Community Engagement in Schools: Evidence from a Field Experiment in Pakistan

Salman Asim[¥]
Amina Riaz[*]

Abstract

We present findings from a field experiment in rural Sindh, Pakistan, where half of school-age children (6-10 years) are out of school. This paper presents results of testing alternative approaches to strengthen engagement of communities with schools: face-to-face dialogue at externally-facilitated community meetings; and ongoing, anonymous dialogue via text messages. The interventions increased communities' interest in education as measured through an improvement in the number of functioning schools and, in the case of the text message treatment, substantial gains in retention of students in Grades 2, 3 and 4. On the supply side, schools were able to significantly increase staffing and reduce the share of one-teacher-schools; however, teacher absenteeism increased, and there was no substantial impact on basic school infrastructure. An additional reform to an existing institution, elections and capacity building for school committees, was implemented in a cross-over experimental design; the intervention undermined the participation of communities in meetings and reduced impacts on all indicators except new admissions and availability of toilets in schools. Overall, we find no evidence of impact on measured test scores for any intervention. The findings suggest that community engagement can achieve some impacts on school access and quality, but even when communities are engaged with schools, there is a limit to what they can do to fill in the administrative gaps left by the state.

JEL Codes: D78, I21, I28, H75, O15

¥ Corresponding author: Salman Asim, Senior Economist, World Bank; email: sasim@worldbank.org;
* Graduate student at the University of Chicago.

Acknowledgments

We thank the Government of Sindh's Education and Literacy Department and the Reform Support Unit, in particular, Dr. Fazlullah Pechoho, Saba Mushtaq, Faisal Ahmed Uqaili, Dr. Hamzo Khan Tagar, Muhammad Nasim Qureshi and Sania Khursheed, for extensive collaboration and assistance; the World Bank's intervention design team, in particular Priyanka Pandey, Lorenza De Icaza and, Mariam Nusrat Adil; the Weitek Group for field implementation; Amar Fateen and Saindad Joyo for implementation support; RCons for school and household surveys; Ravinder Casley Gera for extensive support to the drafting process; Dhushyanth Raju, Nazmul Chaudhury, and Syeda Shahbano Ijaz for substantive technical discussions and edits; and Ilagno Patchamutho, Rachid Benmessaoud, Amit Dar, Keiko Miwa, Umbreen Arif and Margo Hoftijzer for their encouragement and support for this project. Zunaira Mughal, Robbie Dean, Joshua Gill, Cameron Friday, Ali Abbas, Amn Nasir, Tooba Akhtar, Xu Wang and Abhijit Banerjee provided outstanding research assistance at various stages of the project. The authors thank Alaka Holla, Amer Hassan, Juliea M. Trias, Deon Filmer, Owen Ozier, Achim Schmillen and David Evans for helpful discussions and insightful comments. We thank discussants and participants at the 2016 RISE Annual Conference at the Blavatnik School of Government, University of Oxford, and the 2018 PacDev Conference at University of California, Davis. We gratefully acknowledge the financial support provided for this project by the Strategic Impact Evaluation Fund (SIEF), the Education Sector Development Trust Fund. The findings, interpretations and conclusions expressed in this paper are entirely those of the authors, and do not necessarily represent the views of the World Bank, its Executive Directors or the countries they represent.

Dedication

The study is dedicated to the living memory of Mr. Muhammad Nouman Bashir, Managing Partner of Weitek Group, Pakistan, the implementation support firm. Mr. Bashir met an untimely demise during the implementation of the intervention.

1. Introduction

Engagement of citizens in dialogue, monitoring, and decision-making in public services has been widely adopted since 2000 to address state failure in developing countries (World Bank, 2004; Heller and Rao, 2015). Community focused projects in Asia, Africa and South America have demonstrated that participation of the community can improve outcomes in the right set of conditions. In India, for example, *gram panchayats*, village councils, were introduced in the 1990s to devolve decision-making over local finance to village level; villages which held regular annual village-wide meetings performed better at distributing poverty benefits to the most disadvantaged citizens (Mansuri and Rao, 2013). In Uganda, meetings between residents and health clinic staff improved health outcomes (Bjorkman and Svensson, 2009); in Kenya, giving parents training to interview, recruit, and monitor performance of contract teachers reduced nepotism in hiring (Duflo et al., 2015). In Indonesia, development of links between school committees and local governing bodies, such as village councils, had substantial impacts on quality of schools (Pradhan et. al., 2014).

In the case of education, attempts to engage communities have focused primarily on School Management Committees (SMCs)—elected local bodies that create agency and voice for community members, allowing them to have more say in the decision-making process. These committees are typically given control over school grants to respond to immediate needs of the school. Committee members may also be given responsibility to monitor teachers and ensure school-age children attend school. Such reforms have been associated with increased enrollment, reduced dropout, and improvements in school facilities and staffing in countries including Cambodia, Honduras and Mexico (Benveniste and Marshall, 2004; di Gropello and Marshall, 2005; Gertler et al, 2011). However, other countries continue to face several barriers in strengthening community engagement in public services, including low awareness, high opportunity costs, cultural and social restrictions, and low levels of trust in the state (Abraham and Platteau, 2002; De Grauwe, 2005; Joshi, 2013; Croke et al., 2015). Thus, SMCs have not always proven capable of naturally emerging as credible channels of collective action as envisaged, and in many cases suffer from low participation rates and elite capture as noted in reviews of empirical studies (Barrera-Osorio et al., 2009; Bruns et al., 2011).

Government-sponsored reforms to induce community participation in schools have faced similar challenges in Pakistan. The context for this study is Sindh – the second-largest province of Pakistan, which has a dense network of public schools that frequently lack basic facilities. More than half of these are one-teacher schools which are prone to temporary closures. Combination of resource shortages, poverty and weak state capacity results in sub-optimal provision of education in Sindh, leaving a large majority of rural children out of school (Pakistan Social and Living Standards Measurement (PSLM) survey, 2010-11). Dominance of feudal landlords and high levels of ethnic and caste fragmentation in Sindh frustrate any meaningful attempt to invigorate participation of communities to improve service delivery (Miguel and Gugerty, 2005; Hussain et al., 2013).

Community engagement policies, introduced in 2001, were intended to address this problem by moving decision-making closer to schools. School grants were introduced to transfer authority to elected SMCs, promoting communities' interest in managing schools, making teachers

accountable, and encouraging parents to enroll and retain children in schools. As of June 2012, the Sindh Government was successful in ensuring timely transfers of SMC grants to more than 80% of public primary schools reported to be functioning (open) in the Annual School Census (2011/12). However, these funds remained largely underutilized, and in certain cases were misappropriated. Transfer of funds alone was inadequate in achieving the intended engagement of communities in the school improvement process. A potential explanation is the lack of a mechanism to induce inclusive, broad-based participation of communities with SMCs and schools. In the absence of such mechanisms, Bardhan and Mookherjee (2000) note, decentralized structures are prone to elite capture.

This paper provides evidence from a field experiment on the impact of alternative approaches to create an inclusive, community-wide interface for parents to engage with other community members. The first intervention, MEET, supports external facilitation of traditional face-to-face dialogue with community members. In a community-wide meeting, participants were inspired about the salience of education through a locally adapted audio drama clip; informed about ways to support, monitor, and improve their local schools; provided a space to discuss and deliberate on issues related to school performance; and encouraged to arrange follow-up meetings and reach out to SMC members to demand action on specific issues in local schools.[1] It was expected that external facilitators, responsible for social mobilization, meeting administration, and moderated dialogue, would create an inclusive space for the community members to talk openly about challenges faced by local schools. This dialogue was expected to support the incorporation of local knowledge and preferences in school management.

Our second intervention, SMS, tests an alternative low-cost approach using Information Communication Technology (ICT), specifically a text message-based Community Dialogue Platform (CDP). In place of an externally facilitated discussion in a village meeting, SMS supported continuous discussion via text message over a period of five months. Community members, registered on the CDP via a village-wide demonstration meeting, had the ability to send text messages expressing opinions or concerns about their local school. Messages were then anonymized and shared with all registered participants in the form of a weekly message summarizing key issues raised. Compared to the standard face-to-face dialogue supported by MEET, the SMS intervention was intended to lower the transaction costs for participation by allowing continuous, low-cost dialogue at the convenience of users.

The key innovation of this approach was anonymous exchange of views to encourage traditionally excluded voices to participate. The intervention had the explicit objective to reverse power relations by providing a platform where anyone could voice discontent without the need to acquiesce to power elites. In addition, registered CDP participants, even those who did not actively send out text messages, received a regular summary message of key issues raised on the CDP. This diffusion of local knowledge and preferences could potentially start informal discussions in community spaces about the salience of education, applying pressure to authorities for

[1] Under the Sindh system, the wider school community is considered part of an SMC General Body, which has a deliberative and ratifying role through village-wide meetings; while a five-person elected SMC Executive Body carries out the majority of duties stipulated in Government of Sindh, School Management Committee guidelines, 2009. We use *SMC members* to refer to members of the Executive Body and *village-wide meeting* to refer to the General Body.

improvement in schools and strengthening the capacity of community members to undertake self-initiated school improvement activities.

Our third intervention, ELECTIONS-CAPACITY, seeks to reform the existing SMC institution to be more effective and more responsive to the preferences of the wider community. This intervention implements elections of SMC members in community-wide meetings, to address elite capture, improve alignment between committee members and the community, and bring new members into committees. We employ a cross-over design with the third intervention, implementing it as an additional treatment in almost half of randomly assigned villages in each of the dialogue interventions.

In recent years, a number of experimental studies have tested the impact of various mechanisms to boost community engagement. Two studies from India suggest that furnishing citizens with information about existing educational institutions through meetings conducted in the presence of village administrators had no impact on school attendance or student learning outcomes (Banerjee et al., 2010); yet when similar information was provided by external facilitators through an innovative medium (scripted video drama), it had a significant impact on teachers' attendance and on student learning outcomes (Pandey et al., 2009). Cerdan-Infantes and Filmer (2015) find that engaging parents through facilitated village meetings increased awareness of a school management program significantly, as did providing information through text messages.

An experiment in Uganda which provided participatory training to SMC members to monitor school quality had significant impacts on student and teacher attendance, and on learning outcomes, while an alternative less participatory training had smaller effects (Barr et al., 2012). However, other studies in low-income settings have found null impacts from attempts to facilitate communities to monitor school quality (Blimpo and Evans, 2011; Beasley and Huillery, 2015). Pradhan (2014) provided grants to SMCs and either a program of classroom-based training, or more participatory facilitation, for the use of the grant. Neither approach led to significant improvements in school learning outcomes. However, a sub-treatment which directly supported linkages between SMCs and village councils led to significant improvements to community contributions and learning outcomes. An additional treatment, which supported elections for SMC members, strengthened the observed effects.

We contribute to the literature on facilitation of community participation in education in several ways. To our best knowledge, this is the first study that credibly estimates the causal impact of an inclusive, community-wide dialogue interface and benchmarks the results against the standard treatment of externally-facilitated face-to-face meetings. In addition, like Pradhan et al (2014), we add to the literature by estimating the causal impact of elections of SMC members using a randomized evaluation design.

All three interventions were piloted in 284 villages, containing 479 schools, in three representative districts of Sindh, between January and June of 2013. The interventions were expected to have an impact on several school characteristics symptomatic of systemic problems in education service delivery in Sindh (Asim, 2013): school functioning (the open/closed status of school on day of visit), enrollment, availability of teachers, teacher presence, infrastructure (e.g., classrooms, drinking water and toilets), and learning outcomes.

We find evidence of increased access to schooling for both the MEET and SMS interventions. Schools participating in MEET were, on average, more likely to be open at endline than those in the control group. Additionally, these schools were less likely to be closed if previously found open at baseline. Furthermore, schools participating in the SMS intervention were more likely, on average, to have reopened if closed at baseline. The SMS intervention was also associated with a significant increase in the number of new Grade 1 students enrolling in schools. Both interventions increased the number of registered teachers at schools, and significantly reduced the proportion of one-teacher schools, potentially contributing to some of the observed impact on school functioning. However, neither intervention improved overall teacher presence. In fact, teacher absenteeism rates, on average, increased in the treatment groups, which reflects a reduction in effort by existing teachers. We are unable to find impact on test scores in English and Mathematics from either of these interventions.

While the addition of the SMC-focused elections and capacity building intervention led to improvements in pre-primary and Grade 1 enrollment, it reduced both the impact of MEET on school functioning, and the impact of the SMS intervention on retention of pupils and number of registered teachers. These results stand in sharp contrast to earlier findings by Pradhan et. al. (2014) where elections strengthen the impacts on learning outcomes. In this particular context, holding of SMC elections appears to have reduced the level of community participation in meetings; a potential explanation is that the presence of a government education official, Taluka Education Officer (TEO) overseeing the elections, reduced the perceived independence of the interventions.[2]

The rest of this paper is structured as follows. Section 2 describes the experiment design, the sampling framework, and the study timeline. Section 3 discusses the data and estimation strategy. Section 4 presents results. Section 5 discusses the findings. Section 6 concludes.

2. Experiment

2.1 Context

Sindh, Pakistan's second largest province, provides a particularly informative environment to test the impact of our interventions, owing to widespread service failures in access to education and a local political economy which holds back community engagement in local institutions.

Sindh has a feudal political economy in rural areas: 76 percent of families are landless and work as tenants, while less than one percent of rural Sindhi families own land holdings of 50 acres or more and are considered influential (Bengali, 2015). These landlords often dominate economic and political life, exercising control over land, irrigation, credit and service delivery. Many families are bound to their landlord by generations of debt (Hussain et al., 2013). Sindh is also

[2] Community meetings took place in the main settlement of the village, in most cases at/outside the premises of the larger, more central school in the village. While the attendees were walking to the announced venue of the meeting (publicized three days in advance), they were able to observe the presence of external facilitators, and in this case of the Taluka Education Officer (TEO). The sharp decline in number of attendees observed in the ELECTIONS group appears to be the result of general distrust in the Government officials.

subject to a particularly large degree of ethnic and caste fragmentation. Villages function as the lowest administrative unit, but are typically comprised of spatially dispersed hamlets, *paros*, each primarily representing a caste or kinship group.

These feudal landholding patterns and fragmentation have led to the development of a dense network of small public primary schools in rural areas. In 2012, Sindh had 42,114 officially established government primary schools, approximately 1.08 schools per 1,000 inhabitants (Asim, 2013). In **Figure 1**, we illustrate some typical characteristics of a sample village, Seri, in rural Sindh. Seri has households from 11 different castes living in 12 settlements. A total of five public primary schools serve these settlements. The presence of multiple schools within a single village reflects, in part, the extent of social fragmentation: most settlements contain households of one or two castes. As a result, the social barriers against sending children, particularly girls, to a school outside of their home settlement are high (Jacoby and Mansuri, 2011). Public primary schools typically feed 1-2 settlements, with a larger school often located in the main settlement, the traditional hub of the village.

