



Incentives and Accountability in Education: A Review

August 2014

This publication was produced for review by the United States Agency for International Development. It was prepared by RTI International and NORC at the University of Chicago.

Incentives and Accountability in Education: A Review

August 2014

Prepared for:

Office of Education
Bureau for Economic Growth, Education, and Environment (E3)
United States Agency for International Development (USAID)

This paper is a synthesis of two analytical studies prepared for USAID:

“Incentives and Accountability in Education: A Literature Review”
RTI International
Measurement and Research Support to Education Strategy Goal 1, RTI Task Order 20,
Activity 3, Education Data for Decision Making (EdData II)
June 2014

and

“Incentives Systems Literature Review”
NORC at the University of Chicago
Reading and Access Evaluations
May 2014

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Authors

This report is based on two original papers written by RTI International and NORC at the University of Chicago. The paper by RTI International was written by Gustavo Arcia. The paper by NORC at the University of Chicago was written by Alejandro Ome, Vi-Nhuan Le, and Alicia Menendez. Sections 1 and 2 are based on the RTI paper, Sections 3 and 4 are based on the NORC paper, and Sections 5 and 6 are based on both papers. The combination of the RTI and NORC papers was led by RTI International in collaboration with NORC at the University of Chicago.

Table of Contents

List of Figures.....	v
List of Tables.....	v
Abbreviations	vi
1 Introduction.....	1
1.1 Is there a need for incentives?.....	1
1.2 The need for a review of incentives and accountability in education	2
2 A Framework for Incentives and Accountability in Education	3
2.1 Do incentives matter? Incentives and accountability and the provision of public education.....	3
3 Incentives for Teachers and Their Effects on Teacher Accountability	5
3.1 The results of pay for performance programs are mixed	8
3.2 Teacher incentives in the United States have not been very successful.....	10
3.3 Improving teacher attendance may affect student outcomes.....	11
3.4 Incentives that attract better teachers may be more effective than those trying to change teachers' behavior	12
4 Incentives for Parents and Students.....	13
4.1 Financial incentives can have significant impact on the attendance or enrollment rates of eligible children	17
4.2 CCT and FFE have null to weakly positive impact on learning outcomes	17
4.3 If you pay them, they will learn: Pay for performance incentives for students can be effective	18
4.4 Financial incentives for students do not decrease intrinsic motivation and their effects can be sustained over time.....	19
4.5 Larger incentives do not necessarily lead to larger impacts.....	20
4.6 Financial incentives can have unintended consequences	20
5 School Report Cards and School Accountability	20
5.1 School report cards are effective for informing parents and making teachers accountable.....	21

5.2	Scorecards designed with parent participation improve school management.....	22
5.3	Parent-controlled school grants tend to be effective	23
5.4	Contract teachers can produce better results	23
6	Incentives and Accountability in Education: Key Lessons Learned	24
6.1	Incentives for teachers can improve test scores	24
6.2	Scaling up the experimental evidence on incentives and accountability in education requires attention to systemic issues	26
	References.....	28

List of Figures

Figure 1.	Long and short routes of accountability in public education.....	4
-----------	--	---

List of Tables

Table 1.	Overview of studies focusing on teachers.....	6
Table 2.	Incentives aimed at parents and students.....	13
Table 3.	Incentives to improve accountability	21

Abbreviations

CCT	conditional cash transfer
DID	differences in differences (modeling)
E3	USAID Bureau for Economic Growth, Education, and Environment
EdData II	USAID Education Data for Decision Making project
EDUCO	Educación con Participación de la Comunidad [model, El Salvador]
EGRA	Early Grade Reading Assessment
EMIS	education management information system
FFE	food for education
OECD	Organisation for Economic Co-operation and Development
PEC	Programa de Escuelas de Calidad [Quality Schools Program, Mexico]
PFP	pay for performance
PISA	Programme for International Student Assessment
PSM	propensity score matching
RCT	randomized controlled trial
RTI	RTI International (trade name of Research Triangle Institute)
SD	standard deviation
SFP	school feeding program
SMC	school management council
TAP	[Chicago] Teacher Advancement Program
THR	take-home rations
TTI	Talent Transfer Initiative
USAID	United States Agency for International Development

1 Introduction

In 2011, USAID issued a new four-year education strategy to ensure that investments are coordinated to achieve measurable and sustainable educational outcomes. To this end, USAID selected three strategic goals:

- Goal 1: Improved reading skills for 100 million children in primary grades by 2015;
- Goal 2: Improved ability of tertiary and workforce development programs to produce a workforce with relevant skills to support country development goals by 2015; and
- Goal 3: Increased equitable access to education in crisis and conflict environments for 15 million learners by 2015.

Reaching the first goal will require the optimization of educational resources—finance, teachers, school management—and the use of institutional and personal incentives that will motivate education practitioners and students to improve their commitment and performance. The objective of this review is to learn from the successes and failures of implementing incentive and accountability mechanisms in education worldwide, in order to draw lessons that can be applied specifically to developing countries. The approach to the review is simple: to examine some recurrent threads or themes that suggest successful approaches and to typify to any extent possible the circumstances under which incentives and accountability succeed or fail.

A growing literature documents the effects of providing incentives in schools. These incentives can target teachers, school principals, and other administrators, as well as students and their parents. Inversely, these incentives become mechanisms for school and teacher accountability, since they require the measurement and reporting of learning outcomes and school performance. This report reviews the recent literature on incentives and accountability in education with the purpose of assisting practitioners and stakeholders with a frame of reference for operational use.

Mimicking market forces, incentives are used as prizes for improving performance above the existing threshold under the simple notion that school structures are not motivating enough for teachers, students, or parents to automatically increase their dedication and improve student learning. How well incentives improve learning depends on the mechanism itself, the institutional and cultural framework, the ways in which performance is measured and reported, and the willingness of stakeholders to apply rewards and sanctions.

1.1 Is there a need for incentives?

In broad terms, an *incentive* is the promise of a reward for doing a good job, and *accountability* is the acceptance of the rewards and sanctions that come with measuring and reporting the results obtained. Since accountability implies the evaluation of the results obtained, incentives and accountability are just different expressions of the same concept. In principle, producing student learning is the default purpose of a teaching job, and regardless of incentives, the evaluation of performance leading to accountability is the norm. That said, every institutional framework has built-in incentives and disincentives. The application of additional incentives above and beyond

what exist at the school level is, in the end, a simple recognition that the inherent inducements are insufficient to spark teacher and student effort.

1.2 The need for a review of incentives and accountability in education

This report summarizes information on relevant analyses on incentives that can be applied in developing countries, and their implications for school accountability. Recent results on controlled experiments on incentives and accountability, and on statistical analyses of international test scores within the context of school-based management, have yielded important information that can be used for designing similar programs in other countries. However, the number of these studies is relatively small.

Most current efforts on incentives and accountability rely on two different approaches: (1) participatory school-based management (Hanushek, Link, and Woessman 2013), and (2) trust-based institutional incentives (Di Gropello 2004; Arcia et al. 2011). Participatory school-based management is a practical and tested way of localizing teacher incentives and of enforcing school-level accountability in countries where institutions are far from perfect. Participatory school-based management allows parents to link teacher rewards or sanctions to good or bad performance. Parent-enforced accountability for teacher or school performance can be thought of as the manifestation of an educational system that does not enforce sanctions or, if it does, it fails to enforce them in a fair and timely manner. Institutional incentives based on trust are more prevalent in high-performing countries, where education systems have well-established processes for measuring and reporting results, and with the application of rewards attached to them.

In the case of trust-based accountability, systems require a functioning legal process that is charged with enforcement of fair and timely consequences. The institutional realities of many developing countries suggest that the rewards or sanctions tied to accountability require a careful alignment of personal incentives with institutional managerial incentives, with education legislation, and with everyday school management. As Crouch and Winkler (2007) point out, accountability is a necessary but insufficient condition for good governance, and implementing accountability requires paying attention to school management, teacher quality, and symmetries in information among all stakeholders.

This review is organized as follows: Section 2 reviews the concept and the definitions of accountability in education. Section 3 reviews the implementation of incentives for teachers, and section 4 reviews the incentives for parents and students. Section 5 examines the incentives applied to schools and the results for school accountability, and section 6 summarizes the results so far on incentives and accountability in education, and the implications for scaling them up in developing countries.

2 A Framework for Incentives and Accountability in Education

In education, incentives and accountability can be interpreted as tools for improving the effectiveness of teachers and schools, with the ultimate goal of improving student outcomes. Stakeholders can provide for additional rewards for good performance, or sanctions for bad performance. As such, incentives and accountability require (1) defining the scale of rewards and the sanctions attached to specific results, and (2) implementing the rewards or sanctions triggered by the results (Hooge, Burns, and Wilkoszewski 2012).

Defining incentives is important, as the definitions send signals to teachers and students about what the rewards are for good performance or behavior and, indirectly, whether the rewards are worth their effort. The defining process also allows stakeholders to discuss what it is possible to do with existing resources. Implementing rewards and sanctions is important because without implementation there is no accountability, just a recounting of events.

2.1 Do incentives matter? Incentives and accountability and the provision of public education

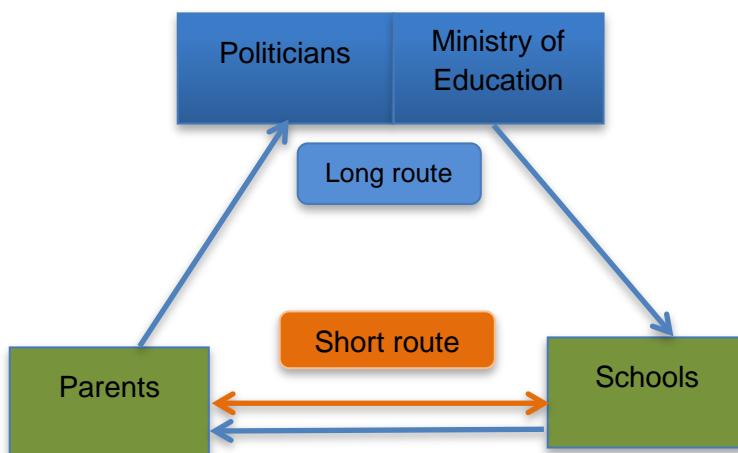
A description of the built-in incentives and school accountability in public education is found in the World Bank's *World Development Report 2004*, which deals with the efficiency and effectiveness of service provision for the poor (World Bank 2003). This report is important because it developed the conceptual framework for accountability in education that is currently being used by most donors, and that now is the reference point for discussing incentives and accountability in many controlled experiments around the world. At the heart of this framework is the use of accountability as a key policy instrument for improving student performance and increasing student learning.

In a public school system, the government, as the main stakeholder, delegates to schools the provision of education services, and finances teacher pay, infrastructure, and operational expenses. Schools deliver educational services under agreed-upon performance standards, informing the government and society about the schools' performance. The government uses these performance results to assign rewards and sanctions to the education providers.

