge to devising an evidence-based policy in the case of school inputs is to have cost information that renders cost-effectiveness comparisons feasible. Intuitively, for every element in the vector s , (e.g., class size), one asks what the

¹¹ For more background see Imbens and Lemieux (2008) and Lee and Lemieux (2010). For discussions on complications see McCrary (2008) and Urquiola and Verhoogen (2009).

¹² See also Muralidharan (2013) for a review focused on India.

causal contribution to skill is. These contributions can be compared to their costs, and resources reallocated across inputs until the “bang for the buck” is equalized across inputs. An example of this is provided by Banerjee et al. (2007), who show that while computer-assisted instruction improves learning twice as much as a remedial teacher, the latter is significantly cheaper, and so is still a better investment.

4.4 Bringing it all together

The above discussion shows that the literature has made significant progress in terms of producing information on the impact of school inputs. It is clearly desirable that such work continues and be expanded. In many countries, expensive school input initiatives are still implemented without evaluation. In some cases, these may have very small or even negative effects, and finding out as much is obviously useful.

At the same time, the promise of this type of work has been to guide the choice of education inputs. In a salient example, Dhaliwal et al. (2011) consider a number of interventions that might increase enrollment. Thus, their emphasis is on enrollment rather than skills, but nonetheless a discussion of their paper is useful. Specifically, Dhaliwal et al. (2011) draw on the evaluations done by MIT’s Abdul Latif Jameel Poverty Action Lab, along with data on the costs of these interventions. They can thus present cost-benefit comparisons that explicitly aim to guide the allocation of an educational budget—indeed; the publication is more along the lines of a policy brief than an academic paper. To cite one result, the authors suggest that expenditure on deworming is much more cost-effective than that on conditional cash transfers. This paper thus illustrates the thrust of much of the research on school inputs and, therefore, provides a useful setting in which to consider some of the challenges to this type of work.

4.5 Remaining challenges

4.5.1 External validity

A first, relatively well-understood challenge concerns external validity. Specifically, it might be that the impact of a given school input in one setting does not necessarily generalize well to other settings—initial conditions or institutional setups may matter. For example, deworming and conditional cash transfers are unlikely to have large effects on attainment in areas with low prevalence of worm infections or relatively high enrollment rates.

A related challenge is that comparisons across different contexts will be more complicated for some educational outcomes. For example, even if a certain input is found to generate a 0.25 standard deviation gain in a certain test in a certain country, it is difficult to determine what that would be equivalent to elsewhere. Tests are often administered at different levels in different countries, and there may be variation across settings regarding the impact of test score gains on, for example, wages.

The evidence in McEwan (2013) shows that these issues may be particularly challenging in terms of using existing research to construct an evidence-based policy for Latin America. The vast majority of experimentation has taken place in other settings, and in the extreme the literature may lead one to use parameters that do not apply well to the region.

4.5.2 Behavioral responses and equilibrium effects¹³

Pop-Eleches and Urquiola (2013) suggest a different complication in basing policy on cost-effectiveness comparisons.¹⁴ Specifically, they suggest that behavioral responses and equilibrium effects may render cost-effectiveness rankings such as those in Dhaliwal et al. (2011) unstable, even within a given country. The complication is that a ranking based on generally small-scale, short-lived experiments might not be an accurate guide to the ranking that would emerge if it was derived from interventions implemented on a large-scale and sustained basis.

To illustrate this point, Pop-Eleches and Urquiola (2013) start from the observation that while families cannot completely control the inputs their children receive at school, s_2 , they can influence their level. For example, parents in Latin America often pay for private schools with smaller class sizes. Even parents who do not use private schools can endeavor to get their children enrolled at (often oversubscribed) publicly subsidized Catholic schools that might, for example, offer different teacher effectiveness than regular public schools. Let s^*_2 denote the level of school inputs that households target by such actions. Assume it is a function of endowments, returns, and levels of skill that children acquired in the preschool period:

$$s^*_2(\theta_1, \alpha_0, p_0, r, o). \quad (3)$$

Schools in turn make decisions on how to allocate resources to students. For example, they might have policies that assign smaller classes or less-experienced teachers to weaker students. Formally, suppose they also condition the inputs children receive on their preschool achievement and endowments, and on returns:

$$s_2(\theta_1, \alpha_0, p_0, r, o). \quad (4)$$

The deviation between the investments in children actually receives at school and the level their parents targeted for them is, therefore, $s_2 - s^*_2$. As stated in Sect. 2, parents can react to what they observe in school in setting their own input levels. For instance, parents will know if their child made it into the nontuition-charging Catholic school they desired before they have to help with homework that school year. For period 2, parental inputs are, therefore, given by:

¹³ This section draws on Pop-Eleches and Urquiola (2013). The concepts discussed go back at least to the work of the Cowles Commission in the 1950s. See Heckman (2000) for further background.

¹⁴ See also the evidence in Das et al. (2013), Beasley and Huillery (2013), and Fredriksson et al. (2014).

$$p_2(\theta_1, \alpha_0, p_0, s_2 - s_2^*, r, o).$$

Now note that experimental and regression discontinuity analyses try to ascertain the effect of exogenously changing one element of s_2 —say class size—while holding all other inputs constant. That is, they aim to estimate terms such as:

$$\partial\theta_2/\partial(s_2 - s_2^*) = \partial\theta_2/\partial s_2 = \partial f_2/\partial s_2. \quad (5)$$

A first point to note, as emphasized by Todd and Wolpin (2003), is that the “reduced form” effects estimated by experiments may more typically also reflect changes in inputs provided by other agents, such as parents. For example, they point out that while the Tennessee Student/Teacher Achievement Ratio (STAR) class size experiment may have manipulated class size exogenously, parents were free to adjust their own effort. Fredriksson et al. (2014) provide a concrete example of such reactions. For another example, Das et al. (2013) present evidence from Africa and India suggesting that parents respond to their children’s school receiving grants by reducing their own financial contributions.

The result is that experiments may actually reveal a “policy effect” that includes such parental responses:

$$d\theta_2/d(s_2 - s_2^*) = d\theta_2/ds_2 = \partial f_2/\partial s_2 + \partial f_2/\partial p_2 \times \partial p_2/\partial(s_2 - s_2^*). \quad (6)$$

In other words, the difference between this estimate and that in Eq. (5) is that this estimate also contains the indirect behavioral response coming from parents.

From a policy perspective, this is still a useful estimate of the effect of providing a certain input. At the same time, it begins to raise some questions about conducting policy using experiments. For example, cost-benefit calculations might require ascertaining the relative contributions of school and family inputs. In addition, although in the present framework the behavioral response by parents is instantaneous, in real-world situations it might take time for parents to notice and react to changes in school inputs.¹⁵ As a result, the estimated policy effect denoted in Eq. (6) could vary with the time at which achievement is measured.