[Figure 1 about here]

Many public primary schools are not functioning or are temporarily closed. In 2012, 14 percent of Sindh's public primary schools were found closed for more than six months or had no teachers or students registered (Asim, 2013). Of the remaining schools, 54 percent had only one teacher, leading to multiple-grade classrooms and risk of temporary closure. Of the schools which are functional, lack important resources and basic facilities. Seventeen percent of schools in 2012 had no building, with classes taking place outdoors in a hut or a temporary structure. Only 58 percent had a toilet, 51 percent had drinking water and 40 percent had electricity (ibid).

These failures of service delivery reflect how schools function as a mechanism of patronage for Sindh's feudal elite. In many cases schools serve more as an extension to a landlord's estate in a village than as a functioning public service (Gazdar, 2000). School buildings are frequently used for non-educational purposes, and politically-appointed teachers are protected from discipline and rarely attend schools (Babur, 2016). Since many schools have only one teacher, this results in a significant proportion of schools being non-functional at any given time, owing to lack of staff and basic facilities.[3]

As a result of these governance failures, enrollment remains low in Sindh. According to the PSLM survey, 2010-2011, the Net Enrollment Rate (NER) at primary level (ages 6 to 10) in rural Sindh was 54 percent, compared to 67 percent in rural Punjab. Data from our baseline survey, conducted in three representative districts of Sindh, reveals a similar picture: fewer than half (49.8 percent) of school-age children in surveyed households were enrolled in school, and the majority of those unenrolled had never attended school **[Table 1]**. Although low enrollment often reflects limitations

[3] During the period of this study, the inadequacy of schooling facilities was exacerbated by extensive flooding in both 2010 and 2011. The 2010 floods destroyed an estimated nine percent of Sindh's educational facilities and damaged an additional 19 percent and severely disrupting the livelihoods of about 865,000 Sindh households (Asian Development Bank and World Bank, 2011). The 2011 heavy rains resulted in the extent of damage similar to that of the 2010 floods (Asian Development Bank and World Bank, 2012). Schools which were not damaged were converted into temporary shelters for displaced people, disrupting teaching for several weeks. Overall, 66 percent of schools in the villages sampled for this study reported being affected by the floods at the time of the baseline (EMIS, 2011/2012).

of demand as well as supply, recent evidence suggests that the latent demand for education in Sindh is high; Barrera-Osorio et al. (2017), in an experimental evaluation, found that provision of schooling through public-private partnership to under-served communities in rural Sindh increased enrollment by 30 percentage points for treatment villages.

Table 1: Enrollment by Grade, Children ages (6-10 years old), 2012

	Public	Private	Other	total	Percent
Population 6-10 years				7416	100%
Never attended				3550	47.9%
Attended & unenrolled				171	2.3%
Enrolled	3360	249	86	3695	49.8%
Currently enrolled:					
*Katchi** & Grade1	894	56	21	971	13.1%
Grade2	1079	63	29	1171	15.8%
Grade3	630	55	22	707	9.5%
Grade4	365	34	9	408	5.5%
Grade5	231	22	3	256	3.5%
*Katchi** – Grade 5	3199	230	84	3513	47.4%
Grade 6 and above	161	19	2	182	2.5%

Source: Data from Sindh Baseline Household Survey, 2012.
Note: Katchi is the pre-primary grade in which children are admitted in a public primary school.

Learning outcomes are still poor for enrolled children. At Grade 3, only 24 percent of students can read words while only 32 percent can perform basic subtraction. At Grade 5, only 19 percent of students can read full sentences while 35 percent can perform division (ASER, 2015). Gender disparities in learning are large and similar for both English and Math; boys outperform girls by 6 percentage points (ibid). In this study's baseline survey, in learning assessments designed to capture students' achievement at Grade 1-5 levels, tested children (7-13 years) scored an average of only 27 percent across subjects.

Sindh introduced community management policies for schools in the 2000s as part of a broader agenda of decentralization and community participation in public services. SMCs were established at each school, including the head teacher, as well as four elected community members, two of which are parents and two non-parents (village elders). However, SMCs had little formal control over school decision-making or expenditure and in practice remained largely dormant. In 2006, the Government of Sindh attempted to revitalize SMCs as a formal channel for local communities to engage with schools. Following the reforms, SMCs received annual grants from the Government worth PKR 22,000 (USD 220 equivalent in 2012) to partake in school improvement activities. The SMC was given control over its allocated funds and was empowered to withdraw these resources to implement activities in accordance with a School Improvement Plan (SIP), developed in consultation with the community. The introduction of the annual grants was intended to draw parents and community members to encourage participation in school improvement activities. However, concerns persisted over the transparency and utilization of the grant funds.

In 2009, as part of further reforms, the Government distributed updated SMC guidelines to all public primary schools to avoid any fiduciary and audit concerns with the direct transfer of SIP Grants to schools. These fiduciary controls were meant to assuage concerns regarding transfer of

funds to SMCs but had the unintended effect of weakening their role. The main purpose of school committees was reduced to that of drafting a SIP, maintaining auditable records for procurement of goods, and implementing nominal repairs and maintenance. They were also made responsible for monitoring enrollments and ensuring that children in their community started school at the appropriate age. However, they retained no wider role of engaging the community to improve quality of teaching and learning in schools.

The reforms were successful in reinstating the transfer of SMC funds annually; as of June 2012, over 81 percent of SMC bank accounts were active and received grants from the Government (World Bank, 2012). Evidence from the baseline survey (2012) suggests that the activity level of SMC improved as a result of the reforms, but remained low overall: while 91 percent of schools had an SMC, fewer than half of these (45 percent) had held a meeting in the previous six to twelve months, and fewer than a quarter (21 percent) had held a village-wide meeting during the same time. Only 29 percent of head teachers reported that SMC members at their school were chosen by election. However, underutilization of funds continued to persist. In two of the three districts in which this study takes place,[4] the schools that fell in the 90th percentile in terms of unutilized SMC funds had PKR 77,500 and PKR 62,000, respectively, each equivalent to at least three years of accumulated funds.[5]

Sindh is a suitable context for exploring alternative approaches to community engagement. Private schools are virtually non-existent in Sindh, making it harder, if not impossible, for parents to withdraw their children from public schools and send them to private schools. Ethnic and caste fragmentation means that, despite the large number of public schools, competition between schools is also not a viable source of pressure for improvement. Hence, there is no market-based *exit* option available to parents (Hirschman, 1970). In the absence of accountability through school choice, and in an environment marked by low democratic pressure on service providers, the only option for dissatisfied parents is to directly *voice* concerns with the educators (ibid).

2.2 Sample

Sampling for the study was done in accordance to the districts: (1) being representative of the overall education profile of rural Sindh; (2) having a sufficient number of schools to meet the statistical requirements of our analysis; and (3) not posing excessive security risk for field teams. Using the PSLM survey, we ranked the 28 districts of Sindh by: i) proportion of adults who ever attended school, and ii) net enrolment rates of primary-age children (5-12 years). Also, using ASC (2010/11), we ranked districts by size as measured by the number of schools and villages in each district. In consultation with the Government of Sindh, we then chose one large, one medium-sized

[4] The districts in question were Matiari and Mirpurkhas. See section 2.2, Sample, for details of selected study districts.
[5] We report cost estimates from School Improvement Plans prepared by SMC members during capacity building under ELECTIONS; participants were instructed to prepare SIPs to spend all accumulated reserves (Figure A3). In 2013, as the result of initial findings from the baseline survey and pilot activites, the Government of Sindh revamped the scope of the use of SIP Grants, specifically requiring all SMCs in the province to submit Bank statements and members' composition details to be eligible to receive annual funds. The idea was to stop the accumulation of SIP funds in delinquent accounts and only provide grants to schools with active SMC members and functional accounts. This accumulation of funds at the time of intervention might have contributed to expenditures on classrooms, which cannot be constructed alone from the annual SIP Grants.

and one small district, ensuring representativeness of the selected districts along two key education measures.[6] The selected districts were Matiari, Mirpurkhas and Sanghar.

In the absence of an updated population census, administrative data— ASC (2010/11)—was used to set the population frame for selecting villages. Given the heterogeneity in village size, we employed probability-proportional-to-size (PPS) sampling, creating selection probabilities proportional to the number of students enrolled in primary education in a village.[7] Although PPS makes it more likely to pick up schools in villages with more enrolled children, we used this method to increase the probability of selecting a school with more reliable administrative data for our census exercise. Using PPS, we selected a sample of 550 schools drawn from 377 villages. The administrative data appeared to be unreliable regarding the exact location of schools and the inclusion of schools that are permanently closed. To address this, we mapped 300 of the 377 villages in a village-census exercise. The maximum of 300 villages was reached using a randomized ranking procedure.[8] The process through which treatment schools were selected from the village mapping exercise is outlined in **Figure 2.**

[Figure 2 about here]

The school sampling strategy for the baseline covered public primary schools that were open on the day of the visit or closed for a period of less than one year prior to the day of visit. We sampled all such schools, identified by local community, in the main settlement of our sampled villages; and an additional 15 percent of schools, in other settlements within each village. In villages that did not have any school in the main settlement, a maximum of three schools were sampled from other settlements based on their total enrollment.[9] Our total sampling consisted of 501 schools across 296 villages. This is representative of both the three study districts, and of rural Sindh, in general, across a range of key indicators [Table A2, in Appendix A]. However, the sample differs significantly from study districts in the number of female teachers per school, and from all of rural Sindh in school enrollment and availability of drinking water; these indicators are linked to heterogeneity in population, as well as location-specific endowments.[10]

Following data collection and analysis, the final sample of the study is 284 villages containing 479 schools.[11] Of these, 387 schools in 249 villages were consistently open at baseline and endline.

---

[6] Matiari was ranked the third smallest district, Mirpurkhas was ranked 12th, and Sanghar was ranked 18th, according to the number of schools and villages in each district. In terms of education indicators, Mirpurkhas had one of the lowest levels of education outcomes, Matiari was in the middle, while Sanghar was among the highest. They were deemed relatively safe by the Government for field teams to visit.

[7] Virtually every village in Sindh has a Government public school. Small schools with no teachers and few students, regarded as ghost schools, generally have problematic location data.

[8] In order to select 300 villages from 377 (per our evaluation design), a village mapping firm identified and listed all schools in each village in the rank-order specified, until they hit the maximum sample of 300 villages. They were allowed to skip villages only when i) schools were not found in listed revenue villages due to noisy administrative data; or ii) field-teams were denied permission by local resident to enter the village. However, the number of such cases was small, with total replacements amounting to less than 20 villages.

[9] For villages that did not have either a functional or temporarily closed school in any of its settlements, we included all schools, 16 in total, even if they were closed for a period of more than one year.

[10] See Table A2, in Appendix A.

[11] Field teams were not allowed to enter four villages at the time of baseline. In addition, eight villages had to be dropped at data analysis stage for incomplete and inconsistencies in collected data.

Therefore, measurement and analysis of outcomes of interest was restricted to this sample, except for the school status on an unannounced visit. Schools that switched status from closed to open at endline are excluded, as the characteristics of these schools are largely different from schools that were consistently open both at baseline and endline [Table A3, in Appendix A].

2.3 Treatments

Our selected treatments were designed to assess the relative effectiveness of three different approaches to strengthen community engagement in the context of low empowerment and high barriers to community participation. The first two interventions, MEET and SMS, test alternative platforms for facilitation of community dialogue around education: MEET introducing external facilitation through face-to-face meetings, and SMS introducing a novel ICT-based approach for an ongoing dialogue. The third intervention, ELECTIONS-CAPACITY, targets the existing SMC institution, and tests the additional impact of holding of elections for SMC members, and raising the capacity of SMCs for collective action by providing participatory training for members.

The evaluation follows a clustered randomized control trial (RCT) for causal identification of the impact of three interventions on school-level outcomes. Villages are randomly assigned to one of the treatment groups or the control. The unit of randomization is the village and the school is the unit of inference.

In a cross-over design, the third intervention is combined with each of the dialogue interventions in a randomly selected subset of schools, creating four distinct treatment combinations. **Table 2** summarizes the structure of the interventions and treatment groups.

Table 2: Cross-over RCT design

		Dimension 1: Community mobilization and facilitation of dialogue		
		No facilitation of dialogue	Meeting-based community dialogue (MEET)	SMS-based community dialogue (SMS)
Dimension 2: Elections and capacity building of SMC Members	No SMC elections and capacity building	Control	MEET	SMS
	SMC elections and capacity building (ELECTIONS-CAPACITY)		MEET-E	SMS-E

All interventions employed a similar structure and institutional arrangements. A local community mobilization firm visited villages and announced a community-wide meeting via loudspeakers, posters and conversations with residents[12]. Various dissemination practices were pre-tested in non-sample villages to identify the most appropriate and effective ways to mobilize the community for village-wide meeting. At the meeting, participants completed an attendance sheet and provided a mobile phone number. They then watched a performance by local schoolchildren and listened to a specially-prepared 10-minute audio clip presenting a dramatized story which promoted the

[12] There was no placebo meeting or sensitization efforts done in control communities. All treatment groups' findings therefore include the impact of these common sensitization activities, which were not carried out in control villages.

importance of education and provided basic information on SMCs, their structure, purpose, and membership, as well as the amount of funds available for their use. The clip and delivery instructions were developed by an international communication firm, in consultation with stakeholders, to maximize effectiveness.[13]

The content of the remainder of the meeting, and any follow-up activities, varied according to the intervention.

MEET

In the first intervention, MEET, we test the impact of dialogue around education and schools in an externally facilitated village-wide meeting. Following the introductory activities described above, meeting attendants were introduced to another audio clip, a 20-minute dramatized discussion highlighting specific actions community members could take to improve education outcomes in schools. The audio clip emphasized the need for parents and the broader community to take an active interest in the schooling of their children by participating in further village-wide education meetings and visiting the school regularly to check the presence and activity of teachers, monitor the use of SIP Grants, and note any need for repairs and maintenance. A moderated discussion was then held amongst meeting participants around key ideas communicated in the audio script. At the end of the discussion, existing SMC members were introduced to the villagers and their names and phone numbers given to participants. Participants were encouraged to independently organize and conduct a follow-up meeting.

SMS

In the second intervention, SMS, we assess the impact of an alternative ICT-based approach to facilitation of dialogue, employing a specially-designed text message-based Community Dialogue Platform (CDP). Following the standard introduction as described above, facilitators played a customized audio clip which introduced the CDP, its purpose and core functions, followed by a hands-on demonstration by field facilitators.

The CDP was designed to offer a similar combination of information and facilitated discussion as MEET, in virtual form. First, key information shared in the MEET intervention, regarding rights and responsibilities and the functions of the SMC and SIPs, was sent through SMS messages to the community members [see Appendix A, Figure A2, for sample messages]. Second, participants were provided with a free number to which to send comments and responses to questions regarding conditions at local school. Users could send messages consisting of any comment or text; these were then categorized according to topic based on key words used. Participants could send any number of messages. Through this, the CDP elicited preferences and concerns about education from community members.