This framework for the provision of public education leads to *long and short routes for incentives and accountability*, in which the process of assigning rewards and sanctions for good and bad performance is filtered by the ability of politicians and parents to intervene in the process. As shown in Figure 1, parents can participate in a long route of accountability by expressing their concerns to politicians, who in turn use the educational structures to transform the parents' voice into action and eventual changes in education delivery. In this case, politicians manage the incentives and the sanctions. This process can take years, however, even in high-performing, mature democracies.

Figure 1. Long and short routes of accountability in public education

(Adapted from World Bank 2003)



The short route to accountability is one in which parents are given enough power (legal, managerial, or purchasing power) to affect school behavior directly and produce a more rapid change in education delivery *with the clear purpose of improving learning outcomes among students*. In the short route to accountability, parents have a more direct role in defining and applying incentives, in evaluating results, and in administering rewards and sanctions. Obviously, maximal accountability is achieved when both the long and short routes are aligned with each other.

The framework for accountability shown in Figure 1 is instructive, but it needs some additional context. For incentives and accountability to work, schools must be given clear objectives and sufficient resources, and teachers must have adequate technical capabilities and personal incentives. Otherwise parental power cannot buy anything. Similarly, the expectations about the education system's performance should be aligned with the conditions on the ground; otherwise, schools and teachers will be asked for results that they cannot give, even if they want to. Schools and teachers can respond to incentives, and can also be made accountable, if they have clear objectives and an internally coherent operational strategy with clear expectations about results. In turn, parents should know what accounts to ask for, how to ask for accounts, and how to use clearly defined incentives and clear paths to accountability as a way to fix problems and motivate teachers and students.

An increase in complexity in school accountability tends to create unintended effects on school performance. The pressure from stakeholders to receive clear accounts tends to produce a narrowing of teaching and learning ("teaching to the test"), a narrowing of the curriculum in order to focus on what is tested, and an emphasis on failure, where stakeholders associate accountability with sanctions and little else. Moreover, teachers tend to pay more attention to those students whose performance is just under the threshold of proficiency, since that is where

the potential performance gains are highest (Hout and Elliott, 2011). These unintended effects may increase if accountability is nested within rigid bureaucratic structures.

In contrast, accountability based on a more nuanced set of rules, applied within a context of personal negotiation—as typically happens at the school level—can be more effective in correcting simple mistakes and in motivating teachers. Test-based accountability in schools with operational autonomy should rely less on bureaucratic rules and more on interaction among teachers, parents, and principals, using test results as information for the negotiations (Crouch and Winkler 2007).

It is worth highlighting that, in this framework, parents are seen as agents who are already taking every possible action to improve their children’s education, which suggests that providing incentives for parents may not have an important effect on students’ outcomes. However, a large body of literature documents that providing cash transfers conditioned on school attendance does have important effects on attendance, which suggests that there may actually be some room for incentives to affect parental behavior, too. Similarly, it is not necessarily guaranteed that students are sufficiently motivated to put forth their best effort toward learning. This can be achieved by encouraging students’ intrinsic motivation and by providing external rewards for high-quality academic performance. Along these lines, in this report we also review the literature on incentives for parents and students, in order to provide a comprehensive view of this matter in the context of basic education.

3 Incentives for Teachers and Their Effects on Teacher Accountability

Financial incentives to teachers are subject to basic scrutiny from parents and government: why should society give additional pay to someone to do the job they were hired to do in the first place? The simplest answer is that education systems use money to induce teachers to work around the disincentives of bureaucratic procedures, bad education laws, or union agreements that foster low performance. However, the judgments teachers make to exchange rewards for better performance are not well understood because of the covariate effect of the curricula, the teaching materials, their own professional training, and the effect of the incentives themselves.

Vegas and Umansky (2005) argued that the structure of teacher incentives is much broader than just financial, being more of a basket of goods that, besides more money, may include recognition and prestige in the community, job stability, pension benefits, professional growth, good school facilities, and personal growth. A recent review of the available evidence on incentive structures supported this assertion (World Bank, 2013).

Among the best known structures are pay for performance (PFP) incentives, which are the most commonly analyzed teacher incentive and are closely tied to test-based accountability (e.g., Glewwe et al. 2010; Muralidharan and Sundararaman 2011). Incentives can also be lump-sum awards to attract more qualified teachers (Steele, Murnane, and Willett 2010; Glazerman et al. 2013). Furthermore, job stability, or more specifically, threats to it, can also be thought as an

incentive, and is the main instrument behind tenure reforms and short-term teacher contracts (Duflo et al., 2012; Dee and Wyckoff, 2013).

Table 1 presents a summary of the papers reviewed for this section. The first column names the authors of the study and the location of the intervention. In the second column we describe the structure of the program; for PFP interventions we highlight whether incentives were provided as a function of teacher individual performance, or were based on group performance. For interventions other than PFP, we describe the main characteristics of the incentive(s) provided. In the last column we show the evaluation technique used and the main findings of each study.

Table 1. Overview of studies focusing on teachers

Study and location	Structure	Evaluation method; results
Alvarado, Duarte, and Nielson 2013, Chile	Scholarship for highly qualified high school graduates to join a teaching school	Regression discontinuity design; increased by 8 percentage points the probability that eligible high school graduates would go to teaching school
Contreras and Rau 2012, Chile	Pay for performance: Group	Weighted differences in differences (DID); standardized test results increased between 0.16 standard deviations (SD) and 0.24 SD for math and 0.14 SD and 0.26 SD for language for students in fourth, eighth, and tenth grades
Cueto et al. 2008, Peru	Incentives based on teacher attendance, both individual and at the school level	Propensity score matching (PSM); teacher attendance increased by 17 days per year. No effect on students' outcomes ¹
Dee and Wyckoff 2013, USA	Pay for performance: Individual teacher performance affects the probability of being dismissed or gaining a bonus	Regression discontinuity design ² ; teachers at risk of being dismissed were 11 percentage points more likely to leave voluntarily. Teachers at risk of being dismissed increased their test scores by 0.27 SD and teachers who had a good chance of gaining a bonus increased their test scores by 0.24 SD
Duflo, Hanna, and Ryan 2012, India	Pay for performance: Incentives linked to teacher attendance and monitoring of classroom practices	Randomized controlled trial; reduced teacher absenteeism by 21 percentage points. Increased student test scores by 0.21 SD in math and 0.16 SD in language
Fryer 2011b, USA	Pay for performance: Group; monetary	Randomized controlled trial; no significant effects for elementary and high schools. For middle school, significant effects of -0.03 SD in reading and -0.05 SD in math

¹ Propensity score matching is a statistical technique that simulates randomization in nonrandom samples by selecting cases that have covariates comparable with those of a control group. See Ch. 7 in Gertler et al (2011).

² Regression discontinuity design refers to a quantitative method used to assess statistical significance in programs that use a cutoff point on an exogenous variable to determine eligibility to the program. See Ch. 6 in Angrist and Pischke (2009).

Study and location	Structure	Evaluation method; results
Fryer et al. 2012, USA	Pay for performance: Individual and group monetary incentives framed as losses (teachers were paid in advance and asked to give back the money if their students did not improve sufficiently)	Randomized controlled trial; between 0.2 SD and 0.4 SD higher test scores than controls in math, no effects on reading, no major differences between individual and group incentives. No effects for treatment arm where incentives were framed as gains
Glazerman and Seifullah 2012, USA	Pay for performance: Individual and group incentives	Randomized controlled trial and PSM; no impact on student achievement, teacher attitudes, or school climate. Teacher retention rates 12 percentage points higher than in comparison schools
Glazerman et al. 2013, USA	Incentive for high-performing teachers to transfer to low-performing schools	Randomized controlled trial; transferred teachers increased student achievement by 0.22 SD in math and 0.25 SD in reading
Glewwe, Ilias, and Kremer 2010, Kenya	Pay for performance: Group; in-kind incentives	Randomized controlled trial; treated schools scored 0.14 SD higher on a general exam. No effect on non-incentivized exams. No effect on incentivized exam 1 year after program ended
Lavy 2002, Israel	Pay for performance: Group incentives	Regression discontinuity design; higher test scores by 0.12 SD
Lavy 2009, Israel	Pay for performance: Individual incentives	Regression discontinuity design and measurement error; higher math test scores by 14% and 4% in math and English
Martins 2009, Portugal	Pay for performance: Individual incentives based on students' test scores and other outcomes. Teachers competing for a few awards	Difference-in-difference; decrease in standardized national exams and inflation in non-standardized school grades. Competition within the school led to non-cooperation among teachers
Muralidharan and Sundararaman 2011; Muralidharan 2012, India	Pay for performance: Individual; monetary	Randomized controlled trial; effects for treatment group: 2-year follow-up: 0.33 SD in math and 0.24 SD in language. 5-year follow-up: 0.54 SD in math and 0.35 SD in language
	Pay for performance: Group; monetary	Randomized controlled trial; effects for treatment group: 2-year follow-up: 0.22 SD in math and 0.09 SD in language. 5-year follow-up: Results were not significant
Springer et al. 2010, USA	Pay for performance: Individual; monetary	Randomized controlled trial; no significant effect on test scores. No significant effect on teachers' behavior
Springer et al. 2012, USA	Pay for performance: Group; monetary	Randomized controlled trial; no significant effect on test scores. No significant effect on teaching attitudes and practices
Steele, Murnane, and Willett 2010, USA	Fellowship for highly qualified novice teachers to go to low-performing schools	Regression discontinuity design; 28% increased likelihood that eligible teachers would go to low-performing schools

SD = Standard deviation

The main methodological problem with analyzing education incentives is that it is difficult to obtain data of sufficient quantity and quality to discern the impact of a given incentive and correctly attribute it to each of the different participants in the education process (Gertler et al. 2011). To help with this methodological problem, as noted in Table 1 above, some investigators have started implementing randomized controlled trials (RCTs), which should be better for determining the exact impact of rewards and sanctions on teacher and student behavior. This method of impact evaluation is one of the most commonly used approaches in Table 1. From that evidence some basic conclusions can be drawn, as elaborated in the subsections that follow.

3.1 The results of pay for performance programs are mixed

The most convincing evidence about the effectiveness of Pay for Performance (PFP) programs comes from randomized controlled trials, where participants are chosen at random and their performance compared against a control group. The additional pay received by teachers is a function of the student test results. In the international arena, perhaps the most cited works are by Glewwe, Ilias, and Kremer (2010) and Muralidharan and Sundararaman (2011).