Pop-Eleches and Urquiola (2013) raise a further complication. Recall that s_2 is a vector of different school investments. To make things explicit, suppose there are two inputs: s_2^x and s_2^y . A randomized experiment might be able to vary one of these, while controlling the level of the other. In that case, the resulting impact will still resemble Eq. (6). For example, Duflo et al. (2011) manipulate the peer quality of the classes children have access to, say s_2^x , while at the same time constraining changes to other school inputs (e.g., teachers are randomly assigned to high- or low-achievement classes).

Now suppose the increase in s_2^x originates not in an experiment but from an extensive and sustained policy. In such cases not just parents but the school system will have a chance to react, and the total effect is:

¹⁵ For example, in Das et al. (2013) the response varies with parents’ information sets in a way that is intuitive. When the grants schools receive are unexpected, parents do not adjust their behavior; when they are expected, they do.

$$\begin{aligned}
 d\theta_2/d(s^x_2 - s^{*x}_2) &= d\theta_2/ds^x_2 \\
 &= \partial f_2/\partial s^x_2 + \partial f_2/\partial s^y_2 \times \partial s^y_2/\partial s^x_2 \\
 &\quad + \partial f_2/\partial p_2(\partial p_2/\partial (s^x_2 - s^{*x}_2) + \partial p_2/\partial (s^y_2 - s^{*y}_2)), \quad (7)
 \end{aligned}$$

which differs from Eq. (6) in also including responses within the school system.

Such responses, which Pop-Eleches and Urquiola refer to as “equilibrium effects”, may only emerge once interventions are taken to scale and sustained for a period of time.¹⁶ For example, if tracking is sustained, more experienced teachers may sort toward the higher achievement classes, and their ability to do so may gradually become enshrined in norms or even union contracts.

To summarize, Todd and Wolpin (2003) make a useful distinction between production function parameters (Eq. 3) and policy effects (Eq. 6). This raises some complications, but experiments can still provide at least rough guidance on both. Pop-Eleches and Urquiola (2013) further emphasize that policy effects might be different in situations where behavioral responses take time to unfold, or where these responses only appear when certain interventions reach a certain scale—Eq. (6) versus Eq. (7).

This matters because behavioral responses and equilibrium effects may limit the ability of extensive experimentation to deliver evidence-based policy. The basic argument is that, essentially by design, experimental research deals with small-scale intervention. For example, the very fact that a control group must be constructed requires some constraint on the scale of implementation. Further, some agents such as parents or teachers may not be given time to react, or such reactions may be deliberately precluded to identify the causal effect of a school input. Exercises such as Dhaliwal et al. (2011) deliver a ranking of interventions under these conditions, but the ranking may change as interventions are taken to scale or sustained.

4.5.3 Long-term effects

A related challenge arises because the effects of different inputs—whether these arise from direct effects or from behavioral responses—may only be observed in the long run. In the case of class size, for example, the STAR experiment highlights a situation in which early effects on test scores faded out by higher grades, yet effects on college attendance emerge later in life (Schanzenbach 2007).

5 Preschool inputs

While evidence has increased substantially on the effect of inputs in K-12 schooling— s_2 in the framework of Sect. 2—there has been less work identifying the causal impact of inputs on preschool outcomes: the effect of s_1 on θ_1 . There is casual evidence that nutritional and other interventions during the preschool period can have a significant impact on developing skills, both in developed countries

¹⁶ One could complicate Eq. (7) further and have s^x_2 and s^y_2 responding further to parental inputs. This is one reason these are labeled equilibrium effects. See, for example, the discussions in Banerjee and Duflo (2008), Acemoglu (2010), and Deaton (2010).

(Currie and Thomas 1995; Ludwig and Miller 2007; Heckman et al. 2012) and developing countries (Grantham-McGregor et al. 1994; Behrman et al. 2009; Attanasio et al. 2012). That said, there has been less work that focuses on specific inputs and mechanisms.

As discussed in Sect. 2, further examination of the preschool period raises a host of issues in terms of designing policy. For example, it brings into focus the question of relative productivities—at what time does an intervention via, say, public provision of preschool inputs, best help disadvantaged children? For another example, Cunha and Heckman (2007) show that once one allows for skill levels to be the product of investments during multiple periods, the information one would ideally want to formulate policy grows significantly. They point out that such considerations may help explain the higher productivity of investments in disadvantaged young children relative to productivity of investments in those same children when they are older. If at a later period of a child's life the child's skill level is low, the productivity of inputs directed toward raising those skills may also be lower.

Thus, in addition to establishing the impact of inputs on the development of skills in preschool, exploring the existence and magnitude of complementarities is an important topic for research. If complementarities are small, then the precise period in which children receive investments (say period 1 or 2 in the framework in Sect. 2) is not crucial. But if they are large, then the period of investment is indeed important. This is a challenging avenue of research. For an interesting example, Aizer and Cunha (2012) present evidence on these issues while acknowledging the substantial challenges involving both data (human capital is rarely measured at multiple points in children's life) and identification (human capital investments by parents are endogenous).

6 Parental inputs

While it has long been recognized that the impact of parental inputs may be crucial to skill development—if only because parents have the most contact with children early on—there is much less well-identified evidence of the actual effect of those inputs. This reflects, among other factors, the fact that it is difficult to experimentally manipulate parental activities. Nevertheless, there is work on how the home environment affects outcomes for school children (Carneiro et al. 2012). There are also emerging but salient examples of experimental work. For example, Attanasio et al. (2012, 2013) report on a randomized study in Colombia that changes children's nutritional intake at home, and also tries to manipulate the way children's parents relate to them. The intervention consists of two components: nutrition and stimulation. The nutritional component provides “sprinkles” that parents can dissolve into food and provide vitamins and minerals. The stimulation component consists of weekly sessions by a home visitor who shows a child's mother different types of activities (e.g., songs, rhymes, and games with puzzles and toys) with which she can engage the child. The visitor encourages the mother to participate in these activities with the child during the following week.

A notable aspect of this study is that it was designed with scalability in mind. The sprinkles, for example, are inexpensive and easy to procure. The stimulation training for the mother is carried out by “lead mothers” (*madres lideres*) selected via Colombia’s “Families in Action” (*Familias en Accion*) Program, which is the country’s main conditional cash transfer mechanism. These women have community leadership roles in Families in Action but otherwise have received only the relatively brief training provided by the program. An academic paper is not yet available on this work, but preliminary results suggest that the stimulation component had significant positive effects on a range of outcomes relevant to cognitive language and motor development, sociability, and inhibitory control.

There is also little work on parental inputs at later educational stages, although, as reviewed above, some recent work explores how different types of parental participation respond to changes in the supply of school inputs. The bottom line is that better understanding of the supply and impact of parental inputs seems like a worthwhile focus for future research.