[13] Based on a feedback cycle, the audio clips were revised multiple times, from the original, poem-based format, and finally to an audio drama. An international marketing and communication firm, JWT Worldwide, was engaged to develop the story plot in drama format. The design team also sought to make the language of the audio clips easy to understand, and to adapt the clip to the context of rural Sindh. Facilitators were provided with timed scripts to ensure consistent implementation of the meeting.

A key aspect of the CDP was to combine this *elicitation* of preferences with *diffusion* of opinions to all registered participants on the system. Over the five months following the meeting, participants received messages on a weekly basis summarizing the topics and content of the messages received by the system. This was intended to enable users, even those who did not send messages to the platform, to be kept informed about dialogue going on through the platform about local schools and of each other's main concerns. It was envisioned that salient issues would then be further discussed by community members in informal gatherings.

The selection of an SMS-based platform for the intervention was informed by the high rate of mobile phone penetration in Pakistan – 72 percent in mid-2013 (Pakistan Telecommunication Authority, 2014) – which provided a low-cost and accessible way to engage communities. The design of the system was intended to reduce barriers to participation: messages were anonymous, in order to minimize any personal risk or consequences to participation for users, and airtime was provided to participants to defray the cost of sending messages. It was anticipated that these low barriers to participation would mean that activity on the CDP would be sustained throughout the five-month period of operation[14].

A potential threat to the effectiveness of an SMS-based approach was the low level of literacy in Sindh: only 60 percent of males above the age of 10 in rural Sindh, and only 22 percent of females, were literate prior to the start of our interventions (PSLM 2010-11). This was addressed through two methods. First, volunteers, selected by communities, were appointed to support mobile-illiterate members in contributing to the system. Volunteers could send messages on behalf of users with literacy or other barriers and receive airtime credit as compensation. About five percent of messages sent to the CDP were sent by these volunteers on behalf of other users, in addition to other users who sent their own messages with the support of volunteers. Second, in order to ensure that all users could participate in the CDP in their own language, messages were made available in both Sindhi and Urdu. An automated voice call was used to inform users that they had a choice of language and to collect their preference.

ELECTIONS-CAPACITY

The third intervention, ELECTIONS-CAPACITY, targeted the SMC institution. Many of Sindh's SMCs were not fully and regularly elected, meaning that even if the MEET and SMS interventions increased the level of community dialogue, the SMC could not be expected to respond to the concerns and preferences expressed. At the same time, the low level of utilization of SIP Grants, under the direct purview of SMCs, suggested a lack of enthusiasm or capacity among members to fulfill their roles. Therefore, the ELECTIONS-CAPACITY intervention included two aspects: transparent elections for SMCs, and participatory capacity building for SMC members.

The elections took place at the end of the introductory village meeting. Taluka Education Officers (TEOs), government sub-district education officials, were required to attend meetings and provided with an honorarium to officiate elections. It was intended that TEOs would afford official legitimacy to the election and provide letters to enable the newly elected SMCs to assume official

[14] The budgets for the pilot intervention were made available through Education Sector Development Trust Fund, and the allocated funds were available only for FY-13. The intervention started in January with completion date of June 30, 2013.

authority and claim SIP Grant funds. We find no evidence of any interference from the TEO in the proceedings of elections or the transfer of charge to newly elected members.

Following the elections, the SMC members were provided with hands-on training during three structured meetings conducted over a three-week period. The meetings familiarized members with the functions of the SMC and provided participatory training on how to develop an SIP, with members drafting one during the meeting, and provided guidance on how to present the plan at the village-wide meeting for ratification. In addition, members were trained on how to withdraw and manage grant funds and oversee implementation.

2.4 Theory of change and anticipated impacts

It was anticipated that the MEET and SMS interventions could impact schools through several channels. First, it was expected that the information provided on the importance of education through the MEET and SMS interventions would lead to increased salience of education within communities, and demand for schooling by parents. Second, the information provided, within both interventions, on the community's role in school management was anticipated to improve the level of effort by communities to monitor the use of resources within schools.

Third, it was anticipated that the elicitation and diffusion of local knowledge and preferences carried out in both MEET and, in particular, SMS could lead to increased social pressure by communities on local elites to improve the quality of local schools. We postulate that a platform that encourages a multiplicity of opinions through anonymous exchanges, and diffusion of preferences among all members, vitiates elite control over framing of public opinion on quality of service delivery in villages. We anticipated elites would respond to community-identified needs, favorably, mobilizing more resources to school in a bid to secure their status as the ones perceived as working for the benefit of the community.

It was expected that the ELECTIONS-CAPACITY intervention, which focused on SMCs, would improve the effort of SMC members and their focus on the priorities of the wider community, and improve the utilization of SIP Grants.

We identify five key anticipated impacts on schools from the interventions: school functioning; availability of teachers and teacher attendance; school infrastructure; enrollment; and learning outcomes.[15]

School functioning: We anticipate impacts from community engagement on school functioning as a response by local elites and school administrators to the increased salience of and demand for education. Such a response could be the result of the removal of barriers, e.g. the misappropriation and misuse of school buildings (Gazdar, 2000); in the case of Sindh, where lack of teachers is a key contributing factor in school closures, it could also reflect increased supply of teachers, specifically a reduction in the number of schools with only one teacher.

[15] The pre-analysis plan was not registered in a secure independent register in 2011 when the experiment was designed, as is the best practice today.

Enrollment: We anticipate impacts on school enrollment because of increased parental interest in education, and in the case of ELECTIONS-CAPACITY, from improvement in SMC efforts to enroll out-of-school children. We also anticipate greater retention of students because of improvements in school quality as perceived by parents.

Availability of teachers and *teacher attendance:* As described above, we anticipate improved provision of teachers to schools as a response by elites and authorities to deliberation on education issues within communities. In addition, we also anticipate improvements in teacher attendance because of enhanced monitoring of school resource use by communities.

Infrastructure: We anticipate possible improvements in school infrastructure as a result of increased effort by SMC members. Although the annual figure of US$ 220 equivalent (2012) for SIP Grants is insufficient for large infrastructure investments, in schools with accumulated carry-over funds it was expected that increased community engagement, coupled with the election of new committee members, would result in substantial spending to address infrastructure gaps, particularly shortages of toilets and classrooms in schools. Furthermore, improvements in simple infrastructure such as toilets could reflect increased contributions of labor by both SMC members and other community members.

Learning outcomes: We anticipate potential improvements in learning outcomes primarily as a result of improvement in teacher availability and reductions in teacher absenteeism. However, given the expectation of improved enrollment, we also anticipate a potential reduction in average learning outcomes because of compositional change in the student body (Crawfurd, 2018).

[Figure 3 about here]

2.5 Randomization

Table 3 shows balance checks between treatment groups on key indicators at baseline. The treatment and control groups are balanced in terms of baseline statistics along a range of covariates and outcome variables. None of the treatment groups differs significantly from the control group in any indicator. We run a regression of each treatment dummy on all variables and test the joint significance of all coefficients using an F-test. We fail to reject the null and conclude that there is no difference between any groups.

[Table 3 about here]

2.6 Timeline

Figure 4 summarizes the experiment sampling frame and timeline.

[Figure 4 about here]

The village census and school mapping exercise, and sampling, were carried out in 2011. Baseline data collection was carried out between April 2012 and January 2013. The introductory village meetings for all four treatment groups were completed within the first quarter of 2013. For the

groups receiving the SMS intervention, the CDP was then operational until June 2013. For groups also receiving the ELECTIONS-CAPACITY intervention, capacity building took place in the month following the village meeting. Endline survey activities were conducted in schools from January 2015 to March 2015. Endline data collection at household level was conducted between January and June 2017.

2.7 Treatment Fidelity

Table 4 illustrates fidelity to treatment for the various interventions. We first measure attendance at the community-wide meetings, which provided the forum for facilitated discussion under MEET, introduction of the CDP under SMS, and the electing of SMC members under ELECTIONS-CAPACITY.[16] The population was estimated from the household census conducted for the listing exercise. The take-up rate is measured as the number of attendees as a proportion of the number of adult individuals residing in the main settlement of each village.[17] Overall, on average the participation rate was 63 percent among treatment villages. Mirpurkhas had the highest participation rate with 75 percent of households, followed by 57 percent in Sanghar and 51 percent in Matiari. These district-level differences in participation rates are controlled in the proceeding estimations using district fixed effects.

[Table 4 about here]

The MEET and SMS groups have higher participation rates, on average 67 percent and 68 percent respectively, compared to 58 percent for MEET-E and SMS-E. This suggests that some aspect of the ELECTIONS-CAPACITY intervention actually reduced attendance at village meetings. The differences in attendance are particularly stark for Matiari district, where average participation rates were only 38 percent in groups receiving ELECTIONS-CAPACITY, and 63 percent for MEET and SMS treatments alone[18].

For the SMS intervention, the other key aspects of treatment fidelity are registration and participation on the CDP platform. Across SMS and SMS-E villages, 63 percent of village meeting attendees registered with a valid cellphone for the CDP. The average registration rate in non-ELECTIONS-CAPACITY villages (65 percent) was higher than that in ELECTIONS-CAPACITY villages (61 percent). A total of 4,981 unique users registered on the CDP portal. All registered users received weekly messages summarizing the opinions expressed by participants anonymously on the platform. Therefore, dialogue via text message encompasses not only active messages sent by informed community members, but also passive receipt of summary messages by all registered users.

[16] The project team were instructed to revisit any village for which attendance was under 20 percent and hold a second meeting. Reported attendance is for second meeting in these cases.
[17] As discussed in Section 2.2, sample includes all main settlement schools and 15 percent of schools outside main settlements of sampled villages.
[18] Matiari was carved out of a larger district, Hyderabad, in 2005. The expectations were high that the creation of small district with dedicated line authorities will improve the functioning of services. The failure to meet these expectations might be another reason for evidently low confidence afforded to the civil administration officers in this particular district.

A smaller proportion of users provided inputs into the system: 28 percent of registered users, 1,229 in total, sent at least one non-junk message into the portal during the course of the project.[19] Input into the portal reduced in frequency over the course of the operation of the CDP, but activity continued throughout the implementation phase: an average of 16 messages per village were sent in the ten days following the introductory meeting; four messages per village were sent during days 31-40 of operation; and two per village during the last ten days of implementation five months later [Figure A4, in Appendix A]. This decline in sending of messages was expected as, once the set of problems in a particular village was identified, e.g., shortages of teachers and classrooms, and sent to all users in a weekly summary, it was unlikely that other users would repeatedly voice the same concern(s). Rather, the expectation was that these issues would be subject to continued discussion informally within the village.

For the first key aspect of the ELECTIONS-CAPACITY intervention, the holding of elections for SMC members, participation was almost universal: elections took place in all MEET-E villages but were not held in two SMS-E villages, one in Mirpurkhas and the other in Sanghar, owing to the non-availability of the Taluka Education Officer. Given that elections were expected to generate an influx of significant numbers of new members to SMCs, we also measure the extent to which the membership of SMCs changed following elections. SMCs have four elected members, and all four positions were up for election under the intervention. The findings suggest that composition change was large, if not total: the average change was 73 percent, equivalent to almost three new members; all four elected members were replaced in slightly more than one-third of schools (36 percent). The extent of composition change varied slightly between the two ELECTIONS-CAPACITY treatment groups, from an average 74 percent in MEET-E villages to 71 percent in SMS-E villages.

We also measure the attendance rate at the capacity building for SMC members, the second key aspect of the ELECTIONS-CAPACITY intervention. The mean attendance for the three meetings, in both MEET-E and SMS-E schools, was 96 percent, and members from all schools in ELECTIONS-CAPACITY groups participated.

Overall, fidelity to treatment was sufficiently consistent across treatment groups to enable accurate assessment of impact of interventions on key outcomes. The variance in take-up between districts does pose a potential threat to internal validity; this is addressed by the inclusion of district fixed effects.

3. Data and Estimation Strategy

3.1 Data

Data for this study comes from a village-level household census, and school and household surveys conducted both at baseline and endline. First, a census of households and schools was conducted in three study districts in 2011, which provided a population frame for random selection of villages and schools. Second, a baseline survey was conducted in schools and sampled households between

[19] Blank, incomplete, or off-topic messages (i.e. not concerning education) were labelled as junk.

April 2012 and January 2013.[20] Third, an endline survey was conducted in all schools from January 2015 to March 2015,[21] and in households between January and June 2017.[22] [23]

In both the baseline and endline, the school surveys collected detailed data on school-level variables such as pupil enrollment and attendance, school infrastructure, facilities and resources, the school SIP, and SMC functioning and membership. The questionnaire also collected data on the total number of teachers employed in school and a count of those present in school on the day of visit. The visits were held on a day when schools were supposed to be open but were not announced in advance. Information on enrollments, teacher and student presence, and SMC members was collected from school records in interviews with the head teachers. In addition, for the endline round of school surveys, a random sample of 20 students from Standard 3 and 4 from each school were interviewed and given standardized tests for Math, English and Sindhi.

In addition, household surveys for both baseline and endline gathered information on household demographic and socioeconomic characteristics, household schooling choices and engagement with SMC members from the male and female head of the household. A student questionnaire was administered to sampled children regardless of school-going status. Standardized tests were administered at this time to primary school-age children (ages 7-13), for Math, English and Sindhi.[24]

3.2 Outcome Variables

As described in section 2.2, the pilot interventions could potentially impact a number of outcomes: school functioning; student enrollment; availability of teachers; teacher attendance; school infrastructure; and learning outcomes. In order to ensure accuracy and adherence to the study timeline, we measure all variables through directly observed survey data.

For school functioning, we measure whether schools were open at the unannounced visit at baseline and endline.[25] Since a number of schools switched status between the visits, it is important

[20] The initial plan was to collect all baseline data in 2012. However, due to heavy flooding in Sindh province the data was collected in two rounds from April to June 2012, and November-December 2012. The flooding delayed implementation of the pilot activities but did not significantly disrupt data collection beyond this delay. We control for a flood dummy in all regression estimates.

[21] The endline survey exercise included two phases: surprise visits to schools in January 2015, and announced visits made to schools in the sampled villages between January 2015 and March 2015.

[22] Following analysis of the school endline survey data, the endline household survey was conducted for a sub-sample of 160 randomly selected baseline villages. Power calculations and budget limitations motivated the decision to reduce endline household sample. Initial visits were conducted at the household level in January 2017 to obtain tracking data on household members and children, with the full survey conducted from March to June 2017.

[23] The World Bank worked with local survey firms: Research Consultants (RCons) and Weitek Group on various rounds of data collection. The baseline and endline surveys were conducted by Weitek and RCons, respectively, to mitigate any risks associated with biased reporting in interviews by the same firm for both rounds.

[24] Standardized tests for English, Math and Sindhi were administered to children ages 7-13 at the household level both for the baseline and endline. The instruments were designed, pre-tested and piloted based on students' textbooks and curriculum standards for Standard 1-5 of primary school. We fitted a two-parameter item response theory (2pl-IRT) model for English and Math scores to account for what students of similar knowledge and skills ought to know by distinguishing between difficulty level (see Appendix B for details). Sindhi tests include open-ended question items and hence an IRT score cannot be fitted.