Glewwe, Ilias, and Kremer (2010) analyzed a teacher incentive program in Kenya. The program lasted for two years and consisted of in-kind rewards (flashlights, briefcases, tea sets, among others) based on student performance averaged at the school level. Reward values were in the range of US\$26 to US\$51. In-kind teacher incentives raised student scores by 0.14 SD on exams that were directly tied to the incentives. Most of the *score gains were linked to a higher number of students taking the test*; the program penalized teachers for every student that did not take the test. However, the *gains in learning outcomes happened only in the areas being tested, while learning in other parts of the curriculum did not increase*. Test score gains were higher on multiple-choice exams than on fill-in-the-blank questions, indicating that students and teachers had become adept at test taking. Teacher absenteeism did not change; instead, what changed was the time teachers devoted to the areas that were linked to the test. Overall, some of the results were illustrative of unintended consequences of PFP programs, including evidence of teaching to the test and a narrowing of the curriculum.

In a randomized study of school and teacher incentives in Andhra Pradesh, India, Muralidharan and Sundararaman (2011) found positive effects of individual incentives on students' test scores in the academic subjects covered by the incentives and in some of the subjects not covered by the incentives. The authors found that program impact was less dependent on teacher attendance than on teacher effort once the teacher was in the classroom.

In a follow-up study in rural Andhra Pradesh, India, Muralidharan (2012) analyzed the impact of a five-year experiment with teacher pay linked to performance in public primary schools.³ Researchers tracked a cohort of primary school students for the five years of primary school. The experiment evaluated incentives based on both individual and group performance. The results showed that for the group with individual incentives, the program had positive, significant effects on learning outcomes for the covered topics, and positive spillover effects on learning outcomes

³ This was a follow-up study to Muralidharan and Sundararaman (2011).

for non-covered topics. Students served by performance pay programs had test scores that were 0.35 SD higher in language, and 0.54 SD higher in math, than the scores of students in the control group. These effects are roughly equivalent to one additional year of schooling, which is very significant. In sciences and social studies, which were not covered by the program, test scores increased by 0.52 and 0.3 SD higher than for control group students. These gains outside of the program suggest that the effect of performance pay on teacher effort went beyond “teaching to the test.” The effects for the teachers who were offered group incentives, on the other hand, were not significant in the five-year follow-up. Participating teachers would receive 500 rupees (US\$9.20) for every percentage point of average gain in the test scores of their students.

In addition to RCT evaluations, some rigorous non-experimental studies are worth attention. In general, these quasi-experimental works have shown positive results (Lavy, 2002 and 2009; Contreras and Rau 2012), but there is also some evidence of negative effects (Martins, 2009).

Lavy analyzed two incentive programs in Israel and found positive effects for both of them. The first program (Lavy 2002) evaluated a group incentive scheme implemented in a set of schools. This intervention granted financial rewards to schools and teachers according to the performance of the school as a whole on a series of students’ outcomes. Lavy used a regression discontinuity design to conclude that the program increased students’ average test scores by approximately 0.12 SD. In the second program Lavy (2009) analyzed the effect of individual teacher incentives on student test scores. The program was designed as a tournament between teachers in the same subject at the same school. Lavy found positive effects on test-taking rates, passing rates, and test scores. In particular, pass rates in math showed a gain of 10 percent in the regression coefficient for the average score for math and a gain of 4 percent for English.

Contreras and Rau (2012) analyzed a nationwide incentive program for teachers in Chile, where comparable schools competed for a fixed amount of money to be distributed among teachers in proportion to their workload. Awards were granted according to an index that included students’ test scores, retention rates, and other school performance measures. Privately-funded schools were the comparison group. Using a difference-in-difference model, the authors found that, after two years, test scores for students in the fourth, eighth, and 10th grades increased between 0.16 SD and 0.24 SD for math and 0.14 SD and 0.26 SD for language over the scores of students from private schools.

Martins (2009) analyzed a national teacher incentive policy implemented in Portugal that used a two-tiered pay scale, where teachers could go from the lower to the higher pay scale on the basis of their student’s test scores, teacher attendance, attendance at training sessions, and other aspects. The number of spots for promotion at the school level was limited, creating a climate of competition among teachers within the school. Martins found that students’ test scores declined as a consequence of the system and he also found some evidence of grade inflation. As a possible explanation for this decline, the author argued that the fact that the incentive scheme was constructed as a tournament among teachers in the same schools may have disrupted collaboration between them.

The evidence from randomized experiments on PFP incentives suggests that teachers may increase their efforts in subject areas that are being tested, or in areas directly linked to the rewards being offered. The evidence also suggests that when implementing PFP programs, one has to be careful and monitor potential sources of distortion—such as teaching to the test, and reduced teamwork among teachers. Individual incentives show promise because they tend to motivate teachers, at the risk of affecting the performance of teachers of subject areas not being tested. On the other hand, sharing rewards among all teachers based on the performance of a couple of subject areas may create free-rider problems, where the benefits may accrue to teachers who did not improve performance. In summary, PFP programs show promise, provided that they are designed and monitored carefully.

3.2 Teacher incentives in the United States have not been very successful

For the United States, the documented effects on incentives are less promising. Most studies that have been conducted on incentives for teachers have failed to find positive and significant effects (Springer et al. 2010; Springer et al. 2012; Fryer 2011b).

Springer et al. (2010) studied POINT, an individual incentive program conducted in Nashville, Tennessee, and found no impact on student test scores, nor much influence on teacher behavior. It seems that teachers thought that they could do little else to improve their performance. Yuan et al. (2012) analyzed POINT and two other experiments on teacher incentives, and found that teachers did not consider these programs motivating because they did not understand the program, they considered that test scores were not a good metric to measure teaching performance, they did not believe that personal efforts would impact student outcomes because of other influencing factors in student achievement, and they did not consider the bonus amount worth the additional effort.

Other more recent experiments in the United States have also failed to document positive effects on incentives for teachers. Fryer (2011b) found that teacher incentives did not seem to have any positive impact on student performance or change teacher behavior. Schools eligible for a bonus could receive up to US\$3,000 per teacher depending on student performance, to be paid out to teachers in a manner designed by each school. As a result, incentives could be for the group or for individuals. In elementary schools, the effects were negative but non-significant; in middle schools, the effects were negative and statistically significant; and in high schools, the effects were positive but non-significant. No effect was found on teacher attendance or retention. The author suggested that, of the several reasons that could explain why no positive effects were found, the most plausible explanations were that teachers did not know how to help their students, and that they could have spent too much time on inefficient teaching methods and practices, which could have had negative effects on student learning in the treatment schools.

Glazerman and Seifullah (2012) used data from the Chicago Teacher Advancement Program (TAP) to explore the impact of individual teacher incentives on student achievement, teacher attitudes and practices, and teacher retention rates. The program offered an average of US\$2,000

per teacher, based on measures of student achievement and classroom observations. School principals were also eligible for bonuses. The incentive had no effect on student achievement, teacher attitudes, or school climate, but teacher retention increased: TAP teachers were about 20% more likely to be in the same school three years later than teachers in comparison schools.

Although most performance-based incentives are constructed so that teachers receive bonuses (or not) once students' achievement has been observed, another possibility is to give the bonus at the beginning of the experiment, and ask teachers to return the money if students do not improve sufficiently. Framing PFP incentives as losses instead of gains may be effective, although the evidence is scant. Fryer et al. (2012) found that teachers who were given PFP bonuses at the beginning of the experiment had student test scores between 0.2 and 0.4 SD higher than those in the control group. If the bonuses were given at the end, just like in most incentive programs in the United States, the effect on test scores was non-significant. These findings suggest that focusing on loss aversion may be an effective approach for designing teacher incentives.

A more definitive success story in the United States is the highly controversial IMPACT, an evaluation system implemented in Washington, DC. This system ranks teachers according to scores based on student achievement and classroom observations, classifying them as Ineffective, Minimally Effective, Effective, or Highly Effective. Ineffective teachers are dismissed, Minimally Effective teachers are dismissed in the next period if they are not then classified as Effective or Highly Effective, and Highly Effective teachers gain a bonus. Dee and Wyckoff (2013) showed that Minimally Effective teachers tended to leave teaching voluntarily at higher rates than did teachers with a higher classification. Teachers who had been classified as Minimally Effective increased their test scores by 0.27 SD, but teachers who could increase their scores just enough to be classified as Highly Effective, and thereby earn a bonus, increased their test scores by 0.24 SD.

Why have PFP incentives in the United States shown significant percentages of failure relative to the rates of failure in developing countries? From an economist's point of view, the issue may be a simple one of marginal costs versus marginal benefits, where the marginal positive impact on teacher welfare brought about by a potential bonus may be lower than the cost of the additional effort that teachers may have to make to increase test scores within an environment where teacher effort may not be the missing link in student learning. Teacher capabilities and student behavior may be factors that make incentives ineffective in this context. In contrast, in some developing countries, the impact of increasing some inputs, such as teacher attendance, may be sufficient to improve test scores.

3.3 Improving teacher attendance may affect student outcomes

The programs described above incentivized teachers based on outputs, such as students' test scores; but other interventions have framed incentives as a function of teacher inputs, such as teacher attendance. Using output-based incentives is more appealing to many practitioners than using inputs because of the better alignment between incentives and the outcome of interest. However, there are a few reasons why input-based incentives may be a better option in some

contexts. First, if teachers do not know what to do to improve student learning above existing levels, the best that a school can do may be to increase teacher attendance. Also, monitoring inputs may be less costly than monitoring outputs.

Duflo, Hanna, and Ryan (2012) evaluated incentives to boost teacher attendance in India. In treated schools, salary was made a function of attendance, which was photographically recorded. The evaluators also collected data on attendance through unannounced random visits in all schools. Teacher absenteeism was reduced by 21 percentage points, and students' test scores increased by 0.21 standard deviations in math and 0.16 standard deviations in language.

Cueto et al. (2008) analyzed a similar intervention in rural Peru. Participating schools were not randomized but were assigned to treatment and comparison groups. Attendance was monitored periodically, mostly by students' parents. Financial incentives were based on teacher attendance, both individual and at the school level. While the program increased teacher attendance by 17 days per year (roughly 13 percent of the baseline level), there was no clear effect on students' outcomes.

3.4 Incentives that attract better teachers may be more effective than those trying to change teachers' behavior

Two different types of incentive mechanisms can affect the composition of the teacher workforce. First, pay for performance schemes may attract highly qualified individuals (Lazear 2003). Second, signing bonuses and other lump-sum payments for joining the profession (or transferring to a particular type of school) are also intended to change teacher quality—not because teachers are attracted to the type of compensation system, but because of the one-time incentive.

These lump-sum incentives can target high-quality high school graduates so they choose to pursue a teaching degree (Alvarado et al. 2013), or they can target high-quality teachers so they go to low-performing schools (Steele et al. 2010; Glazerman et al. 2013).

To incentivize high-quality high school graduates to join the teaching profession, in 2010 the Chilean government granted scholarships to students with high test scores who chose a career in education. Alvarado et al. (2013) found that the number of high-scoring high school graduates choosing programs in education increased by 8 percentage points as a consequence of this program.

