

**IMPROVING EDUCATIONAL QUALITY THROUGH ENHANCING  
COMMUNITY PARTICIPATION: RESULTS FROM A  
RANDOMIZED FIELD EXPERIMENT IN INDONESIA**

By Menno Pradhan, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya  
Gaduh, Armida Alisjahbana, Rima Prama Artha<sup>+</sup>

July 2013

**Abstract**

*Education ministries worldwide have promoted community engagement through school committees. This paper presents results from a large field experiment testing alternative approaches to strengthen school committees in public schools in Indonesia. Two novel treatments focus on institutional reforms. First, some schools were randomly assigned to implement elections of school committee members. Another treatment facilitated joint-planning meetings between the school committee and the village council (linkage). Two more common treatments, grants and training, provided resources to existing school committees. We find that institutional reforms, in particular linkage and elections combined with linkage, are most cost-effective at improving learning. (JEL I25, O1)*

<sup>+</sup>Pradhan: VU University Amsterdam & University of Amsterdam, Dept. of Development Economics, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, m.p.pradhan@vu.nl; Suryadarma: Australian National University, Arndt-Corden Dept. of Economics, Coombs Building, Fellows Road, Canberra, ACT 0200, Australia, daniel.suryadarma@anu.edu.au; Beatty: Mathematica Policy Research, 955 Massachusetts Ave., Suite 801, Cambridge, MA 02139, USA, amandabeatty@post.harvard.edu; Wong: The Wharton School, University of Pennsylvania, 3620 Locust Walk, 1464 SHDH, Philadelphia PA 19104-6302, USA, maisy@wharton.upenn.edu ; Gaduh: Sam M. Walton College of Business, University of Arkansas, Dept. of Economics, Business Building 402, Fayetteville, AR 72701, USA, agaduh@walton.uark.edu; Alisjahbana: Faculty of Economics and Business, Universitas Padjadjaran, Jl. Dipati Ukur 35, Bandung, Indonesia, armida.alisjahbana@fe.unpad.ac.id; Artha: National Graduate Institute for Policy Studies (GRIPS), Jl. Mangga no. 20 rt 009/03 Kalisari Jakarta Indonesia 13790, phd009016@grips.ac.jp.

A large number of people contributed to the design, implementation and supervision of this research project, especially Vivi Alatas, Desmond Ang, Chris Bjork, Esther Duflo, Scott Guggenheim, Djoko Hartono, Dedy Junaedi, Ben Olken, Siswanto, Rosfita Roesli, Chris Thomas, Jeremy Tobacman, Tri Yuwono and from the Indonesian Ministry of National Education: Bambang Indriyanto, Sri Renani Pantjastuti, Sri Amien, Dasim Budimansyah, Agus Haryanto, Yadi Haryadi, Neneng Kodri, Suparlan, Anen Tumanggung, Yudistira Widadiana, Diana Sufa, Ismulyanto Apriwibowo. We are grateful to Deborah Cobb-Clark, Claudio Ferraz, Tue Gørgens, Andrew Leigh, Harry Patrinos Matt Stevens, and Matthew Grant Wai-Poi for helpful comments and suggestions. We also thank anonymous referees for their comments. Ying Chen provided excellent research support. We acknowledge financial support from the Japan Social Development Fund, and the Dutch Government. Maisy Wong is grateful for support from the Zell/Lurie Real Estate Center. The firms, PPA and Moores Rowland, implemented the intervention and survey respectively. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

There is growing interest to improve education quality in developing countries through community participation (Stiglitz 2002; Mansuri and Rao 2012). Many countries have created local institutions to coordinate this, such as school committees and parent teacher associations. Yet, often these government-sponsored institutions do not live up to their expectations (Banerjee and Duflo 2008; Bruns, Filmer, and Patrinos 2011). Why do these institutions fail? Is investing in them a viable strategy to improve learning outcomes? To date, results from several randomized interventions are mixed.<sup>1</sup>

This paper investigates the role of school committees in improving education quality in Indonesia. Our study is a randomized evaluation comprising 520 schools in Central Java, from 2007 to 2008. We have four main treatments. The first two are novel institutional reforms that improve the social capital of the school committee by strengthening its trustworthiness and relationship with the community (Ostrom and Ahn 2009). The first treatment facilitated democratic elections of school committee members. The second treatment linked school committees to the village council by facilitating joint planning meetings (we call this linkage).

We benchmarked these two novel treatments against more common treatments that improve the financial and human capital of school committees: providing block grants and training. Our institutional reforms are attractive because they are cheap to implement compared to grants and training.

<sup>1</sup> Gertler, Patrinos, and Rodríguez-Oreggia (2010) find that providing grants to parent associations in Mexico improves school quality, but Blimpo and Evans (2010) find that providing grants to schools, or grants to schools and training to community members in Gambia has no effect on test scores. Khattri, Ling, and Jha (2010) find small effects from providing training to principals and teachers, and grants to schools. Duflo, Dupas, and Kremer (2009) find that training school committee members in Kenya only impacts learning if interacted with another treatment that provides contract teachers. Merely informing communities about the state of service delivery seems insufficient to improve services (Kremer and Holla 2009; Banerjee et al. 2010).

We contribute to the literature on community participation and education in several ways.<sup>2</sup> First, analyzing the impact of reforming institutions is hard because randomly assigning a political process is seldom possible (Acemoglu, Johnson, and Robinson 2001). A unique opportunity arose for this study through our partnership with the Ministry of National Education (MONE). This allowed us to randomly assign institutional reforms to these government-sponsored institutions that would not have been possible in Indonesia without the support of the government. For example, linkage involved signing a joint action plan with another government-sponsored entity, the village council.

Another novel institutional reform treatment is the elections of school committee members. To our knowledge, this is the first paper that credibly estimates the causal impact of elections using a randomized evaluation.<sup>3</sup> Schools randomly assigned the election treatment showed an increase in the community's awareness of the school committee and resulted in many newly elected members of the school committee.

We also contribute to the literature by evaluating a rich set of interventions in a common context. We use a large collection of intermediate outcomes that allow us to unpack the dynamics of community participation. We evaluate seven interventions, including the four treatments and three combinations of treatments (linkage and election, linkage and training, elections and training). To do so, schools were randomly assigned into nine comparison groups: one control group, one group that was assigned the grant only, and seven groups that were assigned the grant plus combinations of

<sup>2</sup> See Dupas and Kremer (2009); Pandey, Goyal, and Sundararaman (2009); Banerjee et al. (2010); Blimpo and Evans (2010); Duflo, Khattri, Ling, and Jha (2010); Bruns, Filmer, and Patrinos (2011);. We are also related to the broader literature on community participation and public service delivery, see Olken (2007); Banerjee and Duflo (2008); Björkman and Svensson (2009) and Olken, Onishi, and Wong (2012). See Kremer and Holla (2009) and for a survey of related literature.

<sup>3</sup> Olken (2010) is the randomized evaluation of the impact of plebiscites (community votes on categories of public goods to spend on). Martinez-Bravo et al. (2011) study the impact of elections of village leaders in China by using variation in the timing of the introduction of these elections.

training, linkage and/or elections. These latter groups were always assigned the grant treatment also to ensure that the school committee would have money available to support initiatives that might arise from the treatments.

Institutional reforms of school committees that enhance its social capital in the community show positive effects on learning while interventions that enhance its financial and human capital are less cost-effective. Two years after the start of the project, Indonesian test scores increase by 0.17 standard deviations for linkage and 0.23 standard deviations for linkage plus elections. Training, on the other hand, shows no effect on learning and the effect of grants alone is positive but usually statistically indistinguishable from zero. We further analyze five possible pathways through which learning could be enhanced: stakeholder awareness of school committees, school-based management and contributions from parents, teachers, and communities.