[25] If a school was found closed verification was sought from community members. If a school was closed for one or

to capture the direction of change from open to closed and vice versa, in relation to the treatments. This generates three specific outcome indicators: number of schools open; number of schools open at baseline which were closed at endline; and number of schools closed at baseline which had reopened at endline.

For enrollment, we first measure the number of pupils enrolled in *katchi* – pre-primary – and Grade 1, according to school enrollment records at both baseline and endline. To measure the retention of students into upper grades, we also measure the number of pupils enrolled in the ensuing grades 2-5, using the same records.

For school infrastructure, we measure change in the number of classrooms, and whether a school has any of four key facilities that are important for a safe, secure and comfortable school environment: toilets, electricity, drinking water, and a boundary wall. Electricity, sanitation and water are particularly important given the prevalence of high temperatures in Pakistan, where temperatures during summer term periods can often reach upwards of 40°C, while a wall is important for student safety and to support student attendance (Andrabi et al., 2007).

For availability of teachers, we first measure the average number of teachers registered at schools at endline. In addition, in recognition of the fact that the large number of one-teacher schools appears to be a central driver of school closures in Sindh, we also present the share of schools which have only one registered teacher.

For teacher attendance, we measure the number of teachers present at schools during unannounced visits at baseline and endline[26]; as well as the share of teachers absent during these visits, in comparison to the number registered at the school.

For learning outcomes, we measure IRT scores in Mathematics and English, based on the standardized tests described above[27]. We report findings from tests administered to children tracked between baseline and endline at household level; as a robustness check, we also report cohort level results for standard 3 and 4 students tested at the school level.

3.3 Mediating variables

Very few studies, if any, have tested any structural model to identify the specific channel through which institutions supporting community engagement impact education outcomes. Nonetheless, we report some additional statistics on mediating variables measured robustly from the detailed household and school level surveys to describe possible ways the interventions might have impacted outcomes.

---

two days only, a revisit was made when the school was reopened and the school was not recorded as having been closed. If a school was confirmed by community members to have been closed for some time, and/or was found closed on both of two revisits, it was recorded as closed.

[26] Teacher presence and absence was measured at both surprise and announced visit; data from the unannounced visit was used except in 29 cases where the school was closed at unannounced visit but open at announced visit. The results are robust if restricted to the unannounced sample only.

[27] Sindhi test has a large number of open-ended questions. Only multiple-choice items were used to fit 2pl-IRT model. Results of the multiple-choice section of Sindhi were largely consistent with English and Math tests and are omitted for brevity.

To assess monitoring of school resources by the community, we analyze data on the frequency with which household heads report having visited schools in academic year 2015/16 and 2016/17. A greater number of visits to the school or to see the Head Teacher suggests a greater interest in education, as well as more engagement in the way resources are utilized at the school level. We note that the quantity of visits to school may not fully capture the quality or relevance to school issues of interaction between teachers, head teachers and parents at the school level. We also present descriptive statistics from the endline household survey on the key priorities which household heads discuss with head teachers when they visit school.

To measure the engagement of communities with SMCs, we employ data from household surveys at baseline and endline testing the awareness of household heads of the existence of the SMC, its membership, the frequency of its meetings, the amount of school grants it controls, and the requirement to prepare a SIP.

To measure improvement in alignment of SMC activity with communities' priorities, we analyze the main items of expenditure in SIPs developed during the capacity building exercise provided to SMCs as part of ELECTIONS-CAPACITY, in comparison with the main topics of discussion in SMS messages received by the CDP.

3.4 Descriptive Statistics

Table 5 shows the observed values at endline in the control group villages and the various treatment groups, for a range of indicators.

[Table 5 about here]

The findings suggest that, absent the intervention, the problems of the school system in Sindh have not considerably improved. In terms of school functioning, only 80 percent of schools in the control group were found to be open at the unannounced visit at endline. In terms of enrollment, the average school had 71 students, including all grades from *katchi* (pre-primary) to Grade 5, suggesting the continuation of Sindh's general pattern of multiple small schools. Only 46 percent of students at control schools at endline were female, suggesting continuing gender disparities in enrollment.

Infrastructure remained an area of significant shortage in control schools. While 96 percent of schools in the control group had a building at endline, sixteen percent employed at least one open-air classroom. The average control school had 2.2 classrooms available for use at endline. Toilets appear to be an area of significant shortage: only 75 percent of control schools had toilet facilities available, with an average student-toilet ratio of 59:1. Only 51 percent of control schools had electricity, 60 percent drinking water, and 70 percent a boundary wall.

In terms of staffing and teacher attendance, the control endline data suggests no significant improvement absent the intervention. The average control school had 2.25 teachers[28]; an average

[28] The number of teachers in control schools has not decreased as a result of the interventions. We believe that any impact on new teachers in treatment schools is from assignment of teachers to these schools, and not resulting from other teachers taken away from nearby communities (controls) and transferred to the treatment schools.

five percent of teachers were absent even during announced visits. Fifty-two percent of control schools remained one-teacher schools at endline.

In terms of learning outcomes, test scores remained low. School-age children undergoing learning assessment at household level at endline in control villages scored, on average, 22 percent across subjects, compared to 36 percent for the cohort tested in schools.

3.5 Estimation Strategy

We estimate the school level effect of intervention (intent-to-treat or ITT effect) by fitting Analysis of Covariance Model (ANCOVA):

$$Y_{i,v,d,endline} = \beta T_v + f X_{i,v,d,baseline} + \ u_d + \varepsilon_{i,v,d} \qquad (1)$$

where $Y_{i,v,d,endline}$ is an outcome variable at endline for school *i* in village *v* and district *d,* and T_v is the vector for village-level treatment indicators. $X_{i,v,d,baseline}$ controls for the baseline value of outcome variable to improve the precision of the point estimate. In addition, the vector of controls include school size, and a flood dummy to account for any lagged effect of exogenous shock to schools before the intervention. The term u_d denotes district-level fixed effects which control for any district-level differences in participation rates. We adjust the standard error to account for within-village correlations across schools in outcomes.

The coefficient of interest is β, which indicates the magnitude of the effect of each treatment, MEET, SMS, MEET-E and SMS-E, with respect to the control group. We test the sensitivity of our estimates to inclusion of treatment interaction with participation rates at the village meeting.[29]

ANCOVA is the preferred estimator as standard difference-in-difference (DiD) estimator would require twice as much sample to yield the same power when baseline is taken, and autocorrelations are low (McKenzie, 2012)[30].

To measure the effect of treatment status on student test scores we fit the following ANCOVA model:

$$S_{k,m,t,v,endline} = \beta T_v + f X_{k,m,t,v,baseline} + \ u_d + \varepsilon_{i,v,d} \ \ldots (2)$$

where, $S_{k,i,endline}$ is the test score for student *k* in subject *m* in household *t* in village *v* at the follow-up. T_v is the indicator variable for treatment assignment. $X_{k,\mathrm{m},t,\mathrm{v},baseline}$ controls for the baseline value of same students' test score. The term u_d denotes district-level fixed effects. We adjust the standard errors to account for within-village correlations across students in outcomes. We test the sensitivity of our estimates to inclusion of treatment interaction with participation rates at the village meeting.

[29] We are unable to find any significant interaction effects with participation rates and any outcome of interest. Results are omitted for brevity, but available on request.
[30] The study is underpowered to detect small effect sizes for a standard DiD estimator.

3.6 Attrition

There is some attrition in each round of data collection. We used the sample of 287 villages and 489 schools for randomization. For a handful of villages, the data quality was poor, with incomplete surveys returned at the endline, leading to missing data for 1% of study villages. We report pre-intervention characteristics for 284 villages in **Table 3**. In addition, the survey teams could not collect data on 71 schools and 58 schools in the baseline and endline round of surveys; these schools were temporarily closed and despite repeated attempts data could not be collected from these schools. As shown in Appendix A, Table A4 there are no systematic differences in attrition by treatment group. The difference-in-differences among attritors and non-attritors, across schools in treatment and control groups, reported in the last column, suggests that there is no indicator with a statistically significant value for difference-in-differences. We also test for joint significance of individual coefficients being equal to zero using seeming unrelated regression (SUR), the resulting Chi square statistic suggests we should not reject the null that both samples are balanced[31]. While attrition is not systematically different across treatment and control groups, but the diminished sample size nonetheless may have reduced the precision of the estimates.

4. Results

4.1 Average ITT effects: Community mobilization and facilitation of dialogue

This section presents the impact of the first two interventions, MEET and SMS, both of which focused on facilitation of community dialogue and deliberation on school issues among community members. These results are shown in the first two estimations in each of **Tables 7-11.** In addition to presenting ITT estimates in comparison to control, we also carry out equivalence testing (F-stat) to directly compare the outcomes in MEET and SMS. Where relevant, we also report findings from simple comparison of means between treatment groups at endline to motivate the results **[Table 5]**. For certain outcomes of interest, we also conduct transition probability analysis, measuring the likelihood of a change in school-level outcomes **[Table 6]**.[32]

4.1.1 School functioning

On average, the MEET intervention increased the number of schools that were functional in the treatment group as compared to control **[Table 5]**. To understand the direction of change in school functioning between baseline and endline, we conduct transition probability analysis of the likelihood of the school changing status from closed to open between baseline and endline; to capture any potential reduction in school closures, we also capture the likelihood of schools changing from open to closed **[Table 6]**.

[31] We use similar test as applied for attrition by Barrera-Osorio et al, 2018.

[32] The analysis is directional, assigning a value of zero if no or negative change occurred in a given outcome at a school and one if positive changed occurred. The implicit assumption in transition analysis is that any negative impact on outcome variable is the result of factors unrelated to the intervention. The difference in transition percentages between each treatment group and control represents the additional likelihood of a change in outcomes in the stated direction.

[Table 6 about here]

The results demonstrate that schools in MEET were significantly less likely than control schools to close between baseline and endline. 4.3 percent of MEET schools closed, compared to 13.3 percent in the control group; this is equivalent to a 67 percent reduction in the likelihood of schools closing. However, we do not observe a significant improvement from MEET in the likelihood that schools transitioned from closed to open. The limitation of the analysis is that the population of schools that switched status is a small subset of intervention schools and we do not have the full set of covariates to test the balance between treatment and control groups for this subset.[33] The share of schools found closed on unannounced visit at baseline is balanced across treatment groups, giving a degree of confidence that changes in school status are largely driven by the intervention.

For SMS, the picture is more mixed. The intervention did not significantly reduce the likelihood of a school closing between baseline and endline; however, the intervention does appear to have led to reopening of schools which were closed at baseline. For these schools, the chance of transitioning from closed to open were more than double that for control schools: 22.3 percent versus 10.2 percent.[34]

Next, we present the regression estimations for school functioning **[Table 7]**.[35] The findings reinforce those of the comparison of means and the transition probability analysis. MEET schools were 11 percent more likely to be open at the unannounced visit at endline than control schools; they were also significantly (9 percent) less likely to have changed from open to closed. SMS schools were not more likely to be open at endline, but if closed at baseline, were significantly (13 percent) more likely to have reopened.

[Table 7 about here]

4.1.2 Enrollment

Next, we turn to the impact of the interventions on enrollment. **Table 8A** presents overall and gender-specific impacts on school-wide enrollment. The MEET intervention does not appear to have significantly improved enrollments for either male or female students; however, we do observe a large and significant impact on boys' enrollment from the SMS intervention, with schools having an average 10 more male students at endline than control group schools. This is a large increase, given the small school sizes common in Sindh, equivalent to additional enrollment of 21 percent. Equivalence testing confirms that the SMS effect on total boys' enrollment is significantly larger than that of MEET. No significant impact was observed from either intervention on girls' enrollment.

[33] Schools were found closed on the day of visit for either baseline or endline.

[34] We note methodological limitations of the finding as school status changes are limited to 102 schools in the study sample which were closed either at baseline or endline. We are cognizant of the fact that a small number of villages might be contributing to the observed impacts. Nonetheless, flipping of school status at a much higher rate than the controls is an important effect to be captured given the context of this study. Reduction in the share of one-teacher schools in treatment groups further give credence to the effects observed on school functioning indicator.

[35] All regression estimates control for district fixed effects, to rule out impacts from variation in participation rates between districts; as well as a flood dummy which controls for whether a school had been affected by the floods in the period before baseline, in order to rule out bias from flood-related impacts on outcome variables.

Table 8B presents enrollment impacts by grade, specifically for male students. No significant impact is observed from either MEET or SMS on *katchi* or Grade 1 enrollment, suggesting a lack of effect of the interventions on the likelihood of new school-age boys being admitted to school. However, significant impacts are observed from SMS on boys' enrollment in Grades 2, 3 and 4, suggesting improvements in retention of male students throughout the school cycle because of the intervention. SMS schools had an average 2.2 additional students per grade in Grades 2-4 than control schools, equivalent to an increase in enrollment in these grades of 29 percent. The SMS impacts were significantly larger than those of MEET in Grades 2 and 3.

Table 8C presents impacts on girls' enrollment. No significant impacts are observed in any grade.

[Tables 8A, 8B, and 8C about here]

4.1.3 Availability of teachers

Turning to teachers, comparison of means **[Table 5]** suggests significant impacts from MEET and SMS on the proportion of schools with only one teacher: one-third of MEET schools (32 percent) and SMS schools (35 percent) had only one teacher at endline, versus more than half (52 percent) of control schools. Transition probability analysis also suggests significant impacts from both MEET and SMS on the total number of teachers available at schools: MEET schools were almost twice as likely as control schools to gain at least one additional teacher; 29.7 percent of SMS schools gained at least one teacher **[Table 6]**.

A similar pattern is observed in the regression point estimates **[Table 9]**. MEET and SMS schools were both significantly less likely than control schools to be single-teacher schools at endline. MEET schools had 0.26 additional teachers at endline compared to control schools, although the result is not statistically significant[36]; SMS schools had an additional 0.3 teachers, significant and equivalent to a 13 percent improvement in staffing.

[Table 9 about here]

4.1.4 Teacher attendance

In order to capture intervention impacts on teacher absenteeism, we also present findings for the number of teachers present at a school on the day of unannounced visits at endline. We do not observe any impact from MEET on point estimates **[Table 9]**. Findings for the share of teachers absent confirm that the rate of teacher absence was also significantly higher in MEET schools, suggesting that the gains in staffing were offset by increased absenteeism.

[36] The study is sufficiently powered to identify policy-relevant effect size of 0.2 standard deviations or more. However, given a context of very small school sizes with limited number of teachers and very limited infrastructure, it is important that future studies sufficiently increase the number of observations to allow detection of small effect sizes, particularly for teachers and infrastructure variables.

Turning to SMS we do observe a substantial positive impact on teacher presence from the intervention, with treated schools having an average 0.24 additional teachers present at endline compared to control, equivalent to a 12 percent increase in teacher presence; however, the standard error is large, and the result is not significant. The rate of teacher absence was higher in SMS schools than in control, but the difference is statistically insignificant.