Other policies have focused on attracting high-quality teachers to low-performing schools. Between 2000 and 2002, the U.S. State of California granted fellowships to academically talented novice teachers who decided to go to low-performing schools. Steele et al. (2010) found that this program increased the likelihood that academically talented teachers would go to low-performing schools by 28 percentage points.

While lump sum awards seem to be effective for attracting better teachers (or at least more qualified individuals), evaluations of these programs on students' outcomes are hard to come across. An exception is the Talent Transfer Initiative (TTI), implemented in 10 cities in the

United States. This program offered a one-time bonus of US\$20,000 to high-performing teachers to transfer to low-performing schools in 10 school districts. Glazerman et al. (2013) found that transferred teachers increased student achievement by 0.22 standard deviations in math and 0.25 standard deviations in reading, by the second year of implementation in elementary schools, while no effects were found in middle schools. About 88 percent of TTI teachers stayed in the low-performing school for two years and about 60 percent stayed for at least three years.

4 Incentives for Parents and Students

Generally, incentives designed to motivate parents and students tend to be mostly financial (Table 2). For parents, there are two types of incentive structures: conditional cash transfers (CCTs) and food for education (FFE) programs. CCTs give money to poor households with school-age children contingent on their children’s attending school on a regular basis. FFE programs can be of two types: those that serve breakfast and/or lunch to students every day, and those that provide take-home rations (THR), where students receive a certain amount of food staples, conditional on their maintaining a specified attendance rate. For students, monetary rewards are given when they reach a predetermined threshold of academic achievement.

CCTs and FFEs have proven successful in increasing student attendance, but in few studies have these programs had an impact on *learning*, underscoring the fact that improved learning likely depends on other factors besides attendance, such as teacher quality, the availability of good pedagogical materials, and other factors. Most CCT and FFE programs are implemented as components of poverty-reduction programs. As such, they tend to exclude the measuring and reporting of learning outcomes; poverty programs are generally implemented by agencies outside of the education sector. The evidence on learning outcomes tend to be *a posteriori*, generally showing mixed results.

The empirical evidence reviewed in Table 2 shows that incentives aimed at parents and students tend to have a positive effect on attendance, but may have other impacts as well, as discussed below in the rest of this section.

Table 2. Incentives aimed at parents and students

Study and location	Structure	Evaluation method; results
Ahmed and del Ninno 2002, Bangladesh	Food for education: Take-home rations	Regression; school enrollment: girls increased by 44%, boys by 28%; in non-FFE schools, girls increased by 5.4% and boys by 0.1%; covariate-adjusted regression analysis suggests an increase between 7.9% and 8.4%. School attendance: 70% in FFE schools versus 58% in non-FFE schools. Dropout: 6% of FFE students dropped out versus 15% of non-FFE students

Study and location	Structure	Evaluation method; results
Alderman, Gilligan, and Lehrer 2010, Uganda	Food for education: Take-home rations and school feeding program	Randomized controlled trial; THR generally had the same impact as SFP. School enrollment: no impact by THR or SFP. Morning school attendance: no impact of THR/SFP on children aged 6-13, but a positive impact on children aged 10-17, ranging from 8-12 percentage points. Afternoon school attendance: THR/SFP had impact on children ages 6-9 and 10-17, but not 10-13; impact of approximately 14 percentage points. Grade repetition: THR/SFP decreased grade repetition by 0.099 and 0.115 percentage points, respectively
Angrist and Lavy 2009, Israel	Monetary incentives to individual students	Randomized controlled trial; increased certification rates of up to 8 percentage points
Attanasio, Fitzsimmons, and Gomez 2005, Colombia	Conditional cash transfer	Regression; prior to program, enrollments for 8- to 13-year-olds were 89% and 94% in urban and rural areas; for 14- to 17-year-olds, percentages were 54.4% and 72%; afterward, the percentages were 92.2%, 95.4%, 60.3%, and 77.3%
Attanasio, Meghir, and Santiago 2005, Mexico	Conditional cash transfer	Randomized controlled trial; focused on boys less than 10, aged 10 to 13, and older than 13; the program had an average effect of 3% on school enrollment of boys aged 6 to 17. Effect was much larger (around 7.5%) for older boys and virtually zero for boys younger than 9
Baird, McIntosh, and Ozler 2011, Malawi	Conditional cash transfer	Randomized controlled trial; overall attendance was 8 percentage points higher for CCT than control group. CCT outperformed control group on cognitive ability, math, and reading comprehension; CCT impact was 0.14 SD in reading comprehension, 0.12 SD in math, and 0.174 SD in cognitive ability, relative to the control group
Barrera-Osorio et al. 2008, Colombia	Conditional cash transfer	Randomized controlled trial; CCT students were more likely to attend school (2.8 percentage points), more likely to remain enrolled (2.6 percentage points), more likely to matriculate to the next grade (1.6 percentage points), more likely to graduate (4.0 percentage points), and more likely to matriculate to a tertiary institution (23 percentage points)
Barrera-Osorio and Filmer 2013, Cambodia	Poverty or merit scholarships to students	Randomized controlled trial; poverty scholarship recipients were 18 percentage points and merit recipients 13 percentage points more likely to reach sixth grade than the control group; only merit incentives increased achievement (0.17 SD on the math test and 0.149 SD on the Digitspan test)
Behrman, Sengupta, and Todd 2000, Mexico	Conditional cash transfer	Randomized controlled trial, difference-in-difference; significant enrollment rate increases for the treatment group beyond those for the control of 6.9% for 12-year-olds, 7.7% for 13-year-olds, and 8.9% for 14-year-olds. Also positive increases, although not significant at the 5% level, for all other ages in the 10-17 range. No impact on achievement (Spanish and mathematics) after 1.5 years of exposure to CCT
Behrman, Parker, and Todd 2009, Mexico	Conditional cash transfer	Matched difference-in-difference; 1% reduction in the age of entry to primary grades; 8% to 9% increase in grades of schooling completed
Behrman, Parker, and Todd 2011, Mexico	Conditional cash transfer	Difference-in-difference; no impact on achievement (reading, mathematics, and written language skills) 5.5 years after exposure to CCT

Study and location	Structure	Evaluation method; results
Behrman et al. 2013, Mexico	Monetary incentives to students, to teachers, and to groups of students, teachers, and administrators	Randomized controlled trial; this CCT program is part of a large poverty reduction program. No impact of teacher-only incentive treatment on test scores; payments to students increased by 0.169 SD and group payments to teachers, students, and administrators increased scores by 0.314 SD
Berry 2013, India	Monetary and non-monetary incentives to students or their parents	Difference-in-difference; incentives had impact on test scores, with impact of 0.48 points using raw test scores and 0.53 SD using relative test scores. Students in treatment groups attended after-school classes at rate of 24% whereas control group attended at a rate of 11%
Bettinger, 2012, USA	Monetary incentives to students	Randomized controlled trial; a significant impact of 0.15 SD on the state mathematics achievement test; no impact on reading, social science, or science tests. Intrinsic interest was not significantly lower among the treatment group
Blimpo 2013, Benin	Monetary incentives to students, to groups of students, and a tournament where highest group score won prize money	Randomized controlled trial; incentives had positive impact on achievement, with an effect size of 0.29 SD for the individual group, 0.27 SD for the team incentive, 0.34 SD for the team tournament
Bourguignon, Ferreira, and Leite 2003, Brazil	Conditional cash transfer	Multilogit regression; used simulations to estimate counterfactual and estimated that about 40% of 10- to 15-year-olds not currently enrolled in school would enroll in response to the program. Among poor households, this percentage was estimated to be 60%
Chaudhury and Parajuli 2010, Pakistan	Conditional cash transfer	Regression discontinuity design; female enrollment increased by 10%
Dammert 2009, Nicaragua	Conditional cash transfer	Randomized controlled trial/quintile regression; increased school attendance by 12 percentage points for girls and 18 percentage points for boys. Households with lower expenditures received lower positive impacts from the program
de Janvry, Finan, and Sadoulet 2006, Brazil	Conditional cash transfer	Randomized controlled trial; increased complete-year attendance by 7.8 percentage points. Improved grade promotion by 6.2%
Duryea and Morrison 2004, Costa Rica	Conditional cash transfer	Propensity score matching; improved self-reported attendance by 5 percentage points. No impact on the probability of passing the grade
Edmunds and Tancock 2002, USA	Non-monetary prizes to individual students based on number of books read	Multivariate analysis of variance; incentives had no impact on number of books read. Incentives had no impact on children's value of reading, self-concept as readers, or total reading motivation, as reported by either parents or students
Filmer and Schady 2011, Cambodia	Larger and smaller scholarships to individual students	Regression discontinuity design; the smaller scholarship had a very large impact on school attendance, about 25 percentage points, but the larger scholarship did not significantly raise attendance above this level

Study and location	Structure	Evaluation method; results
Fryer and Holden 2013, USA	Monetary incentives to individual students, their parents, and their teachers	Randomized controlled trial; students in treatment schools mastered 1.087 SD (or 0.031) more math objectives than control students. Treatment parents attended almost twice as many parent-teacher conferences as control group parents. There was a 0.081 SD or 0.025 increase in math achievement on the statewide assessment; there was also a -0.077 SD or -0.027 impact on reading achievement, which was not incentivized
Galiani and McEwan 2013, Honduras	Conditional cash transfer	Regression discontinuity design; increase of 8 percentage points in school enrollment. Effects larger in the two poorest strata, and insignificant in the other three strata
Li, Han, Rozelle, and Zhang 2010, China	Monetary incentives to individual students for course grades	Randomized controlled study; no impact of cash incentives for course grades on standardized test scores
Kazianga, de Walque, and Alderman 2012, Burkina Faso	Food for education: Take-home rations and school feeding program	Randomized control trial; both THR and SFP increased enrollment by 3 to 5 percentage points. SFP increased attendance by 0.9 days for boys and (non-significant) 0.5 days for girls; THR increased attendance by 1 day for boys and 0.8 days for girls (both significant). THR improved the proportion correct on a math test by 8.4%; for SFP, the impact was 9.6%. THR and SFP did not improve performance on cognitive development measures
Kremer, Miguel, and Thornton 2009, Kenya	Scholarships provided to girls scoring within the top 15%	Randomized controlled trial; program raised test scores by an average of 0.19 SD. No decline in intrinsic motivation or in attitudes toward school
Levitt et al. 2013, USA	Monetary (smaller and larger amounts) and non-monetary (i.e., trophy) incentives to individual students	Randomized controlled trial; larger incentives led to test score improvements (0.103-0.132 SD) on a low-stakes diagnostic test. Non-financial incentives also impacted performance, and were only slightly smaller than the large monetary incentives, at least for younger children
Levy and Ohls 2007, Jamaica	Conditional cash transfer	Regression discontinuity design; attendance increased 0.5 days per month, or 3 percent over baseline. No impact on outcomes such as course grades or advancement to next grade
Maluccio and Flores 2005, Nicaragua	Conditional cash transfer	Randomized controlled trial; program produced a net increase in school enrollment of 13 percentage points and in attendance of 20 percentage points. Increased the number of children in grades 1-4 who advanced two grades by 7.3 percentage points, despite the fact that advancement past the fourth grade was not a formal requirement of the program. Nearly all estimated effects were larger for the extremely poor, often reflecting their lower starting points; larger impact was also observed for older children
Meng and Ryan 2010, Bangladesh	Food for education: Take-home rations	Propensity score matching and difference-in-difference; program increased high school participation rates by 15% to 26%. Participants stayed in school 0.7 to 1.5 years longer than their counterfactuals
Ozer et al. 2009, Mexico	Conditional cash transfer	Propensity score matching and regression; program decreased aggressive/oppositional symptoms by 10%. No impact on anxiety/depressive symptoms or total problem behaviors
Ponce and Bedi 2008, Ecuador	Conditional cash transfer	Regression discontinuity design; no impact on mathematics or language tests