Linkage and linkage plus elections mostly improve learning by increasing community contributions by 0.14 and 0.13 standardized units. These effects are driven by reported increases in village councils' collaboration with schools and principals' satisfaction of the village councils' attention to education. We do not observe increases in financial and in-kind donations. Instead of being a passive fundraising vehicle only, the joint planning meetings between the school committee and the village council translated into co-sponsored education initiatives such as the hiring of contract teachers and the establishment of village study hours. Strikingly, linkage and linkage plus elections do not raise stakeholder awareness of school committees but impact learning, whilst interventions that increase awareness show limited effects.<sup>4</sup>

Our results shed light on why some community participation programs improve service delivery but some do not. Considering that community mobilization is time-intensive (Banerjee and Duflo 2008), it is important to

<sup>4</sup> This is consistent with findings in the literature suggesting that information campaigns that raise school committee awareness alone may not improve learning (Banerjee et al. 2010).

know this. For example, Björkman and Svensson (2009) find that community participation improved health care delivery by Health Unit Management Committees (HUMC) in Uganda but Banerjee et al. (2010) find that enhancing community participation through Village Education Committees (VEC) in India did not improve school quality. In a survey article of community participation, Banerjee and Duflo (2008) conjecture that it is because HUMCs are more powerful than VECs politically. Our findings corroborate this.

Linkage possibly has strong learning effects because the school committee in Indonesia, like VECs in India, has no power. Engaging the more powerful village council leads to concrete actions on the ground and increases the legitimacy of the co-sponsored initiatives. In contrast, elections increase awareness and the representativeness of school committee members but enhancing this dimension of community participation alone does not endow the school committee with sufficient power to impact learning.

Our results also speak to cross-country evidence that changes in educational spending show weak correlations with changes in learning outcomes (World Bank 2003). Like Duflo, Dupas, and Kremer (2009), we find that providing resources alone is not cost-effective. However, substantial gains are obtained when grants are combined with relatively cheap institutional reforms.

In short, encouraging community participation and action that ultimately impacts service delivery is a time-intensive process (Banerjee and Duflo 2008). Our results suggest that fostering ties between the school committee and powerful local groups through joint planning activities could expedite this process and is cost-effective at improving learning.

We describe the interventions in the next section, outline the timing, sample and data in Section II, discuss empirics in Section III, report results in Section IV and conclude in Section V.

## **I. Motivation, intervention design and implementation**

After achieving universal primary school enrollment in the 1980s,<sup>5</sup> Indonesia began to shift attention to quality, with reforms such as teacher training, curricula revision, facility improvements, and later on school-based management (Kristiansen and Pratikno 2006). Despite these initiatives, Indonesia awaits marked progress in learning.<sup>6</sup>

Against this backdrop of national efforts at promoting education quality, the Government of Indonesia in 2002 instituted a decree that gave school committees a greater role in advising and supporting school management, and encouraged greater engagement with the community.<sup>7</sup> The decree stipulated that school committees would replace committees known as BP3 (*Badan Pembantu Penyelenggaraan Pendidikan*) in Indonesian. The primary function of the BP3 was to raise funds to support the school, yet the funds were largely handed over to principals. The school committee would go further, by making recommendations on school expenditures, teacher qualifications, and school facilities. In addition, the school committee was expected to act as a mediator between the school and the community, and promote community, especially parental, involvement in the school.

Four years after this decree, school committees were largely still operating under the BP3 model (Fearnley-Sander et al. 2008). This called for cost-effective and scalable ways to help school committees realize the role envisioned in the decree. Field visits and further discussions with MONE led to the development of four approaches tested in this experiment. We discuss these treatments below.

<sup>5</sup> Gross enrollment stood at 114 in the mid-nineties (Behrman, Deolalikar, and Soon 2002).

<sup>6</sup> In reading, it ranks 57<sup>th</sup> out of 65 countries that participated in the Program for International Student Assessment (PISA) in 2009 (Organisation for Economic Cooperation and Development 2010). Meanwhile, the 2007 Trends in International Mathematics and Science Study (TIMSS) shows that only half of Indonesia's students performed above the lowest international benchmark in math (Mullis, Martin, and Foy 2008). Hanushek and Woessmann (2008) find that slightly less than 30 percent of a cohort of grade nine Indonesian students had achieved full literacy.

<sup>7</sup> See Appendix II, Minister of National Education Decree 044/U/2002, 2 April 2002.

This experiment was funded by a grant from the Japanese government to the Ministry. The Ministry contracted out intervention implementation to a consulting firm, Pusat Pengembangan Agribisnis (PPA),<sup>8</sup> for a total contract value of 2.9 billion Rupiah (US\$315,000).

#### A. *Intervention 1: Grant and facilitation*

All 420 treatment school committees were provided a block grant of eight million Rupiah (US\$870). This was the first time that a significant sum of money was directly under the school committee's control.<sup>9</sup> The grant was small relative to school budgets, but comparable to similar treatments.<sup>10</sup>

We hypothesized that the grant could help the school committee catalyze change. The committee could more easily reach out to stakeholders because they had more money to hold meetings. Planning how to spend the grant was also an occasion for meeting. For linkage schools, this topic was included in the discussion with the village council and school management. For schools that were assigned to training, the planning process was used as an exercise to put the material covered in the training into practice. The grant also provided the school committee with an opportunity to contribute to school activities, and thus be a more active participant in the school.

The school committee did not receive the money without strings. With the assistance of facilitators, it was expected to develop an expenditure plan and to post how funds were used on the school notice board.<sup>11</sup> Facilitators

<sup>8</sup> The World Bank supported PPA by making available a consultant for about two months, with the task of assisting in planning the interventions.

<sup>9</sup> Reported yearly parent contributions collected through the school committee at baseline were on average 1 million Rupiah (US\$109). School committees transferred this money to the school and it entered the school budget. The project grant was the only budget under their control.

<sup>10</sup> The average 2005/2006 school budget recorded in the baseline survey was 200 million Rupiah (US\$22,000), of which 146 million (US\$15,900) was for personnel related expenditures. This grant amount is comparable to grant treatments in the literature. Blimpo and Evans (2010) use a grant of US\$500 or five percent of the average school budget. Gertler, Patrinos, and Rodríguez-Oreggia (2010) provide grants of US\$500 to US\$700 per year to schools in Mexico.

<sup>11</sup> For the linkage schools, this plan was developed with the village council. See below.

coached school committees on what grant funds could be used for and what problems they might consider addressing with the grant, approved expenditure proposals from school committees, authorized the transfer of the block grant (once the proposals were approved), ensured the transfer of the grant to school committees' bank accounts, and monitored the use of the block grant. On average, one facilitator was assigned to ten schools, and visited each school committee 13 times.<sup>12</sup> Using these estimates on the number of visits, the time facilitators spent in the schools, and staff salaries to break down the facilitation costs, and taking into account other treatment-specific costs, we estimate that the cost of implementing the grant treatment was about US\$321 (excluding the grant itself) per school.

The block grant was transferred directly from the Ministry into a bank account held by the school committee, in two tranches. This first tranche (three million Rupiah, US\$326) was disbursed in January 2008, three months later than planned, due to budgeting problems at the Ministry. In the midline survey, in April 2008, school committee respondents reported an average spending of 2.3 million Rupiah (US\$250). The biggest expenditure category was for meetings, averaging almost 1 million Rupiah (US\$109).

The second tranche also confronted similar budget delays. It was to be disbursed to the schools subject to sufficient progress achieved by the school committee in using the first tranche of the grant; but, in practice, all schools received the second tranche, and received it ten months late, in December 2008, soon after the endline survey.<sup>13</sup> Thus, these results measure the impact of the first tranche, and the anticipation of getting the remaining funding.