4.1.5 Infrastructure

Turning to physical infrastructure, **Table 6** also presents transition probability comparisons for the addition of classrooms and toilets. We do not observe significant impacts from either MEET or SMS on the likelihood of schools gaining additional facilities in either case. In the comparison of means **[Table 5]**, we include estimations for other key school facilities: electricity, drinking water, and boundary walls at schools. We find no significant impacts from either the MEET nor the SMS intervention on any infrastructure outcome.

In the point estimates **[Table 10]**, we include all infrastructure variables. We find no impact from MEET or SMS on any infrastructure variable which is statistically significant in comparison to control. In the case of SMS, a substantial impact is observed on the number of classrooms, with 0.23 additional classrooms at SMS schools; this is significantly greater than the level in MEET in equivalence testing.

[Table 10 about here]

4.1.6 Learning Outcomes

Finally, we present findings on the impact of the interventions on learning outcomes from standardized tests conducted during household surveys. Comparison of means **[Table 5]** shows no differences in test performance between any treatment group and the control for scores. **Table 11** shows regression point estimates at student-level.[37] We find no significant impact on learning outcomes for either the MEET or SMS interventions. The findings are subject to two potential limitations: the passage of time since the interventions, and the negative impact of out-of-school children. To address these, we also present estimates based on the same tests during school surveys, conducted with enrolled students only in 2015; we again observe no significant impacts from either treatment [Table A5, in Appendix A].

[Table 11 about here]

4.2 Average ITT effects: Support to elections and capacity of SMC members

The third and fourth estimations in each of **Tables 7-11** present the impact of the MEET and SMS interventions combined with the ELECTIONS-CAPACITY intervention.

[37] We also conduct estimations at village level and find no difference in results.

4.2.1 School functioning

We do not observe significant positive impacts on school functioning from the ELECTIONS-CAPACITY intervention. In fact, the addition of the intervention appears to have had a negative impact which offsets the impact of the MEET and SMS interventions. In the comparison of means **[Table 5]** and transition probability analysis **[Table 6]**, the impact of MEET in reducing the likelihood of school closures is smaller and no longer significant in MEET-E. Similarly, the impact of SMS in increasing the likelihood of schools reopening is reduced and no longer significant in SMS-E. The same pattern is observed in the point estimates **[Table 7]**, in which MEET-E and SMS-E schools are not significantly more likely than control schools to be open at endline, or to have reopened, or less likely to have closed, in contrast to MEET and SMS schools. However, the impacts of SMS-E on schools reopening, while not significantly different from control, are significantly larger than the impacts of MEET-E in equivalence testing.

4.2.2 Enrollment

A similar pattern is observed for measures of school-wide enrollment **[Table 8A]**. As with school functioning, the impacts observed from the SMS intervention on overall school enrollment and boys' enrollment are reduced and become marginally insignificant in SMS-E. However, turning to impacts on grade-wise enrollment **[Table 8B]**, SMS-E appears to have had significant impacts on enrollment of *katchi* and Grade 1 boys. The average SMS-E school had four more students in these grades than in the control group, an impact of 21 percent. No significant impacts are observed in girls' enrollment **[Table 8C]**.

4.2.3 Availability of teachers and teacher attendance

We find no significant positive impacts from ELECTIONS-CAPACITY on availability of teachers, and the intervention appears to have reduced the impact of MEET and SMS. In the comparison of means **[Table 5]**, the impact of MEET and SMS on the likelihood of a school having only one teacher is reduced and no longer significant in MEET-E and SMS-E. The same result is observed in the point estimates **[Table 9]**. In terms of overall number of teachers, in the transition analysis, the impacts of MEET and SMS on the likelihood of a school gaining additional teachers are reduced and no longer significant in both MEET-E and SMS-E. Similarly, in the point estimates, in contrast to SMS, no significant impact from SMS-E on teacher numbers is observed in the regression analysis.

As with MEET and SMS, no significant impact is observed from either MEET-E or SMS-E on the number of teachers present at school **[Table 9]**. The share of teachers absent was significantly higher in SMS-E.

4.2.4 Infrastructure

In the case of physical infrastructure, in the transition probability analysis **[Table 6]**, SMS-E schools were significantly more likely than control schools (and substantially more likely than SMS schools) to gain an additional classroom. More than one in four SMS-E schools – 27.7 percent – gained an additional classroom, versus 17.6 percent of SMS schools and 15.1 percent of control

schools. Furthermore, SMS-E had a significant and positive impact on the share of schools with toilets which was not observed for SMS alone. 21.9 percent of SMS-E schools gained an additional toilet, versus 16.4 percent of control schools and 16.2 percent of SMS schools. In the point estimates **[Table 10]**, as with SMS, the impact of SMS-E on classrooms was large but not statistically significant, but SMS-E was again found to have had a positive impact on the availability of toilets, increasing the share of schools with at least one toilet by ten percent. Similar to MEET, no significant impact is observed on any infrastructure indicator from MEET-E in any estimation.

4.2.5 Learning Outcomes

As in the case of MEET and SMS, we find no significant impacts on learning outcomes from either MEET-E or SMS-E in either comparison of means **[Table 5]** or point estimates **[Table 11]**.

4.3 Average ITT effects on household awareness of the SMC

To explain the extent to which the intervention succeeded in improving engagement between communities and SMC members, we measure household awareness of the SMC and its functions and processes. **Table 12** presents impacts from the various interventions on the awareness of household heads about the existence of the SMC; its membership; the frequency of its meetings; the amount of SIP Grants; or the requirement to complete an SIP. We find no significant impacts from any of the interventions on any awareness measure. The evidence suggests that the intervention did not raise community awareness of the existence, functioning, or purpose of SMCs. This finding is in line with evidence from India suggesting that the provision of information alone frequently fails to raise community awareness of the functioning of existing community engagement institutions (Banerjee et al., 2010).

[Table 12 about here]

4.4 Threats to internal validity

Using geospatial analysis (GIS), we find that five percent of intervention villages are located 1-2 kilometers from the nearest control village.[38] This proximity may endanger the validity of the impact estimates if contamination of treatment occurred in control schools and villages. Stable Unit Treatment Value Assumption (SUTVA) requires that the treatment of one unit should be unaffected by the assignment of treatment to other units (Wooldridge, 2012). The relatively close distances between control and treatment schools heighten the probability of contamination, a direct violation of the SUTVA assumption.

In order to test that contamination did not occur, we isolated all treatment schools in villages located at least 1 kilometer from control schools, and we conducted the analysis on this limited sample to identify any significant changes in the size of coefficients and standard errors. The results on additional teachers, teacher absenteeism, and primary school enrollment for boys all

[38] The mean distance between a treatment school and the nearest control school is 7.41km; the median 5.83km, and the standard deviation 12.67km.

continue to be statistically significant [Appendix A, Table A6]. In addition, the results remain consistent in magnitude compared to the main results tables, with negligible changes in effect sizes, suggesting that contamination has not biased the impact estimates reported within the full sample and fully functional sample of schools.

4.5 Limitations of Analysis

Using an ANCOVA estimation strategy assumes that the treatment and control villages exhibit statistical balance not only on baseline characteristics, but also in treatment take-up. We therefore conduct robustness checks in which we include variables that could potentially cause village-level heterogeneities in treatment take-up, including a village-level asset index (proportion of individuals who do not have any land), basic literacy scores for males and females, and cellphone penetration rate.[39]

Any technology- or information-based intervention such as SMS exhibits a literacy and gender bias: take-up is likely to be concentrated amongst educated males. If mothers are the family members most likely to influence girls' enrollment, his skewed take-up might partly affect why primary enrollment increased only for boys and not for girls. To correct for literacy bias, the treatment includes the installment of a village volunteer who works as a liaison to send text messages for the less-literate and technologically challenged villagers. As mentioned previously, we also find no change in our estimates when we control for literacy scores. Gender bias is not directly addressed by our interventions, however, and future studies can focus on treatments targeted at women to test whether women are able to affect policy making and public service delivery, and also to see how their social preferences may differ from those of men.

5. Discussion

5.1 School functioning and availability of teachers

Our findings suggest that facilitation of community-wide dialogue around education can succeed in catalyzing both demand for, and the supply of, schooling. Both MEET and SMS had significant impacts on school functioning. Specifically, while MEET had a predominantly defensive impact on school functioning by preventing closures, SMS was effective in causing closed schools to re-open.

MEET and SMS both significantly reduced the share of one-teacher schools; SMS also significantly improved the overall number of teachers available at schools. These results suggest that both SMS and MEET created conditions where local elites felt pressured to increase the supply of the most important resource, teachers, in schools. Shortage of teachers was one of the most frequently discussed topics on the CDP (see **Figure 5A**).

Given that household awareness of the functioning of the SMC did not increase, it is likely that SMC members were not the primary drivers of these results. We believe that general discussion

[39] Regression estimates omitted for brevity, but available on request.

facilitated by the interventions led to a more concerted effort by the community to work with village elders to lobby authorities for additional teachers.

The addition of ELECTIONS-CAPACITY, which sought to reform the SMC into an institution that improves school outcomes through better engagement with wider community, instead appears to have undermined the positive impacts observed for MEET and SMS on school functioning, student retention and staffing. The results suggest that the intervention undermined the effectiveness of community dialogue around education. A potential explanation is that the inclusion of a government official, the TEO, at the village-wide meeting where elections were held had the unintended effect of reducing community trust in the new SMC, and overall community engagement over time. We find that participation in the village-level meetings in SMS-E and MEET-E villages was around ten percentage points lower than in MEET and SMS villages. The participation of TEO was not announced during village mobilization activities prior to the meeting; however, residents walking to the meeting may have turned back once they saw the TEO approaching the venue. This reflects the lack of public trust in the government functionaries in Sindh, as well as popular perception of widespread failures of service delivery.

Another potential explanation is that the election of new SMC members compromised the ability of the SMC to mobilize resources for schools. The replacement of existing SMC members, often appointed by elites, with transparently elected members may have damaged networks and informal linkages effective in delivering some degree of inputs to schools; this is in line with findings from Indonesia that strengthening linkages between SMCs and powerful village councils had the largest impacts on learning outcomes (Pradhan et al., 2014).

The finding suggests that, in conditions of low trust in government, trade-offs may exist between the twin goals of strengthening community engagement and supporting elections of representative local bodies. This is in line with other studies in South Asia which suggest that government-led efforts to bolster community engagement are less effective than those implemented by external, independent actors (Pandey et al., 2009; Banerjee et al., 2010).

5.2 Enrollment

The ELECTIONS-CAPACITY intervention, in combination with SMS, succeeded in driving improvement in number of newly admitted children (*katchi* and Grade 1). The most likely mechanism of impact is an improvement in effort by the SMC, driven by new members, to identify out-of-school children and enroll them in schools. The SMS intervention alone did increase the number of new admissions, although the point estimate was insignificant; new members appear to have put some effort into activities required of them, complementing community's interest in education sparked by dialogue via SMS. To some extent the continuity of dialogue on the CDP may also have offset the loss of community interest resulting from the TEO's presence in the village-wide meeting. In MEET-E, however, owing to the one-off nature of the discussion, such complementarity effects are absent, and there was no possibility of regaining trust of the community once the village-wide meeting was over.

The SMS intervention also improved the retention of students, as measured through enrollment in Grades 2-4. This suggests a sustained improvement in communities' demand for education over

time which was not achieved by the one-off MEET intervention. The improvement in enrollment may also reflect improvement in the perceived quality of schooling in the eyes of parents, in response to the improvements in availability of teachers achieved by the SMS intervention. In SMS-E schools, where improvements in teacher availability were not significant, we do not observe significant gains in retention despite the improvements in *katchi* and Grade 1 enrollment.

The impacts of the SMS and SMS-E interventions on enrollment and retention did not extend over to female students. This may simply suggest that the interventions, although effective, were not strong enough to overcome cultural barriers to girls' enrollment. Alternatively, this may suggest that the ICT-based approach of the CDP did not overcome barriers to female participation. If women are more likely to respond to efforts to improve girls' enrollment but were limited in their ability to access CDP on phones controlled by male heads of household, it follows that girls' enrollment would not benefit significantly from the SMS intervention (Gigler and Bailur, 2014). Given the much lower rate of literacy among females than males in Sindh, an additional explanation is that the efforts made as part of the CDP to address literacy through the use of volunteers were not successful to engage female participants.[40]

5.3 Infrastructure

We find limited evidence of impacts from the SMS intervention on school infrastructure. The addition of ELECTIONS-CAPACITY appears to have strengthened the impact of SMS on school infrastructure, with a significant improvement in the availability of toilets. The findings suggest that the ongoing community engagement engendered by the SMS intervention increased effort by newly elected SMC members and led to improvement in school facilities identified by communities. In order to test this interpretation, we analyze the content of messages received by the CDP **[Figure 5A]**; the messages demonstrate a clear preference by communities for investment in physical infrastructure, the single largest item accounting for 38 percent of messages received. Analysis of the expenditure priorities identified by SMC members during the participatory capacity building exercise carried out as part of ELECTIONS-CAPACITY **[Figure 5B]** reveals a similar priority placed on physical infrastructure, with building, classrooms and toilets accounting for a total of 34 percent.

The MEET intervention does not appear to have driven improvements in infrastructure, either with or without the addition of ELECTIONS-CAPACITY. As in the case of availability of teachers, this suggests a one-off meeting did not provide a sufficiently sustained improvement in community engagement to drive improvement in infrastructure.

[Figures 5A and 5B about here]

5.4 Teacher attendance and learning outcomes

Despite improvements in the availability of teachers, none of the interventions significantly improved the number of teachers actually present at schools at endline. Comparison of means

[40] It is worth noting that the majority of primary schools in Sindh are mixed-gender, in contrast to other provinces of Pakistan. The share of sampled schools which were girls-only was too small to enable robust assessment of whether impacts were observed on enrollment in these schools in particular.

[Table 5] demonstrates that the share of teachers absent at endline was actually greater in MEET and SMS schools, suggesting that gains in teacher availability were balanced out by greater absenteeism. This is in line with findings from India and Kenya, which suggest that increased staffing at schools is often counteracted by a reduction in effort and presence by existing teachers (e.g. Muralidharan and Sundararaman, 2013; Duflo et al, 2015).

The result suggests that the interventions were not successful in catalyzing improved community monitoring of resource use in schools. Data from household surveys at endline reveal a high level of community engagement with schools: more than one-third (34.5 percent) of households reported having met the head teacher of their child's school on at least a monthly basis in the 2015/16 school year, two years after the intervention. However, we find no systematic evidence of greater levels of school visits in either SMS or MEET schools than in control schools; the interventions do not appear to have led to an increase in the quantity of school monitoring activities (although we cannot assess whether the quality and relevance of interactions with the school administration might have improved as a result of the intervention).

Furthermore, analysis of households' reported concerns about their children's schools reveals that teacher absence was not considered a significant problem by household heads, suggesting that community engagement may be of limited impact if misconceptions persist about the severity of problems in service. More generally, the findings are in line with those from Niger (Beasley and Huillery, 2015) that communities are unlikely to successfully monitor school resource use without significant investments in capacity.