Study and location	Structure	Evaluation method; results
Riccio et al. 2010, USA	Conditional cash transfer	Randomized controlled trial; program did not improve school outcomes overall for elementary or middle school students. Significant subgroup difference observed with better-prepared high school students, where program increased the rate of progression into 10th grade (by 5.4 percentage points), increased the fraction of students who earned at least 22 credits (i.e., on track for graduation) by 8.1 percentage points, and increased the percentage of students passing at least two Regents (state-administered end-of-high school) exams (by 5.9 percentage points)
Schady and Araujo 2006, Ecuador	Conditional cash transfer	Randomized controlled trial; probability that a child would be enrolled in school at the time of the follow-up survey was 3.2 to 4.0 percentage points higher among CCT recipients. Program resulted in an increase of between 9.2 and 11.4 percentage points in enrollment among complier households (i.e., those who believed school enrollment was a condition of the transfer)
Schultz 2004, Mexico	Conditional cash transfer	Randomized controlled trial; positive impact on enrollment for grades 1-8, averaging about 3 percentage points overall. Increased enrollment by 9.2 percentage points for girls and 6.2 percentage points for boys

SD = Standard deviation

4.1 Financial incentives can have significant impact on the attendance or enrollment rates of eligible children

Financial incentives generally have a positive impact on student attendance and enrollment rates, as shown by their applications in Brazil (de Janvry et al., 2006), Cambodia (Barrera and Filmer 2013; Filmer and Schady 2011), Colombia (Attanasio, Fitzsimons et al. 2005; Barrera-Osorio et al. 2008), Costa Rica (Duryea and Morrison 2004), Ecuador (Schady and Araujo 2006), Honduras (Galiani and McEwan 2013), Jamaica (Levy and Ohls 2007), Malawi (Baird et al. 2011), Mexico (Behrman et al. 2000; Schultz 2004; Attanasio, Meghir et al. 2005), Nicaragua (Maluccio and Flores 2005), and Pakistan (Chaudhury and Parajuli 2010).

Although CCT and FFE programs could increase school enrollment and student attendance by as much as 26 percentage points (Meng and Ryan 2010; Filmer and Schady 2011), in most cases the impact of these incentives were more modest, increasing attendance rates by 3 to 10 percentage points. Paradoxically, in countries where the education sector has been successful in achieving universal coverage, the impact of CCT and FFE can be very small, as there are very few additional children to enroll.

4.2 CCT and FFE have null to weakly positive impact on learning outcomes

Many CCT and FFE programs have no provisions for benchmarking learning outcomes, because the expectations seem to be that these types of incentives are aimed only at increasing student enrollment and retention. Hence, most findings show increases in enrollment, attendance, and grade promotion. Levy and Ohls (2007), Duryea and Morrison (2004), Riccio et al. (2010), and Alderman et al. (2010) failed to find an impact of financial incentives on grade promotion in Jamaica, Costa Rica, New York, and Uganda, respectively. A few studies have found positive

impacts on learning. In Nicaragua it was found that its CCT increased grade promotion by 7.3 percentage points (Maluccio and Flores 2005) and in Brazil by 6.2 percentage points (de Janvry et al. 2006). In Mexico, Behrman, Parker, and Todd (2009, 2011) found that the CCT program helped increase grade completion by half a grade for younger children, and close to one full grade for older children. In Bangladesh, the FFE program reduced dropout rates by 9 percent and students stayed in school up to 1.1 years longer (Meng and Ryan 2010).

CCT and FFE programs also had mixed impact on standardized test scores and course grades. Examining a CCT program for girls in Malawi, Baird et al. (2011) found weakly positive impact ranging from 0.12 to 0.17 SD on mathematics, reading comprehension, and cognitive ability tests. The Bangladeshi FFE program increased fifth-graders' achievement by 15.7 percent, but it also had a significant negative impact on the achievement scores of fourth-graders (Ahmed and del Ninno 2002). Evaluations of incentive programs in Jamaica (Levy and Ohls 2007), Ecuador (Ponce and Bedi 2008), Mexico (Behrman et al. 2000), and New York (Riccio et al. 2010) also failed to find an impact on course grades or on standardized test scores in mathematics and language.

4.3 If you pay them, they will learn: Pay for performance incentives for students can be effective

The evidence from some experiments on student rewards suggests that well-targeted incentives that pay students for reading and for completing math assignments can have positive significant effects on achievement (Allan and Fryer 2011). The impact of financial incentives may be sustained, even after the removal of the incentives.

Under a randomized incentive program in India, treatment children were eligible to receive 100 rupees if they met their literacy goals. The impact was estimated to be 0.53 SD (Berry 2013). In Kenya, an incentive program provided scholarships to sixth-grade girls within the top 15 percent of their school (Kremer et al. 2009). The scholarship program was estimated to have increased test scores across five core subjects by an average of 0.19 SD.

Angrist and Lavy (2009) examined the impact of a performance-based incentive program, in which low-achieving high-school students in Israel were eligible to receive cash payments for obtaining the high school matriculation certificate. These incentives increased certification rates by up to 8 percentage points in treatment schools. In Benin, Blimpo (2013) examined the impact of individual and group incentive structures for students in the 10th grade. Each of the incentive structures had a positive impact on achievement, with an effect size of 0.29 SD for the individual student incentives, 0.27 SD for the team incentive, and 0.34 SD for the team incentive augmented with a tournament component.

In Cambodia, paid scholarships were offered to students on the basis of poverty or merit. Other than the criteria for selecting the recipients, both incentives were similar in terms of monetary value and in conditions for renewing the scholarship. The merit incentive structure showed an impact of 0.17 SD on the mathematics test and 0.15 SD on a digit span test, but the poverty incentive structure had no impact on either measure (Barrera-Osorio and Filmer 2013). The

authors pointed out that there is an implicit trade-off in CCTs: cash transfers to poor families induce attendance by poor students at schools that are not prepared to improve their learning outcomes. Using CCTs to target high-performing students may increase their learning, but given the clear link between poverty and low student performance, spending limited funds on targeting good students would reduce equity, since students from poor households likely would be left out.

Presumably, education is a desirable investment but, for children, this investment has a high discount rate—that is, any perceived benefits of education may be unknown to them or be too far into the future. Intuitively, educational incentives directed to children must have a simple internal logic: the value of the incentive has to be higher than the children’s discount rate. Unfortunately, this discount rate is not well understood. Hence, implementing incentives to children becomes an empirical exercise where some incentives work and some do not.

4.4 Financial incentives for students do not decrease intrinsic motivation and their effects can be sustained over time

A concern of many policy makers is that paying students for academic performance could reduce their intrinsic motivation. However, multiple studies across contexts and age groups have failed to find a decrease in intrinsic motivation resulting from the financial incentives. Examining primary grade students in the United States, Bettinger (2012) failed to find any differences between the treatment and control students, in terms of either their own self-reports or teachers’ ratings of students’ intrinsic motivation. Similarly, neither Fryer (2011a) nor Fryer and Holden (2013) observed a decrease in intrinsic motivation over time, nor did they find any negative impact on a measure of self-reported “effort” index (e.g., complete homework, ask teachers for help). Studying sixth-grade girls in Kenya, Kremer et al. (2009) also found no decline in students’ intrinsic motivation or attitudes toward school.

There is a notable dearth of studies that have examined whether the impacts from financial incentive programs are sustained after the removal of incentives. Three notable exceptions are Fryer and Holden (2013), Baird et al. (2011), and Kremer et al. (2009). Fryer and Holden (2013) found that two years after the termination of an incentive program that paid students for the number of mathematics objectives mastered, the treatment effect for high-achieving students remained, such that the impact on mathematics achievement was 0.271 SD. Baird et al. (2011) found that one term after a CCT program in Malawi had ended, the impact of the CCT program persisted, such that the enrollment rate continued to exceed that of the control group. By contrast, the enrollment rates of students in an unconditional cash transfer program that was implemented and ended at the same time as the CCT program dropped to levels that were comparable to those of the control group. Similarly, Kremer et al. (2009) found that even one year after the incentives were removed, score gains remained relatively large, suggesting that the initial test score improvements reflected real learning.

4.5 Larger incentives do not necessarily lead to larger impacts

Although the topic is understudied within the literature, there is evidence of diminishing returns with respect to the size of the financial incentives, at least in international settings (see Allan and Fryer 2011 for a dissenting view using U.S. data). Filmer and Schady (2011) found that a US\$45 scholarship in Cambodia had a very large impact on school attendance (i.e., 25 percentage points), but a US\$60 scholarship did not significantly raise attendance above this level. Baird et al. (2011) varied the amounts of transfers to parents (between US\$4 and US\$10) or to girls (between US\$1 and US\$5), but found that increasing the transfer amounts had no effect on any of the outcomes for the CCT treatment. In addition, Levitt et al. (2013) found that younger children (i.e., second and third graders) in the U.S. did not respond differentially to low- and high-financial rewards (US\$10 and US\$20, respectively).

4.6 Financial incentives can have unintended consequences

Despite the generally positive impacts of financial incentives on school enrollment and attendance, Barrera-Osorio et al. (2008) found evidence of a reallocation effect, such that siblings of students who were eligible for cash awards were less likely to attend school and more likely to go to work. Namely, when they compared households that registered two children, they found evidence of lower school attendance and more labor market work for an untreated child with a treated sibling compared to an untreated child with a similarly untreated sibling. In this case, the interplay between education and poverty shows that the income needs of a family may be an incentive in itself, pushing for short-term income needs over the long-term financial benefits of education.