7 Incentives¹⁷

7.1 School choice and competition

As the previous sections make clear, education economists have been interested in the effect of school inputs on educational achievement. The combination of these two concepts has naturally led parts of the research agenda to focus on the concept of school productivity. Hoxby (2002) provides a useful definition of productivity: a school that is more productive is one that produces higher achievement per dollar spent.

One reason to focus on productivity is that many school systems have experienced declines in productivity, at least when measured with test scores as an outcome. For example, Hanushek (1996) describes such a productivity decline in US schools—greater expenditure with no test score improvement—and Pritchett (2003) suggests that this development is common among member of the Organization for Economic Cooperation and Development. Data restrictions make it harder to make analogous statements about developing countries, but the *prima facie* evidence is consistent with many of these countries also having increased real expenditures with at best small test score gains to show for it.

An influential view in economics argues that the way to address this problem is to enhance the incentives and competition faced by schools—essentially by strengthening the incentives and rewards captured by in the framework in Sect. 2. Friedman (1962), for example, suggested that the State could subsidize schooling—perhaps based on equity or efficiency considerations—while leaving the actual production of schooling to the private sector, thereby strengthening incentives and accountability. This general view on how to improve public service delivery is shared by the World Bank (2004), which, while not necessarily advocating outright privatization of

¹⁷ This section draws on MacLeod and Urquiola (2013).

schooling, calls for a greater level of accountability in schools that would emulate the accountability observed in private markets.

Another reason to explore the potential advantages of private schooling is because it is common in the developing world, although the reasons behind this depend on the context. In urban areas in Chile, for example, more than 50 % of children attend private schools. In such middle-income countries, private enrollment rates typically are high if private schools are eligible for significant State subsidies. In contrast, low-income countries sometimes see increased private schooling with little State support, perhaps in response to a public supply that is barely functioning. For example, Andrabi et al. (2008) note that by the end of the 1990s, nearly all wealthy Pakistani children in urban areas, almost a third of wealthy rural children, and close to 10 % of children in the poorest deciles nationally, were studying in private schools. In another instance, Kremer and Muralidharan (2006) pointed out that about 25 % of children in rural India have access to fee-charging private schools. In settings like India and Pakistan, these are mainly for profit schools that charge low tuition and operate at low cost by hiring young, single, untrained local women as teachers and paying them significantly less than the certified teachers more common in public schools.

The next section turns to the evidence on these issues from voucher programs. We follow Epple et al. (2015) in making a distinction between small and large voucher programs. They identify small programs as those that place significant restrictions on who can receive vouchers. The most common restrictions involve income or geography—for instance, vouchers may be made available only to low-income children in a given municipality within a country. By large programs, they mean those in which vouchers are distributed country-wide and with minimal restrictions on the type of children who can use them.

A final note before proceeding to the evidence is that this area of research—unlike those reviewed above—is one in which a disproportionate amount of work has been focused on Latin America. This reflects the fact that Chile and Colombia provide among the most salient examples of large and small voucher systems, respectively.

7.2 Small programs

The literature on small voucher programs most frequently asks if there is a significant advantage in terms of the productivity of private schools. In general, there is no consistent evidence of such an advantage. For example, Barrow and Rouse (2009) conclude that the best research to date finds relatively small achievement gains for students offered vouchers, most of which are not statistically different from zero.¹⁸ The studies that lead to such conclusions are often based on experimental designs. For example, New York City conducted an experiment in

¹⁸ See also Neal (2009). This finding is consistent with a broader literature on the effects of attending a higher-achieving school or class on academic performance, even when these transfers occur within a given (public or private) sector. Here again several papers find little or no effect (Cullen et al. 2005, 2006; Clark 2010; Duflo et al. 2011; Abdulkadiroglu et al. 2011; Dobbie and Fryer 2011). Some papers find positive effects (Pop-Eleches and Urquiola 2013; Jackson 2010), but no uniform pattern emerges. The studies on general equilibrium effects similarly suggest mixed results.

1997 that randomly allocated school choice vouchers to low-income students. The research suggests that winning a voucher to attend a private school had a modest and statistically insignificant impact on student learning, not just on average but across the distribution of preexisting ability (Mayer et al. 2002; Krueger and Zhu 2004; Bitler et al. 2013).¹⁹

There is analogous research in Latin America.²⁰ For example, Angrist et al. (2002, 2006) and Bettinger et al. (2008) look at Colombia. For context, from 1992 to 1997, Colombia operated a secondary school voucher program, a central goal of which was to increase secondary (6th–11th grade) enrollment rates by using private sector participation to ease public sector capacity constraints that mostly affected the poor. As a result, the vouchers were targeted at entering 6th grade students who resided in low-income neighborhoods, attended public school, and were accepted at a participating private school.

The initiative was implemented at the municipal level, with the national government covering about 80 % of its cost, and municipalities contributing the remainder. Resource constraints at both governmental levels resulted in excess demand in most jurisdictions. When this happened, the vouchers were generally allocated via lotteries.

These lotteries make it feasible to estimate the causal effect of winning a voucher to attend private school. Angrist et al. (2002, 2006) and Bettinger et al. (2008) find that, in general, lottery winners have better academic and nonacademic outcomes than lottery losers. This result holds both for achievement measured using administrative data, and for outcomes (such as performance on standardized exams) that the researchers themselves measured.

It should be noted that in terms of identifying whether there is an advantage in private schools regarding test scores, the Colombian voucher experiment has a few problems. First, the vouchers were renewable contingent on grade completion, and thus the program included an incentive component—voucher winners faced a stronger reward for doing well at school. Therefore, it is difficult to rule out that the superior test performance of lottery winners was due to external incentives rather than to their schools' productivity in terms of testing. Second, both lottery winners and losers tended to enroll in private schools, particularly in larger cities. Focusing on Bogota and Cali, Angrist et al. (2002) point out that while about 94 % of lottery winners attended a private school in the first year, so did 88 percent of the losers. This is not surprising to the extent that high private enrollment rate in secondary school was symptomatic of the very supply bottlenecks that the program was implemented to address. Since the reduced-form estimates in these papers are based

¹⁹ Mayer et al. (2002) and Krueger and Zhu (2004) find positive effects for some subgroups, although the conclusion depends on how subgroups are defined.

²⁰ There is a large literature on private/public comparisons in developing countries that extends beyond the case of Colombia covered in this subsection. As is the case in the United States, papers meet with varying success in terms of establishing credible control groups. Some implement only cross-sectional analyses, while others look for explicit sources of exogenous variation. For a review on several countries, see Patrinos et al. (2009); for Latin America, see Somers et al. (2004); for Chile, see Bellei (2007) and McEwan et al. (2008); for India, see Kingdon (1996); for Indonesia, see Newhouse and Beegle (2006); and for Pakistan, see Das et al. (2006).

on a comparison of lottery winners and losers, they in some cases measure a “private with incentives vs. private without incentives” effect, rather than the effect of private vs. public schooling that the literature typically addresses. Finally, the institutional setup implies that many voucher winners (who, again, would have used private school even if they did not win the lottery) used the vouchers to “upgrade” to more expensive private schools. Thus, part of the effect of winning a lottery could reflect the access to greater resources, as opposed to a true test productivity difference.