Neither MEET nor SMS had any significant impacts on learning outcomes. This finding is consistent with the lack of observed impact on teacher effort with no improvement in teachers' presence at school. It is also possible that, while individual students did experience some benefit in learning, these effects were offset by the impacts of increased retention of lower-performing students, reducing average performance; this is aligned with Crawfurd (2018) where the authors find null effects on learning from contracting-out of public schools in Punjab, Pakistan to a private company, despite significant impacts on access. The lack of impact on learning outcomes may simply reflect the relatively short period (two years) between baseline and endline and reinforces the findings of Bruns et al. (2011) that the impacts of school community engagement reforms typically take several years to materialize.

5.5 Cost effectiveness

Figure 6 presents the implementation cost for all three interventions: MEET, SMS and ELECTIONS-CAPACITY. Costs for both MEET and SMS were similar, totaling US$129,059 for MEET and US$138,369 for SMS. Taking the total endline enrollment of treated schools as the population of impact, this suggests a per-student cost of US$16.26 per student for MEET and US$19.41 for SMS. With 25 percent of expenditures invested at the design stage, per-student expenditures for ongoing and scaled-up activities are estimated to be significantly lower: US$12.85 for MEET and US$15.66 for SMS.

[Figure 6 about here]

Taking this variable cost as a starting point, **Table 13** presents cost effectiveness estimates of the intervention. Improvement in enrollment is achieved from two intervention impacts: increase in enrollment in schools which remained open at baseline and endline, and additional enrollment achieved through the opening of closed schools. For every $100 spent on operational costs, SMS intervention contributed to 1.07 additional children enrolled in school compared to 0.91 for MEET. These estimates are net of improvements in enrollment in the control group and are adjusted by scaling-up the gains in control group to the equivalent number of schools in each treatment.

[Table 13 about here]

6. Conclusion

This paper studies the impact of three treatments to test alternative approaches aimed at bolstering community engagement in schools: community-wide dialogue at an externally facilitated meeting; continuous dialogue via text messages, on a high-frequency, anonymous and low-cost text message platform; and elections and training of SMC members. We find that measures aimed at creating community-wide, inclusive platforms to promote dialogue lead to greater interest of the community in education with large impacts on access and staffing. However, attempts to reform an existing institution, the SMC, in a cross-over design, had limited effect. In most cases it undermined the effects of dialogue interventions.

We find both SMS and MEET to be largely successful in drawing more children to school, and also in reducing teacher shortages in schools. SMS proved marginally more effective with its stronger impact on teacher availability in schools as well as on retention of children throughout the primary school cycle. We suspect these impacts arise from anonymity of exchange and continuity in dialogue. However, we caution policy makers to carefully consider the time, effort and investment that is needed for such an ICT infrastructure to be developed, deployed and iteratively adapted to respond to local needs. Even in this controlled experiment, several parameters had to be meticulously calibrated with the local firm to make the ICT platform work. We expect future iterations of the intervention to study more carefully the gender divide in take-up of the intervention and impacts. Technology, in our case, appears to have reproduced the gender inequities both at the stage of engagement with the platform, as well as in outcomes, on access to schooling for girls.

The field experiment raises some interesting questions for future research work. The significant decline in attendance at community-wide meetings which included the TEO, the lowest government administrative functionary, summarizes in a snapshot a possibly broken relationship of citizens with the state. The mere presence of a state official at the time of the village meeting may have reduced the credibility of the intervention among community members. Further research is required to more fully illuminate the ways in which trust in government affects the relative effectiveness of community engagement mechanisms incorporating a greater or lesser degree of official involvement.

References

Abraham, A. and Platteau, J.P. 2002. "Participatory Development in the Presence of Endogenous Community Imperfections." *Journal of Development Studies* 39(2):104-136.

Andrabi, T., Das, J., Khwaja, A., Vishwanath, T. and Zajonc, T. 2007. "Learning and Educational Achievements in Punjab: Insights to inform the education policy debate." Mimeo.

ASER. 2015. "Annual Status of Education Report (ASER) – Pakistan 2015." ASER Centre and South Asian Forum for Education Development.

Asian Development Bank and World Bank. 2011. "Pakistan Floods 2010: Preliminary Damage and Needs Assessment."

Asian Development Bank and World Bank. 2012. "Pakistan Floods 2011: Preliminary Damage and Needs Assessment."

Asim, S. 2013. "The Public School System in Sindh: Empirical Insights." *Lahore Journal of Economics* 18, 49-66.

Asim, S. 2014. "Analysis of Textbook-Curriculum alignment in Pakistan." Mimeo.

Asim, M. and Dee, T. 2016. "Mobile Phones, Civic Engagement, and School Performance In Pakistan." National Bureau of Economic Research (NBER) Working Paper 22764. Cambridge: NBER.

Babur, M. 2016. "Exploring Organizational Learning in the Context of Power and Corruption: a Cultural Historical Analysis of Public Sector Management In Pakistan." PhD Thesis. Njimegen: Radboud University.

Banerjee, A. V., Banerji, R., Duflo, D., Glennerster, R., and Khemani, S. 2010. "Pitfalls of Participatory Programs: Evidence from a randomized evaluation in education in India." *American Economic Journal: Economic Policy* 2(1), 1-30.

Bardhan, Pranab K., and Dilip Mookherjee. 2000. "Capture and Governance at Local and National Levels." *American Economic Review* 90, no. 2 (2000):135–139.

Barr, A., Mugisha, F., Serneels, P. and Zeitlin, A. 2012. "Information and collective action in community-based monitoring of schools: Field and lab experimental evidence from Uganda." Mimeo.

Barrera-Osorio, F., Fasih, T. and Patrinos, H. 2009. *Decentralized Decision-Making in Schools: The Theory and Evidence on School Based Management.* Washington, D.C.: World Bank Publications.

Barrera-Osorio, F., Blakeslee, D., Hoover, M., Linden, L., Raju, D. and Ryan, S. 2017. "Leveraging the Private Sector to Improve Primary School Enrolment: Evidence from a Randomized Controlled Trial in Pakistan." Mimeo.

Barrera-Osorio, Felipe; De Barros, Andreas; Filmer, Deon P. 2018. "Long-term impacts of alternative approaches to increase schooling: evidence from a scholarship program in Cambodia." Policy Research working paper; no. WPS 8566

Beasley, E. and Huillery, E. 2015. "Willing but Unable: Short-Term Experimental Evidence on Parent Empowerment and School Quality." The World Bank Economic Review, Volume 31, Issue 2, 1 June 2017, Pages 531–552

Bengali, K. 2015. *Profiles of Land Tenure System in Pakistan.* Karachi: Pakistan Institute of Labor Education and Research.

Benveniste, L., and J. Marshall. 2004. "School Grants and Student Performance: Evidence from the EQIP Project in Cambodia." Mimeo.

Berlinski, S., Busso, M., Dinkelman, T. and Martinez, C. 2016. "Reducing parent-school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile." Mimeo.

Bjorkman, M. and J, Svensson. 2006. "Power to the People: Evidence from a Randomized Experiment of a Community Based Monitoring Project in Uganda." Mimeo. IIES, Stockholm University

Blimpo, M. and Evans, D. 2011. "School-Based Management and Educational Outcomes: Lessons from a Randomized Field Experiment." Mimeo.

Bruns, B., Filmer, D., and Patrinos, H. 2011. "Making schools work: New evidence on accountability reforms." Washington, D.C.: World Bank Publications.

Carneiro, P., Das, J., and Reis, H. 2010. "Estimating The Demand for School Attributes in Pakistan." Mimeo.

Cerdan-Infantes, P., and Filmer, D. 2015. "Information, knowledge and behavior: evaluating alternative methods of delivering school information to parents." World Bank Policy Research Working Paper 7233. Washington, D.C.: World Bank Publications.

Crawfurd, L. 2018. "Contracting Out Schools at Scale: Evidence from Pakistan." Mimeo.

Croke, K., Grossman, G., Larreguy, H. and Marshall, J. 2015. "Deliberate Disengagement: How Education Decreases Political Participation In Electoral Authoritarian Regimes." Mimeo.

De Grauwe, A. 2005. "Improving the Quality of Education through School-Based Management: Learning from International Experiences." *International Review of Education* 51(4), 269–87.

di Gropello, E. and Marshall, J. 2005. "Teacher effort and schooling outcomes in rural Honduras." In *Incentives to Improve Teaching*, ed. E. Vegas. Washington D.C.: World Bank.

Duflo, E., Dupas, P., and Kremer, M. 2015. "School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools." *Journal of Public Economics* 123, 92-110.

Gazdar, H. 2000. "Public Failure, Political Constraints and Political Resources - Basic Education in Pakistan". Working Paper. London: London School of Economics and Political Science.

Gertler, P., Rubio-Codina, M., and Patrinos, H. 2011. "Empowering Parents to Improve Education: Evidence from Rural Mexico." Journal of Development Economics 99:68–79.

Gigler, B. and Bailur, S. (Eds). 2014. "Closing the Feedback Loop: Can Technology Bridge the Accountability Gap?" Washington, D.C.: World Bank Publications.

Heller, P. and Rao, V. 2015. "Deliberation and Development: Rethinking the Role of Voice and Collective Action in Unequal Societies." *Equity and Development Series*. Washington, D.C.: World Bank Publications.

Hirschman, A. 1970. *Exit, Voice, and Loyalty.* Cambridge, MA: Harvard University Press.

Hussain, G., Mohyuddin, A. and Ahmed, S. 2013. "Nature of Conflicts, Tensions and Exploitation in Sharecropping in Rural Sindh." *Open Journal of Applied Sciences*, 2013(3):482-489.

Jacoby, H., and Mansuri, G. 2011. "Crossing Boundaries: Gender, Caste and Schooling in Rural Pakistan." Policy Research Working Paper 5710. Washington, D.C.: World Bank Publications.

Joshi, A. 2013. "Do they work? Assessing the impact of transparency and accountability initiatives in service delivery." *Development Policy Review* 31(s1):s29-s48.

Mansuri, G. and Rao, V. 2013. *Localizing Development: Does Participation Work?* Washington, D.C.: World Bank Publications.

McKenzie, D. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99:210-221.

Menocal, R. A., and Sharma, B. 2008. "Joint Evaluation of Citizens' Voice and Accountability: Synthesis Report." London: DFID.

Miguel, E and Gugerty, M.K. 2005. "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of Public Economics* 89:2325–2368

Muralidharan, K. and Sundararaman, V. 2013. “Contract Teachers: Experimental Evidence from India.” Mimeo, UC San Diego.

Pakistan Telecommunication Authority. 2014. “Annual Report: 2014.” Islamabad: Pakistan Telecommunication Authority.

Pandey, P., Goyal, S. and Sundararaman, V. 2009. "Community participation in public schools: impact of information campaigns in three Indian states." *Education Economics* 17(3):355-375.

Platteau, J. P., and Abraham, A. 2002. "Participatory Development in the Presence of Endogenous Community Imperfections," *Journal of Development Studies* 39(2):104-136.

Pradhan, M, Suryadarma, D., Beatty, A., Wong, M., Gaduh, A., Alisjahbana, A. and Artha, R. 2014. "Improving educational quality through enhancing community participation: results from a randomized field experiment in Indonesia." *American Economic Journal: Applied Economics* 6.2:105-126.

Pakistan Bureau of Statistics. 2011. *Pakistan Social and Living Standards Measurement (PSLM) Survey, 2010-11.* Islamabad: Government of Pakistan.

Reinikka, R. and Svensson, J. 2005. “Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda.” Journal of the European Economic Association, 3(2):259-267

Wooldridge, J. 2012. *Introductory Econometrics: A Modern Approach.* 5th ed. South-Western, Cengage Learning.

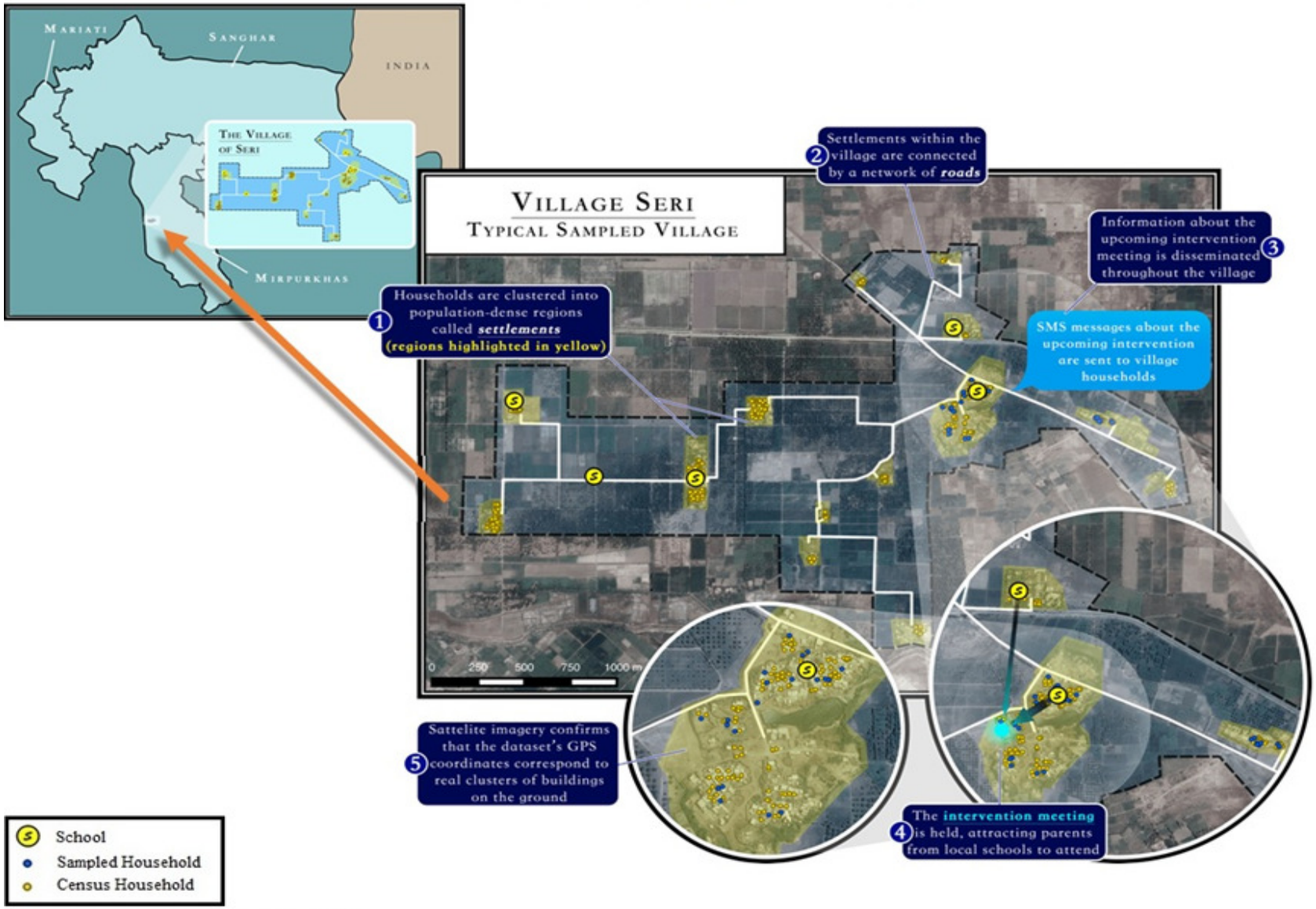
World Bank. 2004. “World Development Report 2004: Making Services Work for the Poor.” Washington, D.C.: World Bank Publications.