<sup>12</sup> Based on interviews by the authors with two district facilitators after completion of the project. According to PPA, the project employed 50 facilitators for a period of 15 months, and six district facilitators who managed the district teams. The interventions were implemented consecutively – elections, linkage, and then training.

<sup>13</sup> The endline survey had to be conducted before the second tranche was disbursed because the grant from the Japan Social Development Fund that financed the survey was about to expire.



### B. *Intervention 2: Training*

We hypothesized that school committees could be failing due to a lack of information, such as a lack of knowledge about the decree or about ways to put it into practice. Thus, a two-day, district-level training attended by four school committee members covered planning, budgeting and steps the school committee could take to support education quality.<sup>14</sup> The budget session focused on a plan for spending the block grant. Materials drew heavily on the Creating Learning Communities for Children (CLCC) model developed by UNICEF, which focuses on active learning, school-based management and community participation, and has served as the foundation for several donor projects promoting school-based management in Indonesia. It was not replicated fully because the cost per school would have been too high. The training also included a visit to a ‘model’ school committee that had been successful in applying school-based management practices. Appendix 1 provides more details on training and our departures from the CLCC model. Schools assigned to training received three additional visits beyond those provided with the grant. One visit was to announce the training and to agree on who would participate; another to deliver an official invitation stamped by the district education office; and a final visit, which took place just before the training, to ensure that those invited would come. Implementing the training cost US\$360 (including the cost of training) per school.

### C. *Intervention 3: Election*

The primary concern to be addressed by the election intervention was that school committee members were often handpicked by school management and did not represent parents or the community. We hypothesized that a democratic mandate and more diverse members would enhance the social

<sup>14</sup> The four members included the principal, teacher, parent, and one village representative. For the schools that were also assigned the linkage treatment, one additional representative from the village council was invited.

capital of the school committee (Ostrom and Ahn 2009), by helping it gain legitimacy, better communicate with parents and community members, and act in their interest with more authority.

The intervention introduced two primary changes to the process outlined in the decree. First, the intervention introduced quotas to tighten the membership guidelines in the decree,<sup>15</sup> designating that the committee be comprised of six parents, three community members, one teacher, the principal and the head of the village council. The rationale for the quota was to ensure better representation.

Second, the intervention redefined the role of the election committee and gave more structure to the election process.<sup>16</sup> The guideline stipulated in the 2002 decree recommends that schools form an election committee to organize the election of school committee members. However, the election committee was often selected by principals, thus making the school committee members also indirectly selected by principals. The intervention facilitated a participatory process to pick election committee members, and supported them to organize a more democratic election of the school committee.

Project facilitators organized two meetings – at the school and village levels – to establish the election committee.<sup>17</sup> After a meeting to establish the

<sup>15</sup> The 2002 decree stipulates that the school committee include at least nine members, including community representatives (with a maximum of three from the village government), teachers, parents and the principal (although he/she cannot be the head); and the community must propose these candidates.

<sup>16</sup> The election process was modeled after that used in the World Bank/Government of Indonesia Urban Poverty Project (UPP), known by the Indonesian acronym P2KP. The experience of UPP pilots, in which membership of village government was put to a popular vote, was that elites stood for election and were elected. Thus, the UPP project modified election processes to mobilize candidates from different sectors of the community, which is the model used in this experiment.

<sup>17</sup> The school-level meeting comprising teachers and the school principal chose two parents and one teacher as members of the election committee. The teacher elected here automatically became a school committee member. Meanwhile, the village-level meeting, comprising various community leaders, came up with two community members (that were not among those recommended at the school-level meeting) as members of the election committee. Invited community leaders include the heads of the village, village council, and the neighborhood associations (RW's), as well as the leadership of the women association (PKK) and other community leaders specific to the localities.

rules of the election, the election committee, with the assistance of the facilitator, organized school-level and village-level elections of school committee members. In the former, parents elected one parent representative for each grade level in a secret ballot.<sup>18</sup> For the latter, the election was preceded by a human-resource mapping exercise to determine representatives, who then voted in the village-level elections.<sup>19</sup> Subsequent meetings were held to sign a decree establishing the school committee and to develop a work plan. The election process generally took facilitators five visits to schools beyond those necessitated by grant implementation, and took place in three batches to spread out facilitator workloads.<sup>20</sup> The intervention cost US \$174 per school.

Despite the efforts of PPA to encourage communities to remain faithful to the design outlined above, some schools refused to conduct an election. As shown in Table 1, 48 percent of the schools randomly assigned to implement an election actually did as intended. Of those schools that did not fully comply, about seven percent of committees refused to change any members, while the remainder of committees agreed to a compromise of electing representatives of previously unrepresented groups. Some of those that refused or partially refused claimed that their school committee members had just begun their terms, and thus they did not want to start over with new membership after new members had just been appointed or elected.

[Insert Table 1 here]

Nevertheless, schools assigned to hold the elections did experience substantial changes in school committee membership as compared to other

<sup>18</sup> In the school-level election, parents wrote down the names of three parents whose children were in the same grade as theirs. For each grade level, one parent with the most votes became the school committee member.

<sup>19</sup> The election committee first met with the heads of the village and the neighborhood associations to choose five representatives from each neighborhood to be sent to the village-level election. In the village-level election, these representatives then voted among themselves who would become the community representatives in the school committee in a secret ballot. Each representative could write down three names in the ballot.

<sup>20</sup> Batches took place 15 April to 31 May 2007, 1 June to 14 July 2007, and 15 July to 31 August 2007.

schools. In the former group, 54 percent of school committee members recorded in the endline survey had started their term after the baseline, while in the latter group only 1 percent did so. We present ‘intent to treat’ results, which estimates the impact of being assigned the election treatment.

#### D. *Intervention 4: Linkage*

The aim of the linkage intervention was to increase the engagement of the village council (known in Indonesian as *Badan Perwakilan Desa* or BPD). The village council is a democratic village organization elected by villagers and has the power to draft village legislation, approve the village budget, monitor village government, and can even call on the district head to remove the village head (Antlov 2003). By facilitating a series of planning meetings between the school committee and the village council, discussing potential measures to address education issues in the village, it was believed that the school committee could more effectively form a bridge between the community and the school. We hypothesized that this “linkage” of school committee with the powerful village council would first increase the stature of the school committee vis-à-vis school management; and second, result in concrete support from the village council for measures addressing education problems that could not be solved by the school committee with school management alone. The intervention represents the spirit of the decree since the decree even envisions village representation in the school committee but findings from field visits indicated, and baseline data confirmed, that there was little evidence of this collaboration before the linkage intervention. At baseline, 22 percent of all school committee representatives reported collaboration with the village council.

The intervention cost US \$125 per school, mainly for two additional visits to the school beyond those provided with the grant. The first facilitated meeting was between school principal and school committee members to identify measures for improving education quality that they would then

propose to the village council. These measures were discussed in a subsequent meeting with village council representatives and other village officials, and the results of the meeting were documented in a memorandum of understanding signed by the head of the school committee, the head of the village council, and the school principal.

Examples of measures that parties collaborated on included building school facilities, establishing village study hours (two hours in the evening when households would turn off televisions and computer game kiosks would be closed), hiring contract teachers, making land available for school infrastructure expansion, resolving conflicts between two schools in a community, and encouraging social and religious activities at school. In some cases, collaboration even extended to village council representatives becoming school committee members (Bjork 2009).