7.3 Large programs

The studies discussed above can be described in economic terminology as taking a “partial equilibrium” approach in the sense of looking at relatively small interventions—for instance, the distribution of vouchers to a small fraction of the population. This type of work essentially seeks to identify what would happen if one took a small number of children from public schools and transferred them to private schools.

More generally, one would like to consider situations that explore the general equilibrium effects of school choice—settings that give an idea of the consequences of allowing a large number of private schools to enter the market, along with allowing parents to use any of them. This is relevant because the magnitude of a “partial equilibrium” private advantage like that measured in Colombia may not be stable with respect to the private sector’s market share. For example, Hsieh and Urquiola (2006) and Bettinger et al. (2008) point out that if the private productivity advantage originates in positive peer effects, then the magnitude of this advantage may change with growth in the private sector. This in turn reflects the fact that the composition of students in the private and public sector is likely to change with private entry.

A useful setting to ask such questions is Chile. Specifically, in 1981 Chile introduced a universal voucher scheme that resulted in a substantial increase in enrollment in private schools.²¹ By 2009, about 57 % of all students nationwide attended private schools, with voucher schools alone accounting for about 50 %. The latter group combined with a public share of 44 % means that about 94 % of all children attended effectively voucher-funded institutions.²²

The analytical virtue of this reform is that it provides an example of a large-scale introduction of competition; the main drawback is that the simultaneous nationwide implementation makes it difficult to establish counterfactuals. As a result, most studies have adopted quasi-experimental methodologies. Hsieh and Urquiola (2006) apply a difference-in-differences approach to municipalities for the 1982–1996 period. They find that municipalities that experienced faster growth in private sector market share show distinct signs of increasing stratification (with higher income students in the public sector moving to private schools), but do not have higher test scores or average years of schooling.

²¹ For further institutional details see McEwan and Carnoy (2000) and Urquiola and Verhoogen (2009).

²² The “elite” unsubsidized private schools account for about 6 % of enrollments.

Even setting identification issues aside, these estimates do not isolate the effects of competition on productivity in the sense of Hoxby (2002). Many things were changing for Chilean schools during this period, including the distribution of students (and hence potential peer effects) and levels of funding.²³ Taken at face value, however, these findings suggest that competition had a modest effect on average school productivity.²⁴

This research must also be considered along with aggregate trends. If there is a substantial private productivity advantage, then one would expect Chile's relative performance on national and international tests to have improved over the years in which large numbers of children were transferred into the private sector. Furthermore, one would expect Chile to outperform other countries with similar levels of GDP per capita. However, neither of these expectations is supported by the data for at least the first 25 years of the voucher program.

As Epple et al. (2015) point out, Chile's recent performance on international testing has been more favorable. This improvement has coincided with a further expansion in private schooling. But it also coincides with more growth in GDP per capita and educational expenditures, expansions in preschool enrollments, and reforms to rules governing university admission. Thus, it is difficult to causally assign this recent improvement to the voucher program.

To summarize, the evidence from developing countries suggests that large-scale expansion of the private school sector leads to stratification,²⁵ but there is less evidence that it leads to substantial gains in average school productivity. This is consistent with the lack of a systematic private school advantage referenced above, and additionally suggests that the introduction of competition may not by itself have a large impact on public school productivity.

7.4 School choice: further challenges

When it comes to improving school skills, school choice programs have proved somewhat disappointing—the evidence is mixed but clearly not sufficient to consider this a silver bullet (Epple et al. 2015). Going forward two questions for research are:

1. Why is it that the effects of school choice programs have proven smaller than economists might expect?

²³ The value of the school voucher fell significantly during the 1980s and grew substantially during the 1990s.

²⁴ Auguste and Valenzuela (2006) and Gallego (2006) analyze cross-sectional data, using instruments for the private market share. Auguste and Valenzuela use the distance to a nearby city, and Gallego uses the density of priests per diocese (with the reasoning that this lowered the costs of Catholic schools). The results from both papers differ from those of Hsieh and Urquiola (2006) in that both find that private entry results in higher achievement, and concur (in the case of Auguste and Valenzuela—Gallego does not analyze the issue) in finding that it also leads to stratification. Again, however, a key issue is the validity of the instrumental variables. It is possible, for example, that more motivated parents migrate toward cities in search of better schools, or that priests were allocated to communities in a manner correlated with characteristics (e.g., population density) that might affect educational achievement.

²⁵ For other examples of school market liberalization leading to stratification see Bjorklund et al. (2005) and Mbiti and Lucas (2009) for Sweden and Kenya, respectively.

2. How can choice schemes be better designed in terms of generating skill improvements?

Some recent theoretical and empirical studies attempt to make headway in this direction. At the heart of them is the notion that school choice can only be expected to deliver what parents want. In an interesting recent study, Muralidharan and Sundararaman (2013) address this issue while combining some elements of the partial and general equilibrium approaches described above. Specifically, the authors implement a project in which applicants for vouchers were first recruited in a number of towns, with two lotteries carried out subsequently. First, some towns were selected for distribution of vouchers. Second, within the towns selected for treatment, some children were randomly selected to receive the vouchers. This allows Muralidharan and Sundararaman (2013) to go beyond the usual comparison (lottery-winning individuals versus lottery-losing ones) in most studies. As one example, by comparing nonapplicants in towns that did not receive vouchers to nonapplicants in towns that did, the authors can get a sense of negative effects on children “left behind” in the public sector. In the study, the authors do not find much evidence of such externalities.

Moving on to the results for the applicants, Muralidharan and Sundararaman found that after 4 years of treatment, lottery winners did not have higher test scores than losers in five of six subjects. Specifically, they found no effects in Telugu (the local language), Maths, English, Science, or Social Studies; in contrast, they did find significantly higher scores in Hindi. Two aspects are of note beyond these reduced-form results. First, the results are generally consistent with a differential allocation of instruction time: private schools seem to spend more time teaching Hindi than other subjects; public schools essentially do not teach Hindi. Second, the results are consistent with a productivity advantage for private schools, since these schools have lower costs than the public schools the students transferred from.

However, some questions remain. The first is if the positive effects on Hindi are a school effect. The paper argues this is the case, but they could also be a peer effect. The private schools may offer greater exposure to the children of parents who value Hindi, perhaps because they are from other parts of India, or because they live in large cities, or because they speak it more at home. As a result, the voucher winners might learn more Hindi as a result of being exposed to such children rather than because the schools teach it. If the types of parents who use private schools are in fixed supply (at least in the short or medium term), the partial versus the general equilibrium effects of choice could once again differ.

A second issue is that the paper suggests that English as a medium of instruction disrupts learning, that parents may not realize this is the case, and that intervention may be warranted. But another possibility is that parents are aware of this but willing to make the sacrifice, if for example, English has high labor market returns.