Similarly, Fryer and Holden (2013) described a substitution effect such that subject areas that were not incentivized showed decreases in performance, relative to the control group. In Houston, where the number of mathematics objectives mastered was incentivized, Fryer and Holden (2013) observed an increase in mathematics achievement on the state achievement test. However, they also observed a decrease of 0.078 SD on the state reading achievement test. The negative impact on reading achievement persisted for lower-achieving students, even two years after the removal of the incentive program.

5 School Report Cards and School Accountability

The use of report cards is not new. What is new is the participation of parents in their design, based on the parents' need for information. In some cases report cards have been used as accountability mechanisms for monitoring school grants controlled by parents. Table 3 shows a brief summary of the studies on school report cards and other accountability systems.

Table 3. Incentives to improve accountability

Study and location	Structure	Evaluation method; results
Andrabi, Das, and Khwaja 2009, Pakistan	Distribution of report cards with learning scores	Randomized controlled trial; report cards increased test scores by 0.10 SD and reduced private school fees by 23%. Top-ranked private schools did not increase test scores while government schools increased them by 0.10 SD
Barr et al. 2012, Uganda	Training parents on school-based management; participatory report cards	Randomized controlled trial; improved test scores, reduced teacher absenteeism by 14%, and increased student attendance by 9%. Test scores increased by 0.19 SD.
Blimpo and Evans 2011, The Gambia	Parent-controlled school grants	Randomized controlled trial; reduced teacher absenteeism by 23% and increased student attendance by 21%. Best results obtained in villages with higher levels of parent literacy
Bold et al. 2013, Kenya	Contract teachers	Randomized controlled trial; test scores went up 0.05 SD among schools supervised by a nongovernmental organization and did not go up in schools supervised by the government
Duflo, Dupas, and Kremer 2012, Kenya	Contract teachers	Randomized controlled trial; locally hired contract teachers had less absenteeism and improved learning outcomes by 0.18 SD
Piper and Korda 2011, Liberia	Early Grade Reading Assessment (EGRA) with training and materials, and report cards	Randomized controlled trial; <u>full</u> treatment included EGRA materials, teacher training and pedagogical support, and two types of report cards; <u>light</u> treatment included only materials and one report card; <u>control</u> group had no treatment. Full treatment score was 0.82 SD higher in reading comprehension, and 0.39 SD higher in listening comprehension than the control group.
Pradhan et al. 2011, Indonesia	Parent-controlled school grants with training for parents in school management council	Pilot on selected schools; block grants found to be helpful tool when combined with democratic elections of school management council (SMC) members. Training effective if tied to visits to good schools by SMC members. Training + elections of SMC members was associated with a gain of 0.22 SD in test scores.

SD = Standard deviation

5.1 School report cards are effective for informing parents and making teachers accountable

Andrabi, Das, and Khwaja (2009) assessed the market-wide impacts of providing report cards with learning scores to parents in 112 randomly chosen villages in Punjab province, Pakistan. One half of the villages were randomly included in the control group. The experiment covered 823 schools and 12,000 children in third grade, a total of 5,000 teachers, and background information on a sample of 1,800 households. The provision of report cards increased test scores by 0.10 standard deviations and reduced private school fees by 23%. However, there was wide variation across schools. In general, good private schools did not increase test scores while

government schools saw an increase of 0.10 standard deviations. A reduction in fees was observed mostly in the good private schools.

The report card included information on the academic performance of children and their relative performance with children in other schools. The report card was designed in collaboration with parents and schools. Parents wanted to know (1) their child's score and his/her rank relative to other students in the class, (2) the average score for each school in their village, and (3) the scores by category (word recognition, sentence building, etc.) so they would know where their child needed more help.

Because many parents were illiterate, the report cards were discussed in small groups and the content of each card explained to the respective parents. The group also discussed what could be done to help, rather than to emphasize the negative or blame the child. The facilitators gave no advice during the discussions, to allow parents to come up with their own ideas.

Another case study on the use and impact of report cards on early grade reading comes from Liberia, where a randomized controlled trial (under the EdData II project) was conducted in 180 schools (Piper and Korda 2011). A group of 60 schools received a full treatment that included the Early Grade Reading Assessment among randomly selected students in the second and third grades, along with teacher training in the assessment of reading performance, frequent pedagogical support, and books and pedagogical materials. In addition, full-treatment schools produced a reading report card for each student given to parents and a school report card given to the community. A second group of 60 schools received a light treatment, in which only the school report cards based on the Early Grade Reading Assessment were distributed among parents and the community. The control group of 60 schools did not get any of the above interventions.

The results of the experiment yielded an increase of 0.82 SD in reading comprehension and an increase of 0.39 on the rate of listening comprehension among full-treatment students. Students in the light treatment group showed modest improvements over the control group in two of the seven items being measured.

The results suggest that informing parents about school quality and effectiveness on a regular basis makes a positive difference in student performance, although such impact may be modest. It must be noted, however, that although report cards may have a positive impact on learning, their innate purpose is to enhance accountability, not student learning; they should be expected to have an indirect effect on learning outcomes.

5.2 Scorecards designed with parent participation improve school management

The results of a randomized experiment on community-monitoring interventions in primary schools in Uganda showed that *students in schools with strong oversight by school management committees performed better than students in the control group* (Barr et al. 2012). Strong oversight of schools by school management committees seemed to improve test scores by 0.19 SD, reduce teacher absenteeism by 14 percent, and increase student attendance by 9 percent.

All schools had scorecards. However, in the control group, school management committee members received training and support in the use of scorecards developed by the Ministry of Education in collaboration with nongovernmental organizations. In the treatment group, school management committee members designed their own scorecard, defined their own objectives, and established their own indicators of progress. The treatment was called Participatory Scorecards. The authors concluded that participatory scorecards helped engage teachers and parents more than the Ministry's scorecard.

5.3 Parent-controlled school grants tend to be effective

A controlled experiment on school grants aimed at improving school performance was conducted in The Gambia as part of its Whole School Development program (Blimpo and Evans 2011). In this experiment, one group of schools received a grant, with school staff and parents receiving management training; another group of schools received the grant and no training; and a third group was the control, without a grant or the training. A total of 273 schools participated in the program. Teacher absenteeism declined and student attendance increased by more than 20% in treatment schools, but test scores were the same among treatment and control schools. In villages with high initial literacy, the program was very successful; but in villages with low literacy, the program had negative effects. For grant-only schools, there was no impact on student or teacher attendance. Low parent literacy was a hurdle inasmuch it precluded parents from understanding school and student performance. Despite many efforts, school and student performance was not well explained by schools even after four years in the program. The authors cautioned against implementing school-based management programs in poor, undeveloped areas.

A pilot project in Indonesia tested four methods for empowering school management committees: (1) giving block grants to school management councils (SMCs), (2) training SMC members in school management, (3) establishing democratic elections of SMC members, and (4) creating formal linkages between the SMC and the village council (World Bank 2011). The results showed that the provision of block grants was the most helpful tool, followed by democratic elections of SMC members. Training was effective if tied to visits to good schools by SMC members from other schools. The use of elections of SMC members was positively linked to homework supervision by parents and to time on task by teachers. Of the four interventions, the block grants were most effective because they were used to increase spending on student activities, and to hire more temporary staff. Closer parent participation increased teacher attendance. These three actions had a significant effect on raising student scores by 0.22 SD.

5.4 Contract teachers can produce better results

An experiment with contract teachers and reduced class sizes in the first grade in 210 schools in Kenya found that locally hired contract teachers had less absenteeism and improved learning outcomes by 0.16 SD (Duflo, Dupas, and Kremer 2012), a difference roughly equal to an additional 3 months of schooling (Hill et al. 2007). The authors found that local civil service teachers undermined contract teachers because the latter had much lower salaries and produced

better outcomes, thereby becoming a threat to civil servants. During the experiment, civil service teachers worked less and pushed for the hiring of relatives as contract teachers.

Parent council members in the Kenya experiment received 90 minutes of training on monitoring teacher performance, and additional training on the contract-teacher program, including how to supervise the hiring of a teacher, how to do an interview, how to ask other parents for input, and how to monitor teacher attendance. Parent council members were also trained in how to hold meetings, create a calendar of activities, do a final evaluation of contract teachers, and complete a performance report. The results suggest that contract teacher programs work well if school-based management is in place and if parent council associations have some power over school decisions.

In another randomized controlled experiment on the impact of contract teachers on learning outcomes in Kenya (Bold et al. 2013), it was found that test scores went up 0.05 SD among schools supervised by a nongovernmental organization and did not go up in schools supervised by the government. Differences in outcomes seem to have been caused by the hiring of friends and family members of civil service teachers in schools under weak government supervision, a significant problem identified by Kremer, Moulin, and Namunyu (2003).

6 Incentives and Accountability in Education: Key Lessons Learned

6.1 Incentives for teachers can improve test scores

Incentives for teachers seem to have had positive effects on students' test scores. From the studies reviewed above, it seems that giving money to teachers for improved student performance did tend to increase student test scores. Likewise, giving money to schools on the basis of score gains also had positive results.

Can these results be scaled up to all areas in a country or to all countries? It is hard to tell, except on a case-by-case basis. For incentives to work well, they must be aligned with the capacity of teachers and schools to meet the expectations associated with the incentives. The relationship between incentives and performance must be clear to all.

There is abundant evidence on the positive effects of CCT and FFE on school attendance. On the other hand, their impacts on student achievement were not always analyzed in the contexts of these interventions, but the available figures suggest that the effects were not very large. In fairness, the goal of these programs typically was to bring the children to school rather than to affect their performance, and should be evaluated in that light. If the intent of a CCT or FFE program is to increase student learning, this element needs to be introduced at the stage of experimental design, as was done in the case in Andhra Pradesh.

The underlying issue that can be gleaned from the literature is that incentives and accountability must be intimately tied to the assessment and measurement of learning outcomes, to the

incentive structure offered to teachers, parents, or students, to the management of assessment information, and to the expectations about the consequences for bad performance.

For incentives to be effective, parents and students must be convinced that the reward is tied to outcomes that are within their direct control, rather than on outcomes out of their control. If students sense that their efforts will not work, or will be measured improperly, or will otherwise not be taken into account, the incentive structure will not work. Furthermore, if students sense that the reward will come no matter what they do, the incentive structure will not work either. In the case of parents, incentives seem to work when they are tied to actions that are within their reach, such as increased student attendance, and participation in some school decisions. However, if the incentives imply changes in parent behavior that will have a higher cost than the benefits, the incentives are likely to have little effect.

Incentives and accountability measures are most successful in improving learning outcomes when they are part of an overall reform program. Incentives have a positive role in education when trust and the “long route to accountability” do not work. There is therefore a need to create mechanisms that enable and empower parents and teachers to use other means to circumvent the failures of the system in sending the proper signals to teachers, parents, and students about improving learning.