## **II. Sample and timeline**

This study took place in six districts in Central Java and Yogyakarta, a region chosen because there were few large education projects active in the area, enabling the results to be relatively free from the risk of contamination from other projects. Moreover, conditions were hypothesized to be ripe for community engagement to flourish – the area is peaceful, has reasonably high levels of existing social capital, and schools are relatively well equipped (high levels of electricity, adequate number of teachers, etc). The evaluation also focuses on public primary rural schools – public because this evaluation was designed by the Ministry, which has the authority over public schools; rural because the majority of schools in the country are in rural or semi-rural areas, and it was hypothesized that accountability would be easier to engender in smaller, closer-knit areas.

From six districts in two provinces, the sampling frame was further restricted by excluding sub-districts containing fewer than eight villages<sup>21</sup> and schools with parallel classes in grade four.<sup>22</sup> We also dropped schools with extremely good or bad average sixth grade examination scores in mathematics or Indonesian. The school-level averages ranged from 0 to 9.65. We dropped schools where the average scores for mathematics or Indonesian were below four (this cutoff corresponds to the seventh percentile for average mathematics scores and third percentile for average Indonesian scores) or above eight (99<sup>th</sup> percentile for mathematics and 97<sup>th</sup> percentile for Indonesian).<sup>23</sup> We selected on test scores *ex ante* because we thought schools that were weak lacked the resources to fully benefit from our institutional reforms and schools that have high test scores will have less need for our interventions.

From this sampling frame, we first sampled 44 sub-districts. To avoid spillovers between treatment and comparison schools within a village, we sampled one school per village.<sup>24</sup> That is, we first selected 520 villages, then randomly selected one school from each of these villages.

The resulting sample of 520 schools was then stratified into three groups using their average test scores.<sup>25</sup> Within each stratum, schools were randomly

<sup>21</sup> This restriction was imposed because in the initial design, facilitated meetings with sub-district government education officials were envisioned, and too few villages per meeting would make this intervention ineffective. However, this idea was never implemented, making the restriction unnecessary.

<sup>22</sup> Parallel classes are grades with more than one section or teacher. This restriction was imposed because the evaluation was not planning on assigning student IDs or ensuring that the student population was identical over time. With only one class per grade, and low dropout and repetition rates, the evaluation team was confident that the same children interviewed in grade four would be in grade six two years later. However, this actually became an issue, since several schools merged, but the team was able to match student names, see below.

<sup>23</sup> To gauge the extent of the external validity problems due to this selection criterion, we checked that the average scores for the selected schools are not too different from the full sample. The average mathematics score for selected (all) schools was 6.0 (5.8) and the average Indonesian score for selected (all) schools was 6.9 (6.8). The standard deviations are 1.3 for mathematics and 1.2 for Indonesian average scores. The median is also not that different.

<sup>24</sup> The sampling probability was increased accordingly for schools that were located in villages with more than one school to keep the probability of being sampled equal across schools.

<sup>25</sup> Calculated as  $0.5 \times \text{average Mathematics score} + 0.5 \times \text{average Indonesian score}$ .

assigned into the nine treatment and comparison groups (see Table 2).<sup>26</sup> A disproportionate share of the sample was allocated to the cells assigned nothing (control group) and the cells assigned just the block grant, in order to separately identify the effect of the block grant. The linkage, election and training interventions were implemented in combination with the grant. The cell size for the training intervention was made slightly smaller than the non-training cells because training is a relatively costly intervention. We account for these different cell sizes using weights (see the next section).

[Insert Table 2 here]

The baseline survey took place in January 2007, midline in April 2008, and the endline survey in October 2008, as shown in Table A1.<sup>27</sup> Tests in mathematics and Indonesian, designed by the Ministry,<sup>28</sup> were administered to all students in grade four at baseline and grade six at endline.<sup>29</sup>

Attrition occurred both in terms of schools and students. Three out of 520 schools were not included in the endline survey. One school refused interviews. In the others, the implementer implemented the treatment in schools different than those surveyed in the baseline. These schools were assigned to be in the grant plus linkage, grant plus election, and control groups respectively. Our analysis is based on the 517 remaining schools.

We matched students on the basis of student names written on the test sheets and school ID. We were able to match 10,941 students, which is equal to 87 percent of the tests administered at baseline in grade four, and 88 percent of the tests administered at endline in grade six, in the 517 schools that participated in both rounds. To investigate the potential impact of the match

<sup>26</sup> Random allocation conducted by authors using Stata.

<sup>27</sup> The survey was conducted by Moores Rowland, a survey firm selected by the Ministry.

<sup>28</sup> We relied on *Puspendik*, which is the test development department of the Ministry, to design the tests to ensure ownership of the results by the Ministry.

<sup>29</sup> As mentioned above, schools with multiple parallel classes were excluded (so all sample schools started out with one grade 4 class), and thus we did not assign student ID. But by the time of the endline survey, several schools had more than one grade six class, due to schools merging. We only have data on the number of parallel grade six classes in 240 sample schools, and found that 13 grade six classes had two parallel classes (5 percent). This issue was remedied by matching student names.

on the results, we estimated the probability of matching a student in both the baseline and the endline as a function of the interventions, the baseline test score (which sums the math and language test scores) and the baseline values of the summary indices (see equation 2 in the next section for a definition of the indices). As shown in Table A2, there is no impact of the interventions on the probability of matching a student. The interaction between the baseline test score and the intervention is not significant at or below the 5 percent level. Only one baseline summary index appears to be important.<sup>30</sup>

We do find that students with lower baseline scores have a statistically significantly lower probability of being matched, but the size of the effect is small. For instance, for students in the lowest decile of the baseline test score distribution, 84 percent could be matched. We believe that most of the matching problems arose from problems in deciphering names, not student repetition or drop-out. School records show very low repetition and drop-out rates (Table 5), and we find no effect of the interventions on repetition or drop-out (see Section IV). We use the matched panel in the estimations.

We also designed surveys centered around hypothesized mechanisms that could improve learning outcomes. Broadly, these intermediate outcomes relate to stakeholder awareness of school committees, school-based management, parent, community and teacher inputs to education. We interviewed parents, teachers, students, school committee members, and principals. Administrative data and interviewer observations on infrastructure and teacher activities at the start of visit were also recorded. Because we wanted to focus interviews on teachers of students being tested, the teacher sample was restricted to teachers teaching grade four at baseline and grade six at endline. We then randomly selected three students from their classes, and these students' parents, for interview.

<sup>30</sup> Teacher-level inputs have a significantly positive effect on the probability of matching students between baseline and endline. We checked whether the interaction between the teacher-level inputs and the intervention was significant, and it was not.



### III. Impact Evaluation Strategy

In this section we discuss the empirical framework and how we address the issues arising from having many intermediate outcomes.

#### A. *Pairwise impact evaluation*

As discussed above, the objective of this study is to evaluate the effects of four treatments, independently and combined with each other. We analyze seven pairwise comparisons (see Table 3). The grant comparison is between the schools in the control group and those assigned the grant only, while all other comparisons measure the effect of assignment to the other interventions (election, linkage and training) and their combinations, conditional on being assigned the grant treatment. We chose to use seven bivariate comparisons that correspond to the hypotheses we set out to test.<sup>31</sup> All comparisons are ‘intent to treat’ effects. Compliance to the grant, training and linkage treatments is close to perfect, but not for elections (see Section I).

[Insert Table 3 here]

We need to apply weights in our regressions because the assignment of schools to the treatment and comparison groups is not balanced (see Table 2). In particular, the cell with grants only has almost twice as many schools as the other seven treatment cells because we wanted to evaluate the impact of the grant-only intervention by comparing this cell with the control group (the 100 schools assigned nothing).<sup>32</sup> We would not need weights if we had a randomized control trial with balanced assignments. For example, when evaluating the impact of linkage, balanced assignment ensures that the share of schools assigned to training and election treatments is equal for the linkage and no linkage groups. This is not the case for us. Table 2 shows that the

<sup>31</sup> We excluded the comparison that tests all interventions against grant only for lack of power as the former group consists of only 45 schools.