A broader point this illustrates is that choice is likely to produce more of what parents want, and those skills may or may not be the ones along the dimensions of what policymakers prefer. MacLeod and Urquiola (2012, 2015) address this issue theoretically, and suggest that the impact of school choice programs—and any observed private productivity advantage—is unlikely to be invariant with respect to

how choice programs are designed. In their model, parents care about what schools their children go to for two reasons: (1) schools/colleges produce value added, enhancing human capital investment, and (2) schools/colleges serve as a signal of unobserved ability.

MacLeod and Urquiola (2012, 2015) show that if this is the case, parents will want schools with good reputations, as expected. The key, however, is that schools' reputations depend not just on how good they are at teaching, but also on which other students are using them. In these situations, for example, rational parents will not always choose the high-value-added schools, and rational schools will not always choose to compete on value added. These implications are consistent, for example, with the well-identified empirical evidence that selective schools only sometimes produce higher learning and value added (Clark 2010; Abdulkadiroglu et al. 2011; Pop-Eleches and Urquiola 2013).

MacLeod and Urquiola (2013) suggest, for example, that a design in which schools are forced to use lotteries in selecting students (as in Sweden's voucher scheme or the US charter school system) may raise school productivity more than a design that allows private schools to easily turn away students, as does Chile's system. These are system design questions that are not easily analyzed via experiments or quasi-experiments, but which may nonetheless be central to successfully raising the production of skills.

7.5 Incentives for parents and students

In a similar vein, it may be that skill accumulation crucially depends on how system design affects the incentives for student or parental effort. For example, it may be that unless students are willing to study and learn, no amount of school inputs or competition between schools will improve outputs (Bishop 2004). This raises the possibility that research should prioritize learning about terms like $\partial\theta_2/\partial s_2$ and $\partial\theta_2/\partial e_2$, and about how system design affects the incentives needed to encourage parents and students to supply effort.

There is also experimental work in this area, through studies that provide rewards for students who perform well. The results are somewhat mixed. For example, Kremer et al. (2009) find positive effects of such rewards, while Li et al. (2010) find little effect in China (unless rewards are combined with other interventions). The mixed evidence in this area lines up with results from developed countries and other educational levels (Angrist et al. 2009).

Rather than focusing on payments for test scores, MacLeod and Urquiola (2015) present a model emphasizing that incentives for parental and student effort may emerge from the link between educational and labor markets. There is emerging, well-identified empirical evidence consistent with this possibility. For example, Jensen (2010) finds that boys in the Dominican Republic are quite responsive to information on the returns to a high school education. A randomized intervention on this produced gains on the order of a quarter to a third of a year of schooling. Nguyen (2008) finds qualitatively similar results for Madagascar. A concern with these relatively early studies, however, is that the information provided may not reflect the causal returns to schooling, let alone be informative about how these

returns may vary with student characteristics. In the extreme, individuals might have been reacting to misleading information.

In more recent work, Jensen (2012) and Oster and Millett (2011) present situations in which “real world” information on job opportunities affects student investment and behavior. For example, information on the availability of call center employment in India (which is open mainly to young women) affects the probability that girls remain in school.

MacLeod and Urquiola (2013) similarly suggest that the organization of the school system may affect incentives and effort. For example, systems that generally emphasize meritocratic transitions between educational levels (e.g., middle to high school in Romania or high school to university in Chile or South Korea) may be better placed to extract effort and high human capital investment from students and parents. Again, recent research on the economics of education—by placing very high priority on identification and micro-studies—may be paying insufficient attention to these “big” design questions.

7.6 Teacher incentives

Another area where the impact of incentives has been explored concerns teacher behavior. Muralidharan and Sundararaman (2011) find positive effects of teacher performance pay on student learning outcomes. Glewwe et al. (2010) find analogous effects in Kenya, but the impact is focused on incentivized exams. This is similar to findings in the United States. Although not looking directly at the production of skills, Duflo et al. (2012) find that monitoring in the form of pictures/time stamps improved teacher attendance.

There has also been work on the use of contract teachers—instructors who are not hired into normal civil service positions and are generally paid substantially less. In his review of this work, McEwan (2013) points out that the effects here are generally positive. An analytical complication, however, is that the effects are often combined with other treatments such as class size reductions (Muralidharan and Sundararaman 2010; Duflo et al. 2011; Bold et al. 2013), which makes it difficult to isolate the effect of contract teachers.

Three notes are warranted regarding the applications of such findings to Latin America. First, although in many countries of the region contract teachers are rare, they are more common in some of the lower income areas where there are contract teachers (*maestros interinos*) who are sometimes hired locally by parent associations. Second, although absenteeism problems seem to be less severe in Latin America than in Africa or India, they are certainly not irrelevant, and so these interventions may have significant returns. Third, this is also an area where the equilibrium effect may be quite relevant. For example, in the short run it may be possible to increase teachers’ attendance by monitoring them with cameras, or by having the flexibility to substitute regular teachers with contract teachers. Over time such policies may have unintended effects. To illustrate, using cameras to force teachers to show up to work amounts to a reduction in real wages. A significant proportion of teachers might have signed up for work in remote locations with the expectation that they only needed to show up 3 or 4 days a week, leaving the other 1

or 2 days for travel to a larger town. If they find that the attendance policy is enforced beyond an experiment, higher nominal wages may be necessary to attract comparable teachers to remote locations.

Finally, while this discussion has treated incentives for students and for teachers separately, recent work finds that there may be important interactions between them. Specifically, Behrman et al. (2015) consider an experiment that provided test-based incentives for: (1) students, (2) teachers, and (3) students, teachers, and school administrators. They find that the third intervention produced the largest effects, while the second had no impacts. Exploring such interactions may be an important avenue for future research.

8 Conclusion

Producing an evidence-based policy outline for how educational systems might better produce and improve skills is not a simple task. Nevertheless, research has made substantial advances in this direction. In particular, the past decades have seen progress in terms of data availability and the credible estimation of the causal effects of given educational inputs.

These results provide an initial impression of what an evidence-based policy might look like. For example, McEwan (2013) reviews a large number of experiments and concludes that the most promising interventions are found in instructional materials, computers or instructional technology, teacher training, smaller classes, smaller learning groups within classes, ability grouping, student and teacher incentives, and contract or volunteer teachers. In contrast, he finds less impact from monetary grants, deworming, providing information to parents, and improving school management and supervision. Dhaliwal et al. (2011) look at a different set of interventions, but go a step further by incorporating cost information and ranking alternatives based on cost-effectiveness comparisons. The type of careful evaluation work these reviews are based on is desirable and should be sustained, if only because many relatively expensive input initiatives in Latin America are still designed and implemented without much reference to planning for evaluation.