World Bank. 2012. “Implementation Completion and Results Report on a Credit in the Amount Of SDR 232.7 Million (US$ 350 Million Equivalent) to the Islamic Republic of Pakistan for a Sindh Education Sector Project (SEP).” Washington, D.C.: World Bank Publications.

World Bank. 2018. “World Development Report 2018: Learning to Realize Education’s Promise.” Washington, D.C.: World Bank Publications.

TABLE AND FIGURES

Figure 1. Sampled Village, Seri in Mirpurkhas District



Data Source: Census level village mapping, 2011

Figure 2. Population and Sampling Profile

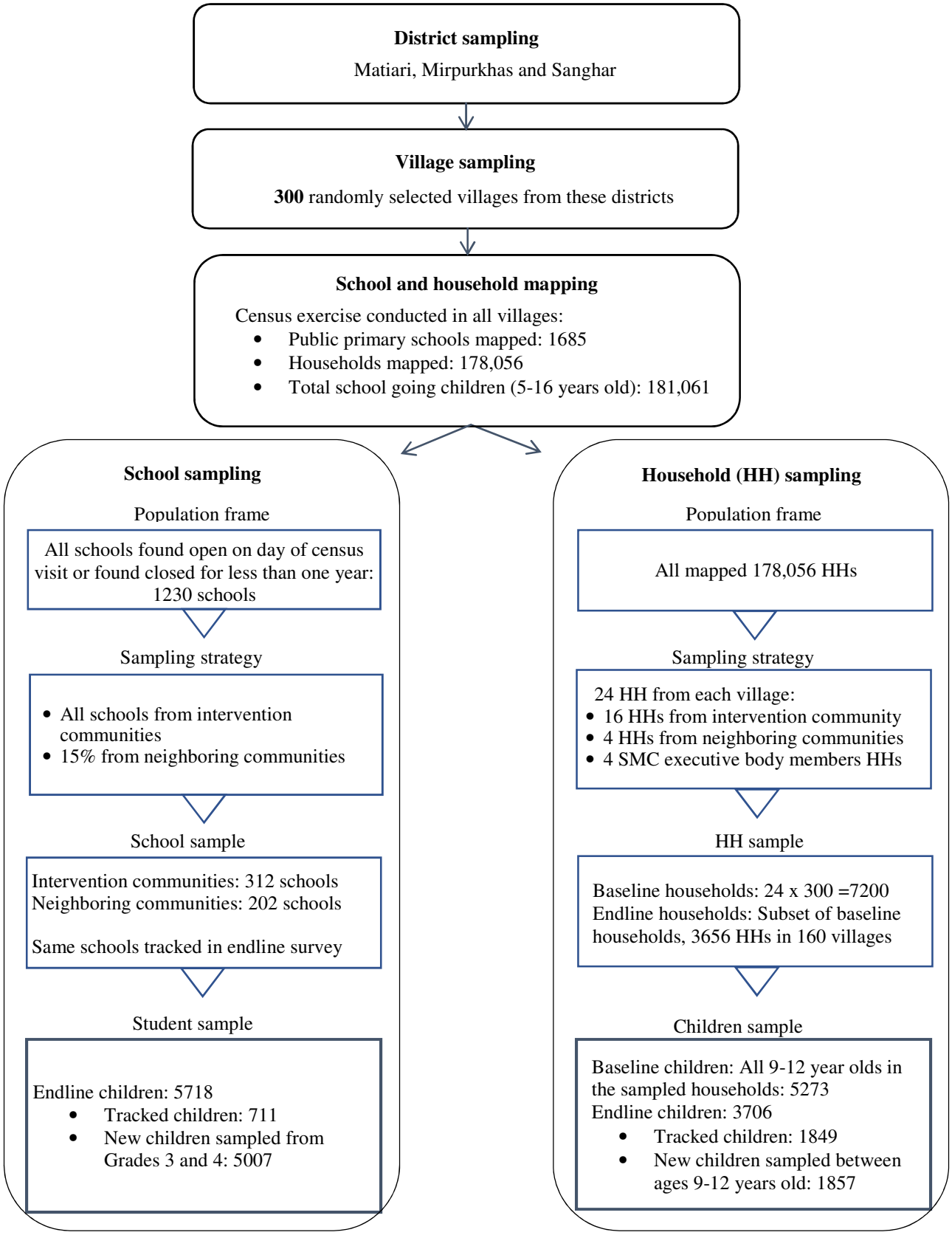


Figure 3. Mechanisms and outcomes

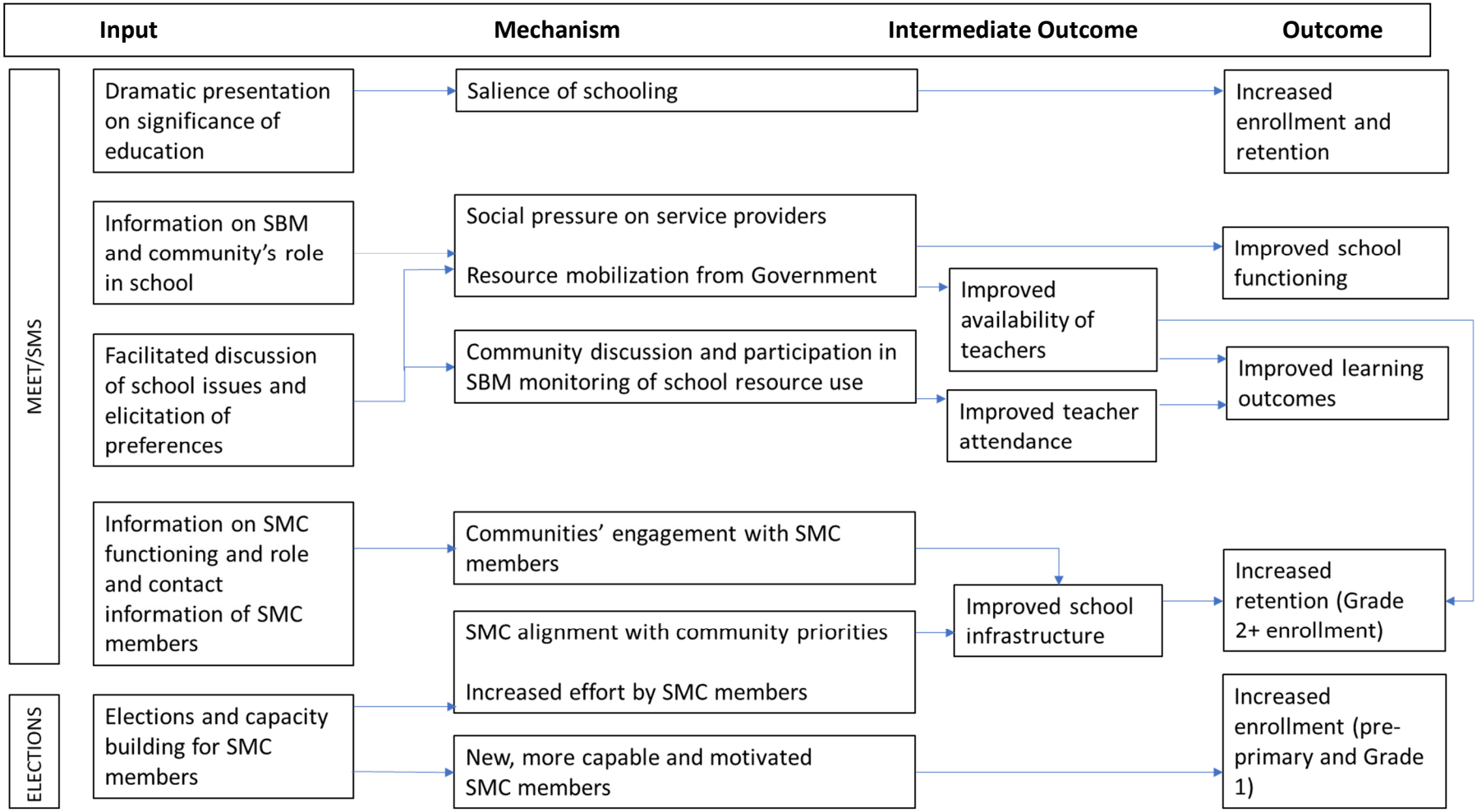


Table 3: Pre-Treatment Balance in Village and School Characteristics across Interventions

	PANEL A: MEAN/SD					PANEL B: Differences in Mean (T-C)				
	Control	**MEET**	**SMS**	**MEET-E**	**SMS-E**	**MEET-C**	**SMS-C**	**MEET-E-C**	**SMS-E-C**	**N**
Village Characteristics:										
Number of primary Schools	5.54	5.28	5.05	5.65	6.13	-0.27	-0.49	0.11	0.58	284
	(3.36)	(3.09)	(2.94)	(3.35)	(3.19)	(0.41)	(0.39)	(0.44)	(0.43)	
School Open (Unannounced)	0.83	0.74	0.74	0.87	0.76	-0.09	-0.09	0.04	-0.07	284
	(0.34)	(0.42)	(0.41)	(0.32)	(0.39)	(0.06)	(0.05)	(0.04)	(0.05)	
Village affected by Flood	0.72	0.74	0.71	0.67	0.66	0.02	-0.01	-0.05	-0.06	284
	(0.45)	(0.44)	(0.46)	(0.48)	(0.48)	(0.06)	(0.06)	(0.06)	(0.06)	
Share of Households with Cell Phones	0.76	0.75	0.72	0.73	0.74	0.01	-0.04	-0.03	-0.03	284
	(0.22)	(0.19)	(0.20)	(0.21)	(0.19)	(0.03)	(0.03)	(0.03)	(0.03)	
School Characteristics:										
School Size (Enrollment)	73.89	85.70	76.21	76.62	74.96	11.81	2.32	2.73	1.08	249
	(43.81)	(50.55)	(44.72)	(44.27)	(39.17)	(7.01)	(6.59)	(6.03)	(5.49)	
Total Boys	47.71	55.44	49.99	52.26	52.46	7.73	2.28	4.55	4.75	249
	(36.38)	(36.57)	(34.82)	(34.27)	(35.61)	(5.07)	(5.14)	(4.66)	(4.99)	
Total Girls	25.63	28.05	25.66	23.42	22.00	2.43	0.04	-2.21	-3.63	249
	(23.74)	(26.70)	(16.78)	(22.25)	(16.94)	(3.70)	(2.47)	(3.03)	(2.37)	
Total Teachers	2.29	2.30	2.28	2.26	2.38	0.02	-0.01	-0.03	0.09	249
	(1.78)	(1.81)	(1.51)	(1.62)	(1.86)	(0.25)	(0.22)	(0.22)	(0.26)	
One teacher School	0.46	0.48	0.40	0.43	0.39	0.01	-0.06	-0.03	-0.07	249
	(0.47)	(0.44)	(0.43)	(0.48)	(0.42)	(0.06)	(0.06)	(0.07)	(0.06)	
Teacher Absence Rate	0.09	0.06	0.08	0.08	0.08	-0.03	-0.01	-0.01	-0.01	249
	(0.18)	(0.17)	(0.17)	(0.17)	(0.18)	(0.02)	(0.03)	(0.02)	(0.03)	
Number of Classrooms	2.13	2.19	2.35	1.88	2.15	0.06	0.22	-0.25	0.02	249
	(1.23)	(1.56)	(1.56)	(1.09)	(1.04)	(0.22)	(0.23)	(0.15)	(0.15)	
Toilets Available	0.68	0.61	0.64	0.61	0.67	-0.07	-0.04	-0.07	-0.01	249
	(0.42)	(0.46)	(0.45)	(0.43)	(0.41)	(0.06)	(0.07)	(0.06)	(0.06)	
Schools with Electricity	0.51	0.45	0.43	0.44	0.45	-0.07	-0.08	-0.07	-0.07	249
	(0.46)	(0.46)	(0.45)	(0.46)	(0.44)	(0.06)	(0.07)	(0.06)	(0.06)	

Schools with Drinking Water Source	0.44	0.40	0.42	0.51	0.55	-0.04	-0.02	0.07	0.11	249
	(0.45)	(0.44)	(0.43)	(0.44)	(0.46)	(0.06)	(0.06)	(0.06)	(0.06)	
Schools with Boundary Wall	0.69	0.70	0.61	0.58	0.79	0.01	-0.09	-0.12	0.10	249
	(0.40)	(0.43)	(0.44)	(0.45)	(0.32)	(0.06)	(0.07)	(0.06)	(0.05)	
Household is aware of:										
School Committee (SMC) exists	0.07	0.08	0.07	0.05	0.07	0.01	0.00	-0.02	0.01	284
	(0.08)	(0.09)	(0.07)	(0.06)	(0.10)	(0.02)	(0.01)	(0.01)	(0.02)	
SMC has five members	0.03	0.05	0.04	0.02	0.04	0.02	0.01	-0.01	0.01	284
	(0.04)	(0.07)	(0.05)	(0.03)	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	
SMC meets at least twice a year	0.04	0.03	0.03	0.02	0.04	-0.01	0.00	-0.02	0.00	284
	(0.06)	(0.05)	(0.05)	(0.03)	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	
SMC receives PKR 22,000 annually	0.06	0.06	0.06	0.04	0.06	0.00	0.00	-0.03*	0.00	284
	(0.07)	(0.07)	(0.07)	(0.04)	(0.10)	(0.01)	(0.01)	(0.01)	(0.02)	
School Improvement Plan (SIP)	0.03	0.04	0.03	0.02	0.04	0.01	0.01	-0.01	0.02	284
	(0.05)	(0.05)	(0.05)	(0.03)	(0.08)	(0.01)	(0.01)	(0.01)	(0.01)	
Children test scores (age 7-13):										
English Test Score (correct percent)	0.32	0.33	0.30	0.35	0.33	0.01	-0.01	0.03	0.01	256
	(0.19)	(0.15)	(0.16)	(0.18)	(0.18)	(0.03)	(0.03)	(0.04)	(0.04)	
Math Test Score (correct percent)	0.27	0.28	0.29	0.32	0.27	0.01	0.02	0.05	0.00	256
	(0.22)	(0.19)	(0.20)	(0.22)	(0.22)	(0.04)	(0.04)	(0.04)	(0.04)	

Sources: Baseline School and Household Survey, 2012 prior to the intervention. Villages were randomly assigned to one of the five groups: (i) No meetings (Control), (ii) village meeting (MEET); (iii) Text Messages (SMS); (iv) village meeting with elections for SMC executive members (MEET-E); and (v) Text messages with election for SMC executive members (SMS E). Panel A reports Mean and Standard deviation for each of the five groups. Panel B reports the difference in means for Treatment against Control with standard errors reported in parenthesis. Grade wise enrollment is also balanced across treatments and control groups but omitted here for brevity.