<sup>32</sup> Our initial power calculations suggested that each hypothesis required approximately 90 schools which is why the grant only and control groups have at least 90 schools

linkage group has 95 out of 190 schools assigned the election treatment but this share is 95/230 for the no linkage group. This is because the no linkage group includes the grant only treatment cell that has more schools.

In our analysis, we assign lower weights to cells in Table 2 that have more schools. Specifically, the weights are calculated as the total number of schools divided by the total number of schools in the cell. This way, observations in the grant only cell that has more schools will have a lower weight so that the weighted number of observations is balanced for each cell, as if we had balanced assignment. In the example above, the weighted number of observations assigned the election treatment would be balanced across the linkage and no linkage groups. This procedure is akin to using weights to correct for oversampling. Here, we have “over-assigned” schools to the grant only cell. We also ran regressions without weights that include controls for the cross-treatments and the results are similar, suggesting that the weights are not driving the results.

### B. *Impact on test scores*

The impact of the intervention on test scores is estimated by

$$(1) \quad y_{i,j, \text{endline}} = \alpha_k + \beta(\text{treatment}_j) + \gamma y_{i,j, \text{baseline}} + \varepsilon_{ij}$$

where  $y_{i,j}$  denotes the standardized test score of student  $i$  in school  $j$  in strata  $k$ . Test scores are standardized by subtracting the mean and dividing by the standard deviation of control group schools. Standard errors were clustered at the school level. The treatment variable equals 1 for the treatment group, and 0 for the comparison group. All regressions include strata dummies because random assignment was within each stratum.

### C. *Impact on intermediate outcomes*

To address the issue of having too many intermediate outcomes, we follow Kling, Liebman, and Katz (2007) and Banerjee et al. (2010) and construct summary indices for each domain of intermediate outcome variables. The advantage is that adding more outcomes to a summary index does not increase the number of hypotheses. The disadvantage is that the summary indices are harder to interpret and less transparent. We report results for the summary indices in Table 6 and results for each individual intermediate outcome in the appendix.

We define the summary index score for school  $j$  over the set of  $N_D$  outcome variables in group  $D$  as the mean of the z scores of the non-missing<sup>33</sup> outcome variables in a group. Each variable is constructed such that it contributes positively to the header or overall concept used for the domain.

$$(2) \quad y_{jD} = \frac{1}{N_D} \sum_{d=1}^{N_D} \frac{y_{jd} - \bar{y}_d}{\sigma_d}$$

where  $\bar{y}_d$  and  $\sigma_d$  are the mean and standard deviations of variable  $y_{jd}$  estimated from the control group schools. The summary index provides an equal weight to each variable that enters the summary index correcting for natural variation as observed in the control group.

Following Banerjee et al. (2010), we condition on the baseline values of all variables included in the summary index. For instance, for outcome variable  $d$  included in summary index  $D$ , the impact is estimated by

$$(3) \quad y_{j,d, \text{endline}} = \alpha_k + \beta(\text{treatment}_j) + \sum_{d=1}^{N_D} \gamma_d y_{j,d, \text{baseline}} + \varepsilon_{jd}$$

<sup>33</sup> There were some difficulties with non-response in some of the schools. We have complete information on 508 schools and partial information on nine schools.

For estimating the impact on the summary index itself,  $y_{j,d, \text{endline}}$  is replaced by  $y_{j,D, \text{endline}}$ .

Our challenge was to group the outcome variables in a way that made intuitive sense as separate pathways to learning. Our strategy was to define domains according to stakeholders – parents, teachers and community – soliciting answers from various respondents on how they support education. Two separate indices that focused on school committee awareness and school-based management, two important factors in the literature, were also included. Each impact domain is represented by a table (see Tables A6-A10). The list of variables in each domain, their definitions and the corresponding questionnaire are found in Table A5. The grouping is as follows:

- Awareness of school committee (Table A6)
- Parent level inputs to education (Table A7)
- Teacher level inputs to education (Table A8)
- School-based management (Table A9)
- Community level inputs to education (Table A10)

#### IV. Results

In this section we check pre-treatment balance, and discuss the impacts on education outcomes and intermediate outcomes.

##### A. *Checking baseline, pre-treatment differences*

Table 4 checks pre-treatment differences. We report estimates of  $\beta$  in (1) where the dependent variable is replaced by the baseline value and only the treatment variable and strata dummies are included on the right hand side. To save space, we present the results for the language and mathematics test scores and the summary indices only.

[Insert Table 4 here]

No difference can be found for the baseline test score comparisons and for most summary measures. Only one coefficient is significant at the five

percent level or below. Treatment schools in the linkage plus elections comparison have a higher value for the school-based management summary index at baseline. Our results are robust to controlling for this baseline difference.

### B. *Education outcomes (Tables 5, A3 and A4)*

We begin by measuring the impact of the interventions on the main education outcome variables — dropout rate, repetition rate, and test scores (Table 5). We do not find any significant effects on dropout and repetition rates, which is not surprising given the very low rates at baseline.

[Insert Table 5 here]

Looking at learning, we find substantial effects on Indonesian test scores, and no effects on mathematics. Linkage and linkage plus elections improve Indonesian scores by 0.17 and 0.23 standard deviations respectively. These two estimates are not statistically significantly different from each other. The first column of Table 5 reports the means and standard deviations (in brackets) of the unstandardized test scores for the control group at baseline and endline. The unstandardized test scores range from 0 to 30 because each test had 30 questions and each correct answer was awarded a point. Therefore, the statistics in Table 5 can be used to convert the effect sizes in standard deviations to the number of correct answers in each test by multiplying and adding the standard deviation and mean of the control group, respectively. We find no impacts on learning for the other comparisons.

The treatment effects for linkage and linkage plus elections are robust to adding various controls (Table A3). The other treatment effects remain insignificant. We tried adding district fixed effects, controlling for quadratic terms in the baseline test score, controlling separately for baseline language and mathematics scores, controlling for baseline summary indices and saturating the model with all these baseline controls.

The estimated treatment effect for the grant only intervention is smaller (0.129) and is not statistically significant. The grant comparison has the least statistical power because it includes the smallest number of schools (190). In our robustness checks, the grant effect remains insignificant for most robustness checks.<sup>34</sup> Grants are at least 2.5 times more costly to implement compared to the incremental costs for linkage and do not have a greater impact on Indonesian test scores.<sup>35</sup>

We think it is harder to detect treatment effects using mathematics scores because the endline mathematics test was particularly hard. We can see this in the data in several ways. First, the unstandardized test score means for mathematics are 16.4 and 8.9 for the baseline and endline tests respectively (Table 5). For Indonesian, the baseline and endline means are more similar (11.7 and 13.1 respectively). Second, there is less variation in the endline mathematics test. The standard deviation is lower (3.2 instead of 5.6 for the baseline). This is not true for Indonesian tests. Third, when we regress math test scores on school fixed effects, we can explain 14 percent of the variation but when we add a post treatment dummy, the R-squared increases to 0.59. For language scores, the R-squared only increases from 0.23 to 0.26. In other words, the post treatment dummy is explaining much of the variation in the math tests, even more than school fixed effects. These findings suggest that the endline mathematics test was not as useful at demonstrating a range of abilities as was the Indonesian test because it was too hard.

Looking at subgroups, the treatment effects for linkage and linkage plus elections are slightly larger for girls than for boys.<sup>36</sup> For girls, linkage also increases math scores by 0.11 standard deviations. Table A4 reports treatment

<sup>34</sup> In Table A3, only when we add controls for baseline summary indices does the grant only treatment effect on Indonesian test scores increase to 0.16 (significant at the 10 percent level).