There is a need for analogous research on the impact of educational inputs at the preschool stage. Such work is less common, but may be quite important if part of the return to preschool investments comes in the form of a greater return to investments at later stages of a child's life. In this case, properly quantifying the impact of preschool interventions may be quite valuable.

Despite such progress, there are likely limits to the extent to which experimental and quasi-experimental evaluations—and reviews like those in Dhaliwal et al. (2011)—can guide policy. One key issue here is that the rankings that emerge may not be stable across or even within countries. For instance, they may vary with the setting and/or scale of implementation. This reflects issues related to external validity and equilibrium effects and suggests that caution in implementation is warranted.

Further progress on understanding the effects of incentives would be valuable. For example, there has been significant work on the effects of introducing competition in school systems. In general, the effects of voucher and similar

initiatives have not been nearly as robust or positive as many economists expected. The key challenge in this area is to understand what accounts for this disappointment, and what this implies regarding how programs might be better designed. The bottom line is that while the evidence is not sufficient to warrant widespread adoption of such initiatives, it certainly points to the desirability of continued experimentation.

Additional research on the impact and determinants of parental and student effort is also desirable. For example, the returns to skills in the labor market, or in admission to universities, may be crucial in terms of determining parents' and students' attitudes toward skill accumulation. A corollary is that the organization of the educational market, and its relation with the labor market, may involve a broad set of institutions that sets an educational system up for either success or failure. If this is the case, then understanding these incentives is not only desirable in and of itself, but may be important to correctly exploit all the knowledge gained from experimentation.

Acknowledgments The author thanks Sebastian Galiani, Patrick McEwan, and an anonymous referee for useful comments and suggestions. This work was supported by the Inter-American Development Bank (IDB). All errors and positions are those of the author and should not be attributed to the IDB.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abdulkadiroglu A, Angrist J, Pathak P (2011) The elite illusion: achievement effects at Boston and New York exam schools. NBER Working Paper No. 17264, National Bureau of Economic Research, Cambridge MA
- Abeberese A, Kumler T, Linden L (2012) Improving reading skills by encouraging children to read: a randomized evaluation of the sa aklat siskat reading program in the Philippines. IZA Discussion Paper No. 5812, Institute for the Study of Labor, Bonn
- Acemoglu D (2010) Theory, general equilibrium, and political economy in development economics. *J Econ Perspect* 24(3):17–32
- Acemoglu D, Robinson J (2012) Why nations fail: the origins of power, prosperity, and poverty. Crown Business, London
- Aizer A, Cunha F (2012) The production of human capital: endowments, investments, and fertility. NBER Working Paper No. 18429, National Bureau of Economic Research, Cambridge, MA
- Andrabi T, Das J, Khwaja A (2008) Students today, teachers tomorrow: identifying constraints on the provision of education. World Bank, Washington, DC
- Angrist J, Lavy V (1999) Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quart J Econ* 114(2):533–575
- Angrist J, Bettinger E, Bloom E, Kremer M, King E (2002) The effect of school vouchers on students: evidence from Colombia. *Am Econ Rev* 92(5):1535–1558
- Angrist J, Bettinger E, Kremer M (2006) Long-term consequences of secondary school vouchers: evidence from administrative records in Colombia. *Am Econ Rev* 96(3):847–862
- Angrist J, Lang D, Oreopoulos P (2009) Incentives and services for college achievement: evidence from a randomized trial. *Am Econ J Appl Econ* 1(1):136–163
- Attanasio O, Fitzsimons E, Grantham-McGregor S, Meghir C, Rubio-Codina M (2012) Early childhood stimulation, nutrition and development: a randomized control trial. Institute for Fiscal Studies, University College, London

- Attanasio O, Grantham-McGregor S, Fernandez C, Fitzsimons E, Rubio-Codina M, Meghir C (2013) Enriching the home environment of low-income families in Colombia: a strategy to promote child development at scale. Bernard Van Leer Foundation
- Auguste S, Valenzuela JP (2006) Is it just cream skimming? School vouchers in Chile. *Fundacion de Investigaciones Economicas Latinoamericanas*
- Banerjee A (2007) Inside the machine: toward a new development economics. *Boston Rev* 32(2):12–18
- Banerjee A, Duflo E (2008). The experimental approach to development economics. NBER Technical Report 14467, National Bureau of Economic Research, Cambridge, MA
- Banerjee A, Cole S, Duflo E, Linden L (2007) Remedying education: evidence from two randomized experiments in India. *Quart J Econ* 122(3):1235–1264
- Barrera F, Linden L (2009) The use and misuse of computers in education: evidence from a randomized controlled trial of a language arts program. World Bank, Washington, DC
- Barrow L, Rouse C (2009) School vouchers and student achievement: recent evidence and remaining questions. *Annu Rev Econ* 1:17–42
- Beasley E, Huillery E (2013) School resources, behavioral responses and school quality: short term experimental evidence from Niger. Paris School of Economics, Paris
- Behrman JR, Birdsall N (1983) The quality of schooling: quantity alone is misleading. *Am Econ Rev* 73(5):928–946
- Behrman JR, Calderon MC, Preston S, Hoddinott J, Martorell R, Stein AD (2009) Nutritional supplementation of girls influences the growth of their children: prospective study in Guatemala. *Am J Clin Nutr* 90(5):1372–1379
- Behrman JR, Parker SW, Todd PE, Wolpin KI (2015) Aligning learning incentives of students and teachers: results from a social experiment in Mexican high schools. *J Polit Econ* 123(2):325
- Bellei C (2007) The private-public school controversy: the case of Chile. Program on Education Policy and Governance Research Working Paper 05–13, Harvard University
- Berlinski S, Galiani S, Manacorda M (2008) Giving children a better start: preschool attendance and school-age profiles. *J Public Econ* 92:1416–1440
- Berlinski S, Galiani S, Gertler P (2009) The effect of pre-primary education on primary school performance. *J Public Econ* 93:219–234
- Bettinger E, Kremer M, Saavedra JE (2008) Are educational vouchers only redistributive? *Econ J* 120(546):F204–F228
- Bishop J (2004) Drinking from the fountain of knowledge: student incentives to study and learn externalities, information problems, and peer pressure. Center for Advanced Human Resource Studies Working Paper
- Bitler MP, Thurston D, Penner EK, Hoynes HW (2013) Distributional effects of a school voucher program: evidence from New York City. NBER Working Paper No. 19271, National Bureau of Economic Research, Cambridge, MA
- Bjorklund A, Clark M, Edin P-A, Frederiksson P, Krueger A (2005) The market comes to education in Sweden. Russell Sage Foundation, New York
- Blimpo MP, Evans DK, Lahire N (2011) School-based management and educational outcomes: lessons from a randomized field experiment. University of Oklahoma, Oklahoma
- Bold T, Kimenyi M, Mwabu G, Ng'ang'a A, Sandefur J (2013) Interventions and institutions: experimental evidence on scaling up education reforms in Kenya. Center for Global Development Working Paper 321
- Card D, Krueger A (1992) Does school quality matter? Returns to education and the characteristics of public schools in the United States. *J Polit Econ* 100(1):1–40
- Carneiro P, Heckman J (2003) Human capital policy. In: Krueger A, Heckman J (eds) *Inequality in America: what role for human capital policies?*. MIT Press, Cambridge
- Carneiro P, Meghir C, Pary M (2012) Maternal education, home environments, and the development of children and adolescents. *J Am Econ Assoc* 11:123–160
- Chay K, McEwan P, Urquiola M (2005) The central role of noise in evaluating interventions that use test scores to rank schools. *Am Econ Rev* 95(4):1237–1258
- Clark D (2010) Selective schools and academic achievement. *BE J Econ Anal Policy Adv* 10(1)
- Coleman JS, Campbell EQ, Hobson CJ, McPartland J, Mood AM, Weinfeld FD, York RL (1966) Equality of educational opportunity. US Department of Health, Education and Welfare, Office of Education
- Cristia J, Ibararan P, Cueto S, Severin E (2012) Technology and child development: evidence from the one laptop per child program. IDB Working Paper No. 304, Inter-American Development Bank, Washington, DC