Educational incentives seem to respond to classic cost-benefit analysis at the institutional and personal levels. Incentives work well at the institutional and/or individual levels when the cost of capturing them is lower than the value of the incentives themselves. That is why it is easier for students to respond to input-related incentives (conduct, uniforms, attendance), than to incentives related to outcomes, like improved test scores. As noted, some educational incentives related to learning outcomes assume that the school, teachers or students know what to do to improve test scores. This assumption may not hold in most poor countries. To make teachers and school accountable in terms of learning outcomes, the education system must also provide them with an infrastructure of assistance that operates all the way down to the classroom level.

Responses to an incentive by teachers, parents, or students seem to have two components: a positive financial component—the value of the reward—and a negative component brought in by uncertainty, which is the value of the probability of not performing well enough to attain the reward.

If the incentive rules are clear to all, and if the rewards are attached to attainable goals, the negative value of uncertainty decreases and the incentive is likely to work. This is important, as it underscores the need to have a support infrastructure aligned with incentives, so teachers and students can appraise the probability of success. If the support system is in place—such as teacher mentoring in the classroom, and access to adequate textbooks—the cost-benefit analysis made by teachers and students is bound to be favorable to incentives.

Finally, like other policies, incentives may have a few unintended consequences. For example, incentives that target only one subject, like math or reading, may be counterproductive as this could lead to an undesired substitution effect, where all other subjects are ignored by teachers, reducing overall student learning. Resources permitting, incentives should be provided to a

number of subject areas to prevent this unintended effect. In this regard, more studies examining students' outcomes after the incentives are no longer in place are needed. Ultimately, the decision to implement an incentive program to improve student achievement depends on how the impact from financial incentive programs compares to that of other types of educational interventions. McEwan's (2013) comprehensive review of educational interventions in developing countries suggests that the impact of financial incentives on achievement appears to be smaller than that for other interventions, such as curricular reform. Whether an incentive program is the best strategy for improving achievement needs to be understood within the larger context of other viable options.

6.2 Scaling up the experimental evidence on incentives and accountability in education requires attention to systemic issues

The experience to date shows that some incentives work and some do not, and success is highly specific to the school environment. As the evidence described above showed, monetary incentives seemed to work well when teachers and learning inputs were aligned, but the magnitude of the results varied greatly. Some interventions had positive effects on attendance and retention, and others positively affected learning outcomes. However, for incentives to be scaled up beyond the level of randomized controlled trials, the following general issues must be taken into account.

1. *Align all the stakeholders with power.* In education, the stakeholders with power may include politicians, community leaders, the Ministry of Education, the Ministry of Finance, funding agencies, and parent associations. Each of these stakeholders has different information needs, different expectations, and different views on the application of rewards and sanctions. Incentives and accountability must rely on common ground that takes into account the interest of these multiple stakeholders. This means that the establishment of *incentives and of procedures for accountability must be clear and agreed upon by these stakeholders.*
2. *Be aware that some stakeholders may extract benefits from the education system.* At the macro level, the Ministries of Education and Finance may use false baseline indicators to qualify for special financing from international agencies; at the mid-level, inspectors and schools may collude to affect enrollment statistics to get larger fiscal transfers; and at the school level, parents and the school may collude to report false enrollment or attendance data to get more funding for the school. These are examples of collusion already found in many countries. This means that entities promoting better educational performance through incentives and accountability must *ensure that baseline data and education management information system (EMIS) are of good quality.* In fact, ensuring the implementation of a good EMIS may be a good goal in itself.
3. *Fix as much as possible the deficiencies in the school support infrastructure.* Incentives require support to schools in order to work well. If the support infrastructure is absent or deficient, the best one can hope for is weak accountability—where information flows up

from the school, but the rewards and sanctions attached to incentives are not applied. Hence, incentives must rely on clear guidelines for the *provision of classroom and student support*.

4. *Fix the misalignment between policy and politics.* Financial incentives for teachers must have a clear link with teacher performance. If this link is unclear, the incentive will not work as intended. Financial incentives for teachers are likely to work better if administered at the school level. Again, parent participation and school autonomy are good mechanisms for aligning policy with politics because decisions will be made at the school level by parents and the school.
5. *Make sure that performance measures are sustainable.* Many performance-based interventions are dependent on “big data”—that is, multiple and recurrent student examinations. Scale-up and sustainability will require that other, less expensive metrics also be explored.

References

- Ahmed, Akhter U., and Carlo del Ninno. 2002. *The Food for Education Program in Bangladesh: An Evaluation of Its Impact on Educational Attainment and Food Security*. FCND Discussion Paper 138. Washington, DC: International Food Policy Research Institute.
- Alderman, Harold, Daniel O. Gilligan, and Kim Lehrer. 2010. *The Impact of Food for Education Programs on School Participation in Northern Uganda*. Washington, DC: International Food Policy Research Institute.
- Allan, Bradley M., and Richard G. Fryer, Jr. 2011. *The Power and Pitfalls of Educational Incentives*. The Hamilton Project Discussion Paper 2011-07. Washington DC: Brookings Institution.
- Alvarado, Macarena, Fabian Duarte, and Christopher Neilson. 2013. *Efectos Preliminares de la Beca Vocación de Profesor* [Preliminary effects of teacher professional scholarships]. Unpublished manuscript. Santiago, Chile: Centro de Estudios Ministerio de Educación.
- Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja. 2009. *Report Cards: The Impact of Providing School and Child Test-Scores on Educational Markets*. Working draft. Washington, DC: Development Research Group, World Bank.
- Angrist, Joshua D., and Victor Lavy. 2009. "The Effects of High Stakes School Achievement Awards: Evidence from a Randomized Trial." *American Economic Review* 99(4), 301–331.
- Angrist, Joshua D., and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton, New Jersey: Princeton University Press.
- Arcia, Gustavo, Harry Patrinos, Emilio Porta, and Kevin Macdonald. 2011. *School Autonomy and Accountability in Context: Application of Benchmarking Indicators in Selected European Countries*. Washington, DC: Human Development Network, World Bank.
- Attanasio, Orazio, Emla Fitzsimmons, and Ana Gomez. 2005. *The Impact of a Conditional Education Subsidy on School Enrolment in Colombia*. London, UK: The Institute for Fiscal Studies, Centre for the Evaluation of Development Policies.
- Attanasio, Orazio, Costas Meghir, and Ana Santiago. 2005. *Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA*. London, UK: The Institute for Fiscal Studies, Centre for the Evaluation of Development Policies.
- Baird, Sarah, Craig McIntosh, and Berk Ozler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *The Quarterly Journal of Economics* 126: 1709–1753.
- Barr, Abigail, Lawrence Bategeka, Madina Guloba, Ibrahim Kasirye, Frederick Mugisha, Pieter Serneels, and Andrew Zeitlin. 2012. *Management and Motivation in Ugandan Primary Schools: An Impact Evaluation Report*. Working Paper 2012-14, Partnership for

- Economic Policy (PEP). http://ageconsearch.umn.edu/bitstream/164412/2/1357671419-WP_2012-14_PIERI-11857.public.pdf
- Barrera-Osorio, Felipe, and Deon Filmer. 2013. *Incentivizing Schooling for Learning: Evidence on the Impact of Alternative Targeting Approaches*. Washington, DC: World Bank.
- Barrera-Osorio, Felipe, Marianne Bertrand, Leigh Linden, and Francisco Perez-Calle. 2008. *Conditional Cash Transfers in Education: Design Features, Peer and Sibling Effects. Evidence from a Randomized Experiment in Colombia*. Washington, DC: World Bank.
- Behrman, Jere R., Susan W. Parker, and Petra E. Todd. 2009. "Schooling Impacts of Conditional Cash Transfers On Young Children: Evidence from Mexico." *Economic Development and Cultural Change* 57(3): 439–477.
- Behrman, Jere R., Susan W. Parker, and Petra E. Todd. 2011. "Do School Subsidy Programs Generate Lasting Benefits? A Five-Year Follow-Up of Oportunidades Participants." *Journal of Human Resources* 46(1): 93–122.
- Behrman, Jere R., Susan W. Parker, Petra E. Todd, and Kenneth I. Wolpin. 2013. *Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools*. Philadelphia: Penn Institute for Economic Research at the University of Pennsylvania.
- Behrman, Jere R., Piyali Sengupta, and Petra E. Todd. 2000. *The Impact of PROGRESA on Achievement Test Scores in the First Year*. Washington, DC: International Food Policy Research Institute.
- Berry, James W. 2013. *Child Control in Education Decisions: An Evaluation of Targeted Incentives to Learn in India*. Ithaca, New York: Cornell University.
- Bettinger, Eric P. 2012. "Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores." *The Review of Economics and Statistics* 94(3): 686–698.
- Blimpo, Moussa P. 2013. *Team Incentives for Education in Developing Countries: A Randomized Field Experiment in Benin*. Stanford, California: Stanford University.
- Blimpo, Moussa, and David K. Evans. 2011. *School-Based Management and Educational Outcomes: Lessons from a Randomized Field Experiment*. Working paper. Washington, DC: Office of the Chief Economist, Africa Region, World Bank.
- Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur. 2013. *Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education*. Washington, DC: Center for Global Development.
- Bourguignon, Francois, Francisco H. G. Ferreira, and Phillippe G. Leite. 2003. *Conditional Cash Transfers, Schooling, and Child Labor: Micro-Simulating Bolsa Escola*. Paris, France: World Bank. <https://openknowledge.worldbank.org/handle/10986/17179>

- Chaudhury, Nazmul, and Dilip Parajuli. 2010. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Programme on Public School Enrolments in Punjab, Pakistan." *Applied Economics* 42: 3565–3583.
- Contreras, Dante, and Tomás Rau. 2012. "Tournament Incentives for Teachers: Evidence from a Scaled-Up Intervention in Chile." *Economic Development and Cultural Change* 61 (1): 219–246.
- Crouch, Luis, and Donald Winkler. 2007. "Governance, Management, and Financing of Education for All: Basic Frameworks and Case Studies." Paper commissioned for the *EFA Global Monitoring Report 2009, Governance, Management and Financing of Education for All*. Research Triangle Park, NC: RTI International.
- Cueto, Santiago, Máximo Torero, Juan León, and José Deustua. 2008. *Asistencia docente y rendimiento escolar: el caso del programa META* [Teacher support and school accountability: The META program]. GRADE Working Paper No. 53. Lima, Peru: GRADE.
- Dammert, Ana C. 2009. "Heterogeneous Impacts of Conditional Cash Transfers: Evidence from Nicaragua." *Economic Development and Cultural Change* 58(1): 53–83.
- Dee, Thomas, and James Wyckoff. 2013. *Incentives, Selection, and Teacher Performance: Evidence from Impact*. NBER Working Paper Series, No. 19529. Cambridge, Massachusetts: National Bureau of Economic Research. <http://www.nber.org/papers/w19529>
- De Janvry, Alain, Frederico Finan, and Elisabeth Sadoulet. 2006. *Evaluating Brazil's Bolsa Escola program: Impact on Schooling and Municipal Roles*. Berkeley, California: University of California, Berkeley.
- Di Gropello, Emanuela. 2004. *Education Decentralization and Accountability Relationships in Latin America*. World Bank Policy Research Working Paper 3453. Washington, DC: World Bank.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2012. *School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools*. NBER Working Paper No. 17939. Cambridge, MA: National Bureau of Economic Research.
- Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102(4): 1241–1278.
- Duryea, Suzanne, and Andrew Morrison. 2004. *The Effect of Conditional Transfers on School Performance and Child Labor: Evidence from an Ex-Post Impact Evaluation in Costa Rica*. Washington, DC: Inter-American Development Bank.
- Edmunds, Kathryn M., and Susan M. Tancock. 2002. "Incentives: The Effects on the Reading Motivation of Fourth-Grade Students." *Reading Research and Instruction* 42(2): 17–37.