<sup>35</sup> Calculated using \$321 as the cost to implement the grant (see Section I) compared to \$125, the incremental cost to implement linkage. If we included the cost of the grant itself (\$870), grants would be less cost-effective. The largest estimate for the grant treatment (0.16 in Table A3) is slightly smaller than the effect size for linkage in our preferred specification (0.17).

<sup>36</sup> The students' gender variable was constructed *ex-post* using the name of the student.

effects by quintile of the baseline test scores. Impacts are generally higher for students who already did well in the baseline test.

Figure 1 shows probability density functions of the scores for all the tests. There is no evidence of ceiling or floor effects for the mathematics scores – in both years the distributions have the expected bell shape curves. The same is true for Indonesian scores at baseline, but the endline distribution has two humps, where the right hump is censored above by the maximum score. There is however no clear excess mass at the maximum score, suggesting limited truncation.

[Insert Figure 1 here]

### C. *Cost-Effectiveness*

Using the costs discussed in Section I,<sup>37</sup> we find that the linkage intervention is the most cost-effective at \$0.44 per child per tenth of a standard deviation.<sup>38</sup> Linkage plus elections costs \$0.78 per child per tenth of a standard deviation. These incremental gains are achieved at small incremental costs compared to the grant only treatment.<sup>39</sup> We use a tenth of a standard deviation because this is widely considered to be the minimum effect size relevant for policy, and has been used in other literature looking at the impacts on test scores compared to cost (He, Linden, and MacLeod 2008; Linden 2008; Burde and Linden 2012).

We consider intervention and facilitation costs. Specifically, as discussed above, linkage alone costs \$125.47 per school, and linkage plus elections costs \$300.43 per school. The per student calculations utilize 164 as

<sup>37</sup> The only cost not mentioned in Section I was overhead costs (team leader, office space, etc.) allocated equally to all intervention schools, which amounted to US\$ 140 per school.

<sup>38</sup> Calculated as  $\$125.47 / (164 \times 10 \times 0.173)$ .

<sup>39</sup> We calculate that the best possible cost-effective estimate for the grant only treatment is \$2.47, which is much higher than the cost-effectiveness for linkage conditional on grants. The \$2.47 estimate was calculated as total outlays as of the endline survey (\$326 of grants disbursed plus \$321 general facilitation cost), divided by the largest possible effect per child per tenth of a standard deviation ( $0.16 \times 164 \times 10$ ), where 0.16 is the largest effect size for the grant only treatment in our robustness checks (Table A3). If we included the full grant amount in the numerator (\$870 instead of \$326), the estimate for grants is even higher.

the average school size, since the entire school was affected by the intervention, not just the students who were tested.

There are several things worth noting that were not included in the cost calculations. First, we did not include the cost of general facilitation and the grant. We think this is appropriate since the linkage and linkage plus elections effects are estimated, conditional on the grant treatment. That is, the numerator in our cost-effectiveness calculation differences out the cost of the grant treatment because the denominator differences out the effect of the grant treatment. While it is possible that the linkage, election and training interventions would not be as effective without the general facilitation and grant being available, this hypothesis is not testable as these interventions were always implemented in combination with the grant. Second, we did not factor in the opportunity costs of the school committee members' time which would decrease cost-effectiveness. Third, we did not amortize the cost of the interventions over time, which could increase cost-effectiveness. We decided against the last two factors because we wanted our cost-effectiveness estimates to be comparable to as large a set of estimates as possible, and most cost-effectiveness estimates for test scores do not account for these factors.<sup>40</sup>

Compared to other learning interventions, linkage and linkage plus elections are extremely cost-effective. A village-based school construction program in Afghanistan costs \$4.80 per child per tenth of a standard deviation (Burde and Linden 2012), an after-school computer-based learning program in India costs \$4.59 per child per tenth of a standard deviation (Linden 2008), and a girls' scholarship in Kenya costs between \$1.77 to \$3.53 per child per tenth of a standard deviation (Kremer, Miguel, and Thornton 2009).

Other programs with similar cost-effectiveness are machine or flash card-based activities that teach English to Indian children that cost \$0.22 per

<sup>40</sup> Dhaliwal et al. (2012) provide a comprehensive set of cost-effectiveness estimates that account for these factors. But, they do not provide estimates for effects on test scores, which is our focus.



student per tenth of a standard deviation (He, Linden, and MacLeod 2008), and an Indian remedial teacher education program that cost \$1 per tenth of a standard deviation (Banerjee et al. 2007, cited in Linden 2008).

#### D. *Intermediate outcomes (Table 6, Tables A6 – A10)*

In this section, we focus on why linkage and linkage plus elections impact learning, then briefly discuss the mechanisms driven by the other interventions. Improvements in these areas may be goals themselves, and they were hypothesized to ultimately lead to improved learning. We only focus on estimates that are significant at the 1 percent and 5 percent levels unless stated otherwise.

#### E. *Linkage and Linkage Plus Elections*

Linkage increases community level inputs by 0.14 standardized units while other summary indices are not significant (Table 6). Linkage plus elections increases community level inputs by 0.13 standardized units but the estimates are only significant at the 10 percent level. Table A10 reports estimates for individual components of the community level index.

[Insert Table 6 here]

All linkage-related interventions enhance collaboration between the school committee and the village council (*SCbpd*, *SPbpd*). Principals demonstrate improved satisfaction with village councils' attention to education in the village (*SPsatbpd*). Parents are more satisfied with support from the community (*Psatcomm*).

Cooperation with the village council appears to be central. School committees in the linkage treatment groups do not report more cooperation with any non-educational community organizations other than the village council (*SCnonbpd*). Interestingly, with the linkage intervention, meetings between the school committee and the local education office decline significantly (*SCmeetdinas*, Table A6), suggesting that the local education

office may have pulled back as a result of greater school engagement by the village council.

The community support appears to be unrelated to donations. School committee representatives perceive no changes in community support in areas such as in-kind or financial contributions (*SCcomfundraise*, *SCcominkind*). This is consistent with our hypothesis that the linkage intervention improved learning because the joint planning meetings translated into co-sponsored education initiatives. This is in marked contrast to the school committees' predecessor, the BP3, which was largely perceived as a passive fundraising arm for the school.

While the summary index for teacher level inputs is insignificant, an additional factor that could have contributed to the learning gains of the linkage plus elections intervention is teacher work hours (*Thours*, see Table A8). This increases by one hour per week for this intervention. We chose to highlight this particular result as teacher work hours seems closely related to the learning experience students have in school.

#### F. *Other Interventions*

Grants increase awareness of school committees (Table A6). Likewise for elections and training plus elections; however, their effects are only significant at the 10 percent level. Grants increase the likelihood that parents know the school committee exists and know the names of the school committee members (*Pknowexist*, *Pknow\_scmem*). The increased awareness, however, does not translate to increased interaction between parents and school committees (*SCmeetparents*). Only internal meetings of the school committee are impacted positively by the grant (*SCintmeettot*).

We do not see much traction in the index for parent level inputs (Table A7). Elections and training plus elections increase parent level inputs but are only significant at the 10 percent level. There is an increase in the total number of minutes that household members accompany a child studying at

home in the past week (*Pallhh\_min*).<sup>41</sup> Training increases in kind contributions of parents to the school committee (*SCparinkind*) but decreases the likelihood that parents ever visit to observe the class (*Pvisit*). There is no impact on stakeholder satisfaction with parents' support for pupils' learning (*Psatparents*, *SCsatparents*, *Tsatpar*, *SPsatpar*), nor do we see parents engage more with schools (*SPparentsinvolve*, *Pmeet\_teacher*).