- Cullen J, Jacob B, Levitt S (2005) The effect of school choice on student outcomes: an analysis of the Chicago public schools. *J Public Econ* 89(5–6):729–760
- Cullen JB, Jacob BA, Levitt SD (2006) The effect of school choice on student outcomes: evidence from randomized lotteries. *Econometrica* 74(5):1191–1230
- Cunha F, Heckman J (2007) The technology of skill formation. *Am Econ Rev* 97(2):31–47
- Currie J, Thomas D (1995) Does head start make a difference? *Am Econ Rev* 85(3):341–364
- Das J, Pandey P, Zajonc T (2006) Learning levels and gaps in Pakistan. World Bank Policy Research Working Paper No. 4067. World Bank, Washington, DC
- Das J, Dercon S, Habyarimana J, Krishnan P, Muralidharan K, Sundararaman V (2013) When can school inputs improve test scores? *Am Econ J Appl Econ* 5(2):29–57
- Deaton A (2010) Instruments, randomization, and learning about development. *J Econ Lit* 48:424–455
- Dhaliwal I, Dufo E, Glennerster R, Tulloch C (2011) Comparative cost-effectiveness with applications for education. Abdul Latif Jameel Poverty Action Lab. <http://economics.mit.edu/files/6959>
- Dobbie W, Fryer R (2011) Exam high schools and academic achievement: evidence from New York City. NBER Working Paper No. 17286, National Bureau of Economic Research, Cambridge, MA
- Duflo E (2001) Schooling and labor market consequences of school construction in Indonesia: evidence from an unusual policy experiment. *Am Econ Rev* 91(4):795–813
- Duflo E, Dupas P, Kremer M (2011) Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in Kenya. *Am Econ Rev* 101(5):1739–1774
- Duflo E, Hanna R, Ryan S (2012) Incentives work: getting teachers to come to school. *Am Econ Rev* 102(4):1241–1278
- Epple D, Romano RE, Urquiola M (2015) School vouchers: a survey of the economics literature. *J Econ Lit* (forthcoming)
- Fredriksson P, Ockert B, Oosterbeek H (2012) Long term effects of class size. *Quart J Econ* 128(1):249–285
- Fredriksson P, Ockert B, Oosterbeek H (2014) Inside the black box of class size: mechanisms, behavioral responses, and social background. Stockholm University, Stockholm
- Friedman M (1955) The role of government in education. In: Solo R (ed) *Economics and the public interest*. Trustees of Rutgers College, Rutgers
- Friedman M (1962) *Capitalism and freedom*. University of Chicago Press, Chicago
- Fuller B, Clarke P (1994) Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy. *Rev Educ Res* 64:119–157
- Gallego F (2006) Voucher school competition, incentives, and outcomes: evidence from Chile. Massachusetts Institute of Technology, Massachusetts
- Glewwe P (2002) Schools and skills in developing countries: education policies and socioeconomic outcomes. *J Econ Lit* 40(2):436–482
- Glewwe P, Kremer M (2006) Schools, teachers, and education outcomes in developing countries. In: Hanushek E, Welch F (eds) *Handbook of the economics of education*. Elsevier, Oxford
- Glewwe P, Kremer M, Moulin S, Zitzewitz E (2004) Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya. *J Dev Econ* 74:251–268
- Glewwe P, Ilian N, Kremer M (2010) Teacher incentives. *Am Econ J Appl Econ* 2(3):205–227
- Grantham-McGregor S, Powell C, Walker S, Chang S, Fletcher P (1994) The long term follow-up of severely malnourished children who participated in an intervention programme. *Child Dev* 65(2):428–439
- Hanushek E (1986) The economics of schooling: production and efficiency in public schools. *J Econ Lit* 24(3):1141–1177
- Hanushek E (1994) Money might matter somewhere: a response to Hedges, Laine, and Greenwald. *Educ Res* 23(5):5–8
- Hanushek E (1995) Interpreting recent research on schooling in developing countries. *The World Bank Research Observer* (August)
- Hanushek E (1996) The productivity collapse in schools. In: Fowler W (ed) *Developments in school finance*. National Center for Education Statistics, Washington, DC
- He F, Linden L, Margaret M (2009) A better way to teach children to read? Evidence from a randomized control trial. Columbia University, Columbia
- Heckman J (2000) Causal parameters and policy analysis in economics: a twentieth century retrospective. *Q J Econ* 115(1):45–97