- Elmore, Richard. 2003. *Knowing the Right Thing to Do: School Improvement and Performance-Based Accountability*. Washington DC: National Governors Association Center for Best Practices.
- Filmer, Deon, and Norbert Schady. 2011. “Does More Cash in Conditional Transfer Programs Always Lead to Larger Impacts on School Attendance?” *Journal of Development Economics* 96: 150–157.
- Fryer, Roland G., Jr., and Richard T. Holden. 2013. *Multitasking, Dynamic Complementarities, and Incentives: A Cautionary Tale*. Cambridge, Massachusetts: Harvard University.
- Fryer, Roland G., Jr. 2011a. “Financial Incentives and Student Achievement: Evidence from Randomized Trials.” *The Quarterly Journal of Economics* 126: 1755–1798.
- Fryer, Roland G., Jr. 2011b. *Teacher Incentives and Student Achievement: Evidence from New York City Public Schools*. Cambridge, Massachusetts: Harvard University and the National Bureau of Economic Research.
- Fryer, Roland G., Jr., Steven D. Levitt, John List, and Sally Sadoff. 2012. *Enhancing the Efficacy of Teacher Incentives Through Loss Aversion: A Field Experiment*. NBER Working Paper 18237. Cambridge, Massachusetts: National Bureau of Economic Research.
- Galiani, Sebastian, and Patrick J. McEwan. 2013. “The Heterogeneous Impact of Conditional Cash Transfers.” *Journal of Public Economics* 103: 85–96.
- Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. 2011. *Impact Evaluation in Practice*. Washington DC: World Bank.
- Glazerman, Steven, Ali Protik, Bing-ru Teh, Julie Bruch, and Jeffrey Max. 2013. *Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Randomized Experiment*. Washington, DC: Mathematica Policy Research.
- Glazerman, Steven, and Allison Seifullah. 2012. *An Evaluation of the Chicago Teacher Advancement Program (Chicago TAP) After Four Years*. Mathematica Policy Research Final Report. Washington, DC: Mathematic Policy Research. http://www.mathematica-mpr.com/~media/publications/PDFs/education/tap_year4_impacts.pdf
- Glewwe, Paul, Nauman Ilias, and Michael Kremer. 2010. “Teacher Incentives.” *American Economic Journal: Applied Economics* 2 (3): 205–227. <http://www.aeaweb.org/articles.php?doi=10.1257/app.2.3.205>
- Hanushek, Eric A., Suzanne Link, and Ludger Woessmann. 2013. “Does School Autonomy Makes Sense Everywhere? Panel Estimates from PISA.” *Journal of Development Economics* 104: 212–232.
- Hermans, Peter, and Jan Wieggers. 2011. *The Use and Misuse of Assessment Data in Test-Based Accountability Systems*. Presented at the 37th International Association for Educational Assessment Annual Conference, Manila, Philippines.

- Hill, Carolyn J, Howard S. Bloom, Alison Rebeck Black, and Mark W. Lipsey. 2007. *Empirical Benchmarks for Interpreting Effect Sizes in Research*. MDRC Working Papers on Research Methodology, New York: MDRC. <http://www.mdrc.org/publication/empirical-benchmarks-interpreting-effect-sizes-research>
- Hooge, Edith, Tracey Burns, and Harald Wilkoszewski. 2012. *Looking Beyond the Numbers: Stakeholders and Multiple School Accountability*. Organisation for Economic Co-operation and Development (OECD) Education Working Paper No. 85. Paris: OECD Publishing. <http://dx.doi.org/10.1787/5k91dl7ct6q6-en>
- Hout, Michael, and Stuart W. Elliott, eds. 2011. *Incentives and Test-Based Accountability in Education*. Washington, DC: The National Academies Press.
- Kazianga, Harounan, Damien de Walque, and Harold Alderman. 2012. “Educational and Child Labor Impacts of Two Food for Education Schemes: Evidence from a Randomized Trial in Rural Burkina Faso.” *Journal of African Economies* 21(5): 723–760.
- Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. “Incentives to Learn.” *The Review of Economics and Statistics* 91(3): 437–456.
- Kremer, Michael, Sylvie Moulin, and Robert Namunyu. 2003. *Decentralization: A Cautionary Tale*. Poverty Action Lab Paper No. 10. Cambridge, Massachusetts: Massachusetts Institute of Technology (MIT).
- Lavy, Victor. 2002. “Evaluating the Effect of Teachers’ Group Performance Incentives on Pupil Achievement.” *The Journal of Political Economy* 110(6): 1286–1317.
- Lavy, Victor. 2009. “Performance Pay and Teacher’s Effort, Productivity, and Grading Ethics.” *The American Economic Review* 99(5): 1979–2021.
- Lazear, Edward P. 2003. “Teacher Incentives.” *Swedish Economic Policy Review* 10: 179–214.
- Li, Tao, Li Han, Scott Rozelle, and Linxiu Zhang. 2010. *Cash Incentives, Peer Tutoring, and Parental Involvement: A Study of Three Educational Inputs in a Randomized Field Experiment in China*. Working Paper 221. Stanford, California: Rural Education Action Project.
- Levitt, Steven D., John A. List, Susanne Neckermann, and Sally Sadoff. 2013. *The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance*. Chicago, Illinois: University of Chicago.
- Levy, Dan, and Jim Ohls. 2007. *Evaluation of Jamaica's PATH Program: Final Report*. Washington, DC: Mathematica Policy Research.
- Maluccio, John A., and Rafael Flores. 2005. *Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social [social safety net]*. Washington, DC: International Food Policy Research Institute.

- Martins, Pedro S. 2009. *Individual Teacher Incentives, Student Achievement and Grade Inflation*. IZA Discussion Paper No. 4051. Bonn, Germany: Institute for the Study of Labor (IZA). <http://ftp.iza.org/dp4051.pdf>
- McEwan, Patrick J. 2013. *Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments*. Wellesley, Massachusetts: Wellesley College.
- Meng, Xin, and Jim Ryan. 2010. "Does a Food for Education Program Affect School Outcomes? The Bangladesh Case." *Journal of Population Economics* 23: 415–447.
- Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119(1): 39–77.
- Muralidharan, Karthik. 2012. *Long-Term Effects of Teacher Performance Pay: Experimental Evidence from India*. Working paper. San Diego, California: Department of Economics, University of California–San Diego. <http://econweb.ucsd.edu/~kamurali/research.html>
- Ozer, Emily J., Lia C. Fernald, James G. Manley, and Paul J. Gertler. 2009. "Effects of a Conditional Cash Transfer Program on Children's Behavior Problems." *Pediatrics* 123(4): 630–637.
- Piper, Benjamin, and Medina Korda. 2011. *EGRA Plus: Liberia—Program Evaluation Report*. Prepared for USAID under the Education Data for Decision Making (EdData II) project, Task Order EHC-E-06-04-00004-00. Research Triangle Park NC: RTI International. <https://www.eddataglobal.org/countries/index.cfm?fuseaction=pubDetail&ID=541>
- Ponce, Juan, and Arjun S. Bedi. 2008. *The Impact of a Cash Transfer Program on Cognitive Achievement: The Bono de Desarrollo Humano of Ecuador*. Bonn, Germany: Institute for the Study of Labor (IZA).
- Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Armida Alishjabana, Arya Gaduh, and Rima Prama Artha. 2011. *Improving Educational Quality Through Enhancing Community Participation Results from a Randomized Field Experiment in Indonesia*. Policy Research Working Paper No. 5795. Washington, DC: World Bank.
- Riccio James A., Nadine Dechausay, David Greenberg, Cynthia Miller, Zawadi Rucks, and Nandita Verma. 2010. *Toward Reduced Poverty Across Generations: Early Findings from New York City's Conditional Cash Transfer Program*. New York City: MDRC. http://www.mdrc.org/sites/default/files/full_588.pdf
- Schady, N., and Araujo, M.A. (2006). *Cash Transfers, Conditions, School Enrolment, and Child Work: Evidence from a Randomized Experiment in Ecuador*. Washington, DC: World Bank.
- Schultz, T. Paul. 2004. "School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program." *Journal of Development Economics* 74(1): 199–250.

- Springer, Matthew G., Dale Ballou, Laura Hamilton, Vi-Nhuan Le, J. R. Lockwood, Daniel F. McCaffrey, Matthew Pepper, and Brian M. Stecher. 2010. *Teacher Pay for Performance*. Nashville, Tennessee: National Center on Performance Incentives.
- Springer, Matthew G., John F. Pane, Vi-Nhuan Le, Daniel F. McCaffrey, Susan Freeman Burns, Laura S. Hamilton, and Brian Stecher. 2012. "Team Pay for Performance: Experimental Evidence From the Round Rock Pilot Project on Team Incentives." *Educational Evaluation and Policy Analysis* 34(4): 367–390.
- Steele, Jennifer L., Richard J. Murnane, and John B. Willett. 2010. "Do Financial Incentives Help Low-Performing Schools Attract and Keep Academically Talented Teachers? Evidence from California." *Journal of Policy Analysis and Management* 29(3): 451–478.
- Vegas, Emiliana, and Ilana Umansky. 2005. *Improving Teaching and Learning Through Effective Incentives. What Can We Learn from Education Reforms in Latin America?* Washington, DC: World Bank.
- World Bank. 2003. *World Development Report 2004. Making Services Work for Poor People*. Washington, DC: World Bank.
- World Bank. 2013. *What Matters Most for Teacher Policies: A Framework Paper*. System Approach for Better Education Results (SABER) Working Paper Series No.4. Washington, DC.
- Yuan, Kun, Vi-Nhuan Le, Daniel F. McCaffrey, Julie A. Marsh, Laura S. Hamilton, Brian M. Stecher, and Matthew G. Springer. 2012. "Incentive Pay Programs Do Not Affect Teacher Motivation or Reported Practices: Results from Three Randomized Studies." *Educational Evaluation and Policy Analysis* 35(1): 3–22.