There is no effect on the summary index for teacher inputs (Table A8). We see some increases in teacher hours (*Thours*) for elections and decreases in the fraction of classes observed to have teachers (*OBfractwithteach*) for election and training plus elections.

We see no effects on the summary index for school-based management (Table A9). Principals in the grant treatment report providing more information to parents about school funding and budgeting (*SPparentsrapbs*). Finally, training appears to have an impact on community level inputs but only if paired with linkage (Table A10).

## V. Conclusion

This paper studies the impact of four treatments aimed at increasing the effectiveness of school committees in public schools in Indonesia. We find that measures that reinforce existing school committees, grant and training, demonstrate limited effects; while measures that foster ties between school committees and other parties, linkage and elections combined with linkage, lead to greater engagement by education stakeholders and in turn to learning. The results suggest that these institutional reforms can be highly cost-effective in an environment where school committees have access to resources.

<sup>41</sup> The survey asked fathers, mothers and anyone else in the household to report the number of hours per day and the number of days a week that they accompanied his/her child studying. Field workers were then instructed to convert hours per day to minutes, so there was the potential for rounding errors, and we see this with peaks in the data on the hour (e.g., 60, 120, 180 minutes).

In considering the mechanisms from treatment to greater learning, we find that the results arise mostly from increased community support. The election intervention was the most promising at reforming the school committee into an institution that engages community members and improves service delivery. Indeed, elections raised the awareness of the school committee, increased parental support for homework and teachers reported more work hours. However, elections alone were insufficient to raise learning outcomes.

We think that the success of the linkage intervention results from the fact that a more powerful community institution, the village council, was involved in the planning of the activities. This provided the legitimacy needed to ensure that actions that could improve learning were implemented.

The grant and training interventions, which focused on fortifying the existing committee, showed little promise. This resonates with the oft-voiced view that focusing on increasing resources, in this case financial and human resources, is not sufficient to make an institution perform better.

## REFERENCES

- Acemoglu, Daron, Simon Johnson, and James A. Robinson.** 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369-401.
- Antlov, Hans.** 2003. "Village Government and Rural Development in Indonesia: The New Democratic Framework." *Bulletin of Indonesian Economic Studies* 39 (2): 193-214.
- Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani.** 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* 2 (1): 1-30.
- Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden.** 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics* 122 (3): 1235-64.
- Banerjee, Abhijit V., and Esther Duflo.** 2008. "Mandated Empowerment - Handing Antipoverty Policy Back to the Poor?" *Reducing the Impact of Poverty on Health and Human Development: Scientific Approaches* 1136: 333-41.
- Behrman, Jere R., Anil B. Deolalikar, and Lee-Ying Soon.** 2002. "Promoting Effective Schooling through Education Decentralization in Bangladesh, Indonesia, and Philippines," In: Asian Development Bank ERD Working Paper No. 23.

- Bjork, Christopher.** 2009. "Improving Educational Quality through Community Participation - Qualitative Study," In: World Bank Report.
- Björkman, Martina, and Jakob Svensson.** 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *The Quarterly Journal of Economics* 124 (2): 735-69.
- Blimpo, Moussa P., and David K. Evans.** 2010. "School Based Management, Local Capacity, and Educational Outcomes Lessons from a Randomized Field Experiment," In <http://www.stanford.edu/~mpblimpo/research.html>.
- Bruns, Barbara, Deon Filmer, and Harry Anthony Patrinos.** 2011. *Making Schools Work: New Evidence on Accountability Reforms*. The World Bank
- Burde, Dana, and Leigh L. Linden.** 2012. "The Effect of Village-Based Schools: Evidence from a Randomized Controlled Trial in Afghanistan," In: National Bureau of Economic Research Working Paper 18039.
- Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch.** 2012. "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education." *Abdul Latif Jameel Poverty Action Lab (J-PAL), MIT*.
- Duflo, Esther, Pascaline Dupas and Michael Kremer.** 2009. "Additional Resources Versus Organizational Changes in Education: Experimental Evidence from Kenya," In: <http://www.poverty-action.org/node/2082>

**Fearnley-Sander, Mary, Pahala Nainggolan, Mike Ratcliffe, Abby Riddell, Simone Seper, and George Taylor.** 2008. "Support to Basic Education in Indonesia. Report of the Joint European Community/Ausaid Pre-Feasibility Study Mission. *Unpublished.*

**Gertler, Paul, Harry Patrinos, and Eduardo Rodríguez-Oreggia.** 2010. "Parental Empowerment in Mexico: Randomized Experiment of the Apoyo a La Gestión Escolar (Age) in Rural Primary Schools in Mexico: Preliminary Findings," Washington, DC.: The World Bank.

**Hanushek, Eric A., and Ludger Woessmann.** 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46 (3): 607-68.

**He, Fang, Leigh L. Linden, and Margaret MacLeod.** 2008. "How to Teach English in India: Testing the Relative Productivity of Instruction Methods within the Pratham English Language Education Program." *working paper.*

**Khatti, Nidhi, Cristina Ling, and Shreyasi Jha.** 2010. "The Effects of School-Based Management in the Philippines: An Initial Assessment Using Administrative Data," In *World Bank Policy Research Working Paper*, edited by The World Bank.

**Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz.** 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83-119.

- Kremer, Michael, and Alaka Holla.** 2009. "Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?," In *Annual Review of Economics*: 513-42.
- Kremer, Michael, Edward Miguel, and Rebecca Thornton.** 2009. "Incentives to Learn." *Review of Economics and Statistics* 91 (3): 437-56.
- Kristiansen, Stein, and Pratikno.** 2006. "Decentralising Education in Indonesia." *International Journal of Educational Development* 26 (5): 513-31.
- Linden, Leigh L.** 2008. "Complement or Substitute? The Effect Of Technology on Student Achievement in India." *downloaded from <http://www.povertyactionlab.org/publication/complement-substitute-effect-technology-student-achievement-india>.*
- Mansuri, Ghazala, and Vijayendra Rao.** 2012. *Localizing Development: Does Participation Work?* Washington DC: The World Bank.
- Martinez-Bravo, Monica, Gerard Padró i Miquel, Nancy Qian, and Yang Yao.** 2011. "Do Local Elections in Non-Democracies Increase Accountability? Evidence from Rural China," In *NBER working paper*.
- Mullis, Ina V.S. , M.O. Martin, and P. Foy.** 2008. "Timss 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades," Boston: TIMSS & PIRLS International Study Center. Boston College, Chestnut Hill, MA.



- Olken, Benjamin A.** 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200-49.
- Olken, Benjamin A.** 2010. "Direct Democracy and Local Public Goods: Evidence from a Field Experiment in Indonesia." *American Political Science Review* 104 (2): 243-67.
- Olken, Benjamin A., Junko Onishi, and Susan Wong.** 2012. "Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia," *NBER working paper*.
- Organisation for Economic Co-operation and Development.** 2010. *Pisa 2009 Results*. Paris: OECD Publishing.
- Ostrom, Elinor, and T.K. Ahn.** 2009. "The Meaning of Social Capital and Its Link to Collective Action" In *Handbook of Social Capital*, edited by Gert T. Svendsen and Gunnar L. H. Svendsen. Edward Elgar.
- Pandey, Priyanka, Sangeeta Goyal, and Venkatesh Sundararaman.** 2009. "Community Participation in Public Schools: Impact of Information Campaigns in Three Indian States." *Education Economics* 17 (3): 355-75.
- Stiglitz, Joseph E.** 2002. "Participation and Development: Perspectives from the Comprehensive Development Paradigm." *Review of Development Economics* 6 (2): 163-82.

**The World Bank.** 2003. "Making Services Work for Poor People," In *World Development Report*, edited by Shantayanan Devarajan and Ritva Reinikka-Soininen. Oxford University Press, The World Bank.