- Heckman J, Pinto R, Savelyev PA (2012) Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. NBER Working Paper No. 18581, National Bureau of Economic Research, Cambridge, MA
- Hedges L, Laine R, Greenwald R (1994) Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educ Res* 23(3):5–14
- Hoxby C (2000) Does competition among public schools benefit students and taxpayers? *Am Econ Rev* 90(5):1209–1238
- Hoxby C (2002) School choice and school productivity (or could school choice be a tide that lifts all boats?). NBER Working Paper 8873, National Bureau of Economic Research, Cambridge, MA
- 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* 90:1477–1503
- Imbens G, Lemieux T (2008) Regression discontinuity designs: a guide to practice. *J Econom* 142(2):615–635
- Jackson CK (2010) Do students benefit from attending better schools? Evidence from rule based student assignments in Trinidad and Tobago. *Econ J* 120(549):1399–1429
- Jamison D, Searly B, Galda K, Heyneman S (1981) Improving elementary and mathematics education in Nicaragua: an experimental study on the impact of textbooks and radio education. *J Educ Psychol* 73(4):556–567
- Jensen R (2010) Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *Quart J Econ* 127(2):753–792
- Jensen R (2012) The perceived returns to education and the demand for schooling. *Quart J Econ* 125(2):515–548
- Kingdon G (1996) The quality and efficiency of private and public education: a case study of urban India. *Oxford Bull Econ Stat* 58(1):57–82
- Kremer M, Miguel E (2004) Identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 72(1):159–217
- Kremer M, Muralidharan K (2006) Public and private schools in rural India. Harvard University, Harvard
- Kremer M, Miguel E, Thornton R (2009) Incentives to learn. *Rev Econ Stat* 91(3):437–456
- Krueger A (1999) Experimental estimates of education production functions. *Quart J Econ* 114(2):497–532
- Krueger A (2003) Economic considerations and class size. *Econ J* 485:F34–F63
- Krueger A, Zhu P (2004) Another look at the New York City voucher experiment. *Behav Sci* 47(5):658–698
- Laboratorio Latinoamericano de Evaluacion de la Calidad de la Educacion (2001) Primer estudio internacional comparativo sobre lenguaje, matematica, y factores asociados, para alumnos del tercer y cuarto grado de la educacion basica. UNESCO, Santiago de Chile
- Lee D, Lemieux T (2010) Regression discontinuity designs in economics. *J Econ Lit* 48:281–355
- Levin H, McEwan P (2001) Cost effectiveness analysis. Sage Publications, Thousand Oaks
- Li T, Han L, Rozelle S, Zhang L (2010) Cash incentives, peer tutoring, and parental involvement: a study of three educational inputs in a randomized field experiment in China. Stanford University, Stanford
- Linden L (2008) Complement or substitute? The effect of technology on student achievement in India. Columbia University, Columbia
- Ludwig J, Miller DL (2007) Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quart J Econ* 122(1):159–208
- MacLeod WB, Urquiola M (2012) Anti-lemons: school reputation and educational quality. IZA Discussion Paper No. 6805. Institute for the Study of Labor, Bonn
- MacLeod WB, Urquiola M (2013) Competition and educational productivity: incentives writ large. In: Glewwe P (ed) Education policy in developing countries. University of Chicago Press, Chicago
- MacLeod WB, Urquiola M (2015) Reputation and school competition. *Am Econ Rev* (forthcoming)
- Malamud O, Pop-Eleches C (2011) Home computer use and the development of human capital. *Quart J Econ* 126(2):987–1027
- Mayer D, Peterson P, Myers D, Tuttle C, Howell W (2002) School choice in New York City after 3 years: an evaluation of the school choice scholarships program: Final Technical Report, Mathematica Policy Research, Inc
- Mbiti I, Lucas A (2009) Access, sorting, and achievement: the short-run effects of free primary education in Kenya. Southern Methodist University, Technical Report
- McCrary J (2008) Manipulation of the running variable in the regression discontinuity design: a density test. *J Econom* 142(2):698–714

- McEwan P (2013) Improving learning in primary schools of developing countries: a meta-analysis of randomized experiments. *Rev Educ Res*. doi:10.3102/0034654314553127
- McEwan P, Carnoy M (2000) The effectiveness and efficiency of private schools in Chile's voucher system. *Educ Eval Policy Anal* 22(3):213–239
- McEwan P, Urquiola M, Vegas E (2008) School choice, stratification, and information on school performance. *Economia* 8(2):1–42
- McMillan P (2005) Competition, incentives, and public school productivity. *J Public Econ* 89:1131–1154
- Meyer B (2005) Natural and quasi-experiments in economics. *J Bus Econ Stat* 13(2):151–161
- Mo D, Swinnen J, Zhang L, Yi H, Qu Q, Bozwell M, Rozelle S (2012) Can one laptop per child reduce the digital divide and educational gap? Evidence from a randomized experiment in migrant schools in Beijing. *Rural Education Action Project Working Paper No. 233*, Stanford University
- Muralidharan K (2013) Priorities for primary education policy in India's 12th 5-year plan. University of California, San Diego
- Muralidharan K, Sundararaman V (2010) Contract teachers: experimental evidence from India. University of California, San Diego
- Muralidharan K, Sundararaman V (2011) Teacher performance pay: experimental evidence from India. *J Polit Econ* 119(1):39–77
- Muralidharan K, Sundararaman V (2013) The aggregate effect of school choice: evidence from a two-stage experiment in India. University of California, San Diego
- Neal D (2009) Private schools in education markets. In: Berends M, Springer M, Balou D, Walberg H (eds) *Handbook of research on school choice*. Routledge, New York
- Newhouse D, Beegle K (2006) The effect of school type on academic achievement: evidence from Indonesia. *J Hum Resour* 46(2):529–557
- Nguyen T (2008) Information, role models, and perceived returns to education: experimental evidence from Madagascar. Massachusetts Institute of Technology, Massachusetts
- Oster E, Millett B (2011) Do call centers promote school enrollment? Evidence from India. University of Chicago, Chicago
- Patrinos H, Barrera-Osorio F, Guaqueta J (2009) The role of and impact of public-private partnerships in education. World Bank, Washington, DC
- Pop-Eleches C, Urquiola M (2013) Going to a better school: effects and behavioral responses. *Am Econ Rev* 103(4):1289–1324
- Pradhan M, Suryadarma D, Beatty A, Wong M, Gaduh A, Alisjahbana A, Artha RP (2013) Improving educational quality through enhancing community participation: results from a randomized field experiment in Indonesia. *Am Econ J Appl Econ* 6(2):105–126
- Pritchett L (2003) Educational quality and costs: a big puzzle and five possible pieces. Harvard University, Harvard
- Reback R, Rockoff J, Schwartz H (2014) Under pressure: job security, resource allocation, and productivity in schools under no child left behind. *Am Econ J* 6(3):207–241
- Rockoff J (2009) Field experiments in class size from the early twentieth century. *J Econ Perspect* 23(4):211–230
- Schanzenbach D (2007) What have researchers learned from project star? In: Loveless T, Hess F (eds) *Brookings papers on education policy*. Brookings Institution, Washington, DC
- Somers M-A, McEwan P, Willms D (2004) How effective are private schools in Latin America? *Comp Educ Rev* 48(1):48–69
- Thistlewaite D, Campbell D (1960) Regression discontinuity analysis: an alternative to the ex-post facto experiment. *J Educ Psychol* 51(6):309–317
- Todd P, Wolpin K (2003) On the specification and estimation of the production function for cognitive achievement. *Econ J* 113:F2–F33
- Urquiola M (2006) Identifying class size effects in developing countries: evidence from rural Bolivia. *Rev Econ Stat* 88(1):171–177
- Urquiola M, Calderon V (2006) Apples and oranges: educational enrollment and attainment across countries in Latin America and the Caribbean. *Int J Educ Dev* 26:572–590
- Urquiola M, Verhoogen E (2009) Class-size caps, sorting, and the regression discontinuity design. *Am Econ Rev* 99(1):179–215
- Vegas E, Petrow J (2008) Raising student learning in Latin America: the challenge for the 21st century. World Bank, Washington, DC
- World Bank (2004) Making services work for poor people. World Bank, Washington, DC