**Table 1: Adherence to design (Percent of intent to treat)**

	Fully Implemented (1)	Partially Implemented (2)	Not Implemented (3)
Election	47.9	44.7	7.4
Grant	98.8	0	1.2
Training	100	0	0
Linkage	98.4	0	1.6

Note: Partial implementation for the election means previously underrepresented groups were elected while other members stayed on.

**Table 2: Allocation of schools to treatments (number of schools)**

Receiving block grant	No election		Election		Total
	Linkage	No Linkage	Linkage	No Linkage	
No Training	50	90	50	50	240
Training	45	45	45	45	180
Total	95	135	95	95	420

Control Group, not receiving block grant, no intervention: 100 schools

**Table 3: Impact evaluation framework**

Comparison	Treatment	Number of schools	Control	Number of schools
Grant	Grant-only	90	No grant	100
Election	Grant + Election	190	Grant + No Election	230
Linkage	Grant + Linkage	190	Grant + No Linkage	230
Training	Grant + Training	180	Grant + No Training	240
Linkage + Election	Grant + Linkage + Election	95	Grant + No Linkage + No Election	135
Linkage + Training	Grant + Linkage + Training	90	Grant + No Linkage + No Training	140
Training + Election	Grant + Training + Election	90	Grant + No Training + No Election	140

**Table 4: Tests of pre-treatment balance in observables across interventions**

	Grant, G	Election, E	Linkage, L	Training, T	L+E	L+T	T+E
	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Language test scores	0.133 (0.093)	-0.047 (0.064)	0.001 (0.064)	-0.054 (0.065)	-0.047 (0.088)	-0.050 (0.092)	-0.101 (0.087)
Mathematics test scores	0.104 (0.102)	-0.059 (0.067)	0.015 (0.068)	0.004 (0.068)	-0.046 (0.101)	0.022 (0.098)	-0.052 (0.090)
Awareness of school committee	-0.047 (0.058)	0.026 (0.038)	0.008 (0.038)	-0.040 (0.038)	0.033 (0.055)	-0.033 (0.052)	-0.014 (0.049)
Parent level inputs to education	-0.005 (0.046)	-0.029 (0.038)	0.002 (0.038)	-0.007 (0.038)	-0.028 (0.056)	-0.005 (0.053)	-0.036 (0.050)
Teacher level inputs to education	0.017 (0.041)	0.035 (0.030)	-0.002 (0.030)	-0.014 (0.030)	0.033 (0.041)	-0.015 (0.038)	0.021 (0.043)
School based management	-0.042 (0.049)	0.061* (0.035)	0.049 (0.035)	-0.043 (0.036)	0.109** (0.049)	0.005 (0.046)	0.018 (0.046)
Community level inputs to education	-0.077 (0.061)	-0.017 (0.045)	0.020 (0.045)	-0.086* (0.045)	0.003 (0.061)	-0.066 (0.063)	-0.103* (0.055)

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Each cell reports the estimated treatment effects using OLS. All estimations include stratum fixed effects because assignment of treatment was within each stratum. Robust standard errors reported in the parentheses. All standard errors for regressions with test scores are clustered at the school level.

**Table 5: Impact on drop out, repetition and test scores**

	Pre/Post Mean and SD (1)	Grant, G OLS (2)	Election, E OLS (3)	Linkage, L OLS (4)	Training, T OLS (5)	L+E OLS (6)	L+T OLS (7)	T+E OLS (8)
<u>Panel A: Drop out and repetition rates</u>								
Drop out	0.002/0.01 [0.01/0.05]	-0.005 (0.005)	-0.003 (0.006)	-0.002 (0.006)	0.007 (0.006)	-0.005 (0.011)	0.003 (0.006)	0.004 (0.006)
Repetition	0.02/0.03 [0.04/0.06]	-0.004 (0.007)	-0.001 (0.005)	0.007 (0.005)	-0.006 (0.005)	0.007 (0.008)	0.001 (0.009)	-0.007 (0.008)
<u>Panel B: Language test scores (average, by gender)</u>								
Average	11.66/13.09 [4.32/6.34]	0.129 (0.094)	0.053 (0.069)	0.173** (0.068)	-0.042 (0.069)	0.234** (0.094)	0.134 (0.087)	0.015 (0.103)
Boys	11.38/13.16 [4.13/6.57]	0.085 (0.105)	0.025 (0.078)	0.156** (0.078)	-0.044 (0.078)	0.187* (0.101)	0.115 (0.101)	-0.020 (0.115)
Girls	11.93/13.04 [4.48/6.13]	0.167* (0.093)	0.073 (0.069)	0.191*** (0.068)	-0.040 (0.069)	0.271*** (0.099)	0.152* (0.088)	0.039 (0.101)
<u>Panel C: Mathematics test scores (average, by gender)</u>								
Average	16.42/8.88 [5.61/3.17]	-0.015 (0.080)	-0.008 (0.050)	0.070 (0.050)	-0.029 (0.050)	0.061 (0.075)	0.040 (0.068)	-0.036 (0.066)
Boys	16.38/8.98 [5.67/3.10]	0.002 (0.085)	-0.030 (0.058)	0.026 (0.058)	-0.021 (0.059)	-0.003 (0.089)	0.001 (0.080)	-0.051 (0.071)
Girls	16.49/8.78 [5.55/3.23]	-0.032 (0.088)	0.009 (0.053)	0.113** (0.053)	-0.035 (0.053)	0.121 (0.074)	0.079 (0.072)	-0.025 (0.074)

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Dropout and repetition are fractions obtained from aggregate, school level administrative data, test scores are z-scores standardized using the baseline and endline means and standard deviations for the control group. Column 1 reports means and standard deviations for the control group in the baseline and endline (for each outcome, the means are in the first row and the standard deviations are in the second row, in brackets). For the test scores, we report the mean and standard deviation of the unstandardized test scores for the control group (what we used to calculate the z-scores). The unstandardized scores are on a scale of 0 to 30 because each test had 30 questions and each correct question was awarded one point. The following columns report the estimated treatment effects using OLS. All estimations include stratum fixed effects because assignment of treatment was within each stratum. Robust standard errors reported in the parentheses. All standard errors for regressions with test scores are clustered at the school level.

**Table 6: Impact on summary index of intermediate outcomes**

	Grant, G	Election, E	Linkage, L	Training, T	L+E	L+T	T+E
	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(4)	(5)	(6)	(7)	(8)
Awareness of school committee	0.154*** (0.054)	0.074* (0.041)	-0.012 (0.042)	0.049 (0.041)	0.048 (0.060)	0.025 (0.060)	0.110* (0.060)
Parent level inputs to education	0.023 (0.054)	0.062* (0.034)	0.025 (0.032)	0.026 (0.033)	0.070 (0.043)	0.048 (0.044)	0.074* (0.044)
Teacher level inputs to education	-0.027 (0.042)	0.011 (0.032)	0.022 (0.030)	-0.001 (0.030)	0.015 (0.046)	0.024 (0.039)	0.016 (0.041)
School based management	0.055 (0.058)	-0.030 (0.035)	0.017 (0.035)	-0.018 (0.036)	-0.018 (0.050)	0.006 (0.050)	-0.068 (0.054)
Community level inputs to education	0.008 (0.073)	-0.019 (0.051)	0.141*** (0.051)	0.095* (0.051)	0.126* (0.072)	0.237*** (0.066)	0.062 (0.070)

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Each cell reports the estimated treatment effects using OLS. All estimations include stratum fixed effects because assignment of treatment was within each stratum. Robust standard errors reported in the parentheses. See Tables A6 to A10 for effects on individual outcomes of each summary index.

Figure 1: Probability density functions of test scores