Figure 4A: Experiment Profile and Timeline

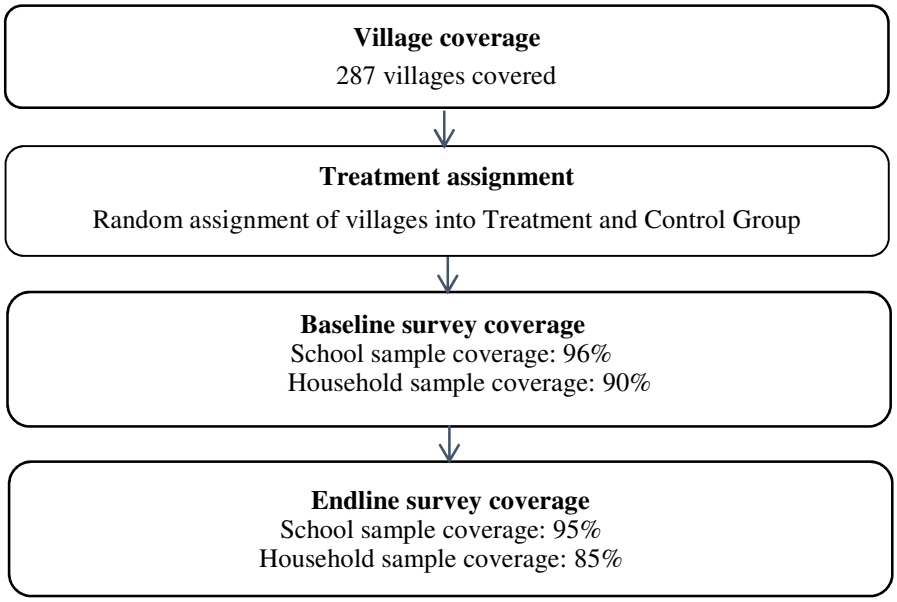


Figure 4B. Sample Coverage by Treatment Status

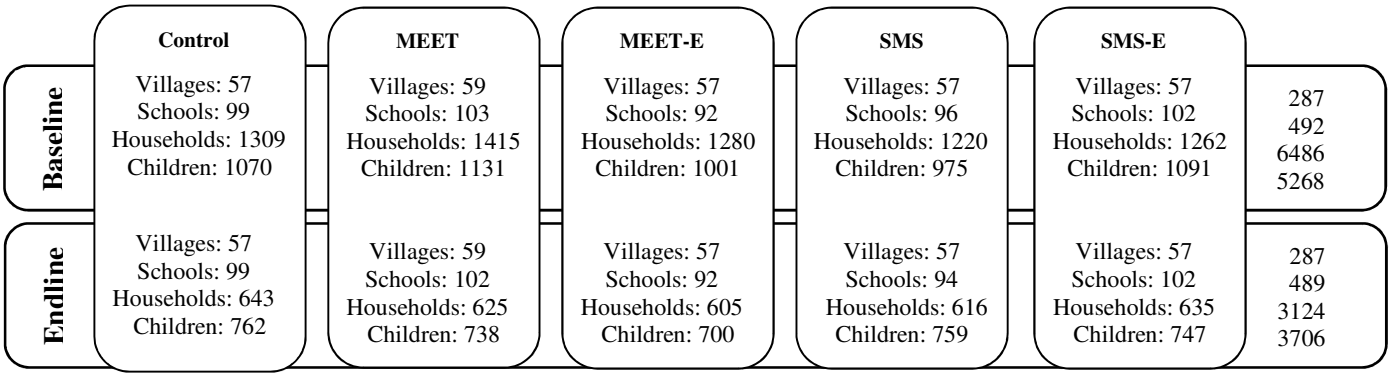
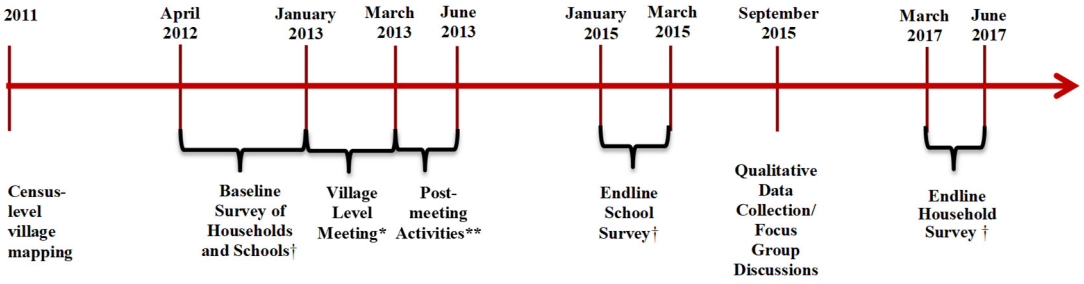


Figure 4C. Experiment Timeline



Notes: * Denotes meeting where MEET intervention; introduction of CDP under SMS intervention; and elections under ELECTIONS-CAPACITY intervention took place. ** Denotes period during which CDP was operational and capacity building for elected SMC members carried out. † Denotes testing of students.

Table 4: Treatment Fidelity

	Total	Matiari	Mirpurkhas	Sanghar
Participation Rates[1] in Community Meetings				
MEET	67%	59%	77%	58%
SMS	68%	68%	77%	62%
MEET-E	58%	42%	69%	54%
SMS-E	58%	34%	76%	52%
SMC Composition Change[2]				
MEET-E	74%	73%	74%	75%
SMS-E	71%	73%	75%	65%
Average Attendance[3] in Training for SMC				
MEET-E	96%	86%	97%	100%
SMS-E	96%	88%	96%	100%

Notes: (1) Community participation rates are estimated as the proportion of residents in a community that attended the intervention meeting in community's school. Population level data on households in a community was obtained from community level census conducted for the study in 2011. (2) SMC composition change is the share of SMC members elected as part of the intervention. (3) Average attendance of the SMC executive body member is the share that attended capacity building training sessions averaged across three meetings.

Table 5. Key Outcomes at Endline: Control and Difference in Means, Treatment-Control

	Control	Difference in means (Treatment-Control)				
		MEET	**SMS**	**MEET-E**	**SMS-E**	**N**
School level						
Open on unannounced visit	0.80	**0.09***	0.02	-0.01	0.06	479
	(0.41)	(0.03)	(0.04)	(0.04)	(0.04)	
School changes from closed to open	0.10	**0.10****	**0.12****	-0.03	0.07	479
	(0.03)	(0.05)	(0.05)	(0.04)	(0.05)	
School changes from open to closed	0.13	**-0.09****	-0.02	0.03	-0.04	479
	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	
School Size (Enrollment)	70.93	15.99	11.56	5.22	4.89	387
	(57.5)	(9.43)	(8.16)	(5.55)	(6.49)	
Enrollment (Boys)	47.04	10.77	11.57	7.91	8.39	387
	(6.66)	(9.74)	(9.43)	(8.38)	(8.70)	
Enrollment (Girl)	23.92	6.18	-0.94	-2.82	-2.94	387
	(3.13)	(5.60)	(4.17)	(4.02)	(3.76)	
One Teacher School	0.52	**-0.20****	**-0.17****	-0.10	-0.15	387
	(0.50)	(0.08)	(0.08)	(0.08)	(0.08)	
Total Teachers	2.25	0.56	0.38	0.18	0.18	387
	(2.01)	(0.38)	(0.36)	(0.31)	(0.32)	
Teachers Present	1.96	0.26	0.22	0.08	-0.04	387
	(0.22)	(0.32)	(0.32)	(0.29)	(0.29)	
Teacher Absence Rate	0.05	**0.06****	0.03	0.03	**0.08****	387
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	
Number of classrooms	2.18	0.20	0.39	-0.02	0.21	387
	(0.16)	(0.24)	(0.26)	(0.21)	(0.21)	
Toilets available	0.75	0.06	0.02	-0.06	0.10	387
	(0.05)	(0.07)	(0.07)	(0.07)	(0.06)	
Schools with electricity	0.51	0.05	-0.01	0.11	0.01	387
	(0.50)	(0.06)	(0.07)	(0.06)	(0.06)	
Schools with drinking water source	0.60	-0.14	-0.07	0.08	-0.07	387
	(0.49)	(0.05)	(0.06)	(0.06)	(0.07)	
Schools with boundary wall	0.70	0.09	-0.04	-0.03	0.12	387
	(0.46)	(0.04)	(0.06)	(0.06)	(0.04)	
Children test scores (household level))						
English IRT scaled scores	0.07	-0.04	-0.06	-0.04	-0.02	2837
	(0.86)	(0.05)	(0.05)	(0.06)	(0.05)	
Math IRT scaled scores	-0.02	-0.08	-0.07	0.04	-0.01	2476
	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)	

Notes: Endline School and Household Surveys 2015-2017. Standard errors are clustered at village level in parenthesis. School and Teacher characteristics reported at school level. Children characteristics are reported at village level.

Table 6. Transition Probabilities by Treatment Status

	Control	MEET	SMS	MEET-E	SMS-E	Transitions
Open to Closed	13.27	**-8.97****	-1.5	3.21	-3.56	53
Closed to Open	10.20	10.23	**12.14****	-2.51	7.28	75
Added a Teacher	17.81	**17.63****	**11.92***	10.40	11.11	109
Added a Classroom	15.07	7.71	2.50	6.72	**12.64****	82
Added a Toilet	16.44	5.08	-0.22	1.51	**5.45***	74

Notes: *** p < 0.01, ** p < 0.05, * p < 0.1 The table reports the odds of treatment schools (1) switching status from closed to open and vice versa; (2) getting an additional teacher deployed; and (3) adding a classroom or toilet, when compared to controls. Mean for control group is reported in column (1).

Table 7: Impact on School Functioning

	School is open on unannounced visit	School changes from closed to open	School changes from open to closed
MEET	**0.11****	0.09	**-0.09***
	(0.05)	(0.06)	(0.04)
SMS	0.05	**0.13***	-0.02
	(0.06)	(0.06)	(0.05)
MEET-E	-0.01	-0.02	0.03
	(0.06)	(0.05)	(0.05)
SMS-E	0.07	0.08	-0.04
	(0.06)	(0.05)	(0.05)
Observations	479	479	479
Adjusted R-squared	0.04	0.07	0.02
Mean of Control	0.79	0.10	0.13
Equivalence Test:			
MEET = SMS	1.4	0.19	**3.43***
MEET-E = SMS-E	1.9	**4.06****	1.89

Notes: Standard errors clustered at village level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Regression estimates control for whether school was affected by flood and for district fixed effects. F-statistic is reported to test the equivalency (H0) for the two dialogue interventions.

Table 8A: Impact on School Size (Enrollment)

	Total	Boys	Girls
MEET	4.56	0.36	3.81
	(5.98)	(4.39)	(3.12)
SMS	**9.22***	**10.06****	-1.06
	(5.55)	(4.83)	(2.53)
MEET-E	2.09	3.39	-1.39
	(4.73)	(3.83)	(2.23)
SMS-E	5.88	5.84	-0.52
	(4.77)	(3.87)	(1.91)
Observation	387	387	387
R-squared	0.67	0.70	0.65
Control group mean	70.88	47.04	23.92
Equivalence Test:			
MEET = SMS	0.42	**2.89***	1.97
MEET-E = SMS-E	0.52	0.31	0.18

Notes: Standard errors clustered at village level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Regression estimates control for whether school was affected by flood and for district fixed effects. F-statistic is reported to test the equivalency (H0) for the two dialogue interventions.

Table 8B: Enrollment Impacts by Grade (Boys)

	Grade1	Grade2	Grade3	Grade4	Grade5
MEET	1.85	-1.49	-0.13	0.57	0.75
	(2.89)	(1.18)	(0.94)	(0.75)	(0.80)
SMS	3.35	2.60*	2.12**	1.77**	0.51
	(2.60)	(1.47)	(1.03)	(0.89)	(0.66)
MEET-E	1.65	1.41	-0.66	1.59*	0.08
	(2.04)	(1.22)	(1.02)	(0.81)	(0.65)
SMS-E	4.05*	1.54	0.32	0.92	-0.63
	(2.10)	(1.16)	(0.82)	(0.72)	(0.66)
Observations	387	387	387	387	387
R-squared	0.54	0.52	0.55	0.50	0.45
Control group mean	18.93	8.74	7.96	6.01	5.39
Equivalence Test:					
MEET = SMS	1.97	**9.43*****	**4.11****	1.68	0.08
MEET-E = SMS-E	0.18	0.01	0.97	0.67	1.10

Notes: Standard errors clustered at village level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Regression estimates control for whether school was affected by flood and for district fixed effects. Grade 1 also include pre-primary (*katchi*). F-statistic is reported to test the equivalency (H0) for the two dialogue interventions.

Table 8C: Enrollment Impacts by Grade (Girls)

	Grade1	Grade2	Grade3	Grade4	Grade5
MEET	2.64	0.84	0.52	-0.05	0.63
	(1.89)	(0.49)	(0.68)	(0.07)	(0.63)
SMS	-1.43	0.60	0.09	-0.00	-0.28
	(1.51)	(0.63)	(0.62)	(0.59)	(0.47)
MEET-E	-1.34	0.28	0.24	-0.34	-0.15
	(1.21)	(0.53)	(0.70)	(0.52)	(0.50)
SMS-E	0.21	-0.11	-0.14	-0.70	-0.43
	(1.15)	(0.44)	(0.53)	(0.43)	(0.42)
Observations	387	387	387	387	387
R-squared	0.45	0.49	0.34	0.36	0.22
Control group mean	9.74	3.97	3.99	3.38	2.92
Equivalence Test:					
MEET = SMS	**3.55***	0.13	0.36	0.00	2.02
MEET-E = SMS-E	1.78	0.53	0.40	0.81	0.52

Notes: Standard errors clustered at village level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Regression estimates control for whether school was affected by flood and for district fixed effects. Grade 1 also include pre-primary (*katchi*). F-statistics is reported to test the equivalency (H0) for the two dialogue interventions.

Table 9: Impact on availability of teachers and teacher attendance

	One-Teacher School	Total Teachers	Teachers Present	Teacher Absence Rate
MEET	**-0.15****	0.26	-0.00	**0.06****
	(0.07)	(0.18)	(0.17)	(0.03)
SMS	**-0.13***	**0.30***	0.24	0.04
	(0.07)	(0.18)	(0.20)	(0.03)
MEET-E	-0.08	0.13	0.10	0.03
	(0.07)	(0.16)	(0.15)	(0.03)
SMS-E	-0.10	0.15	0.00	**0.08****
	(0.07)	(0.17)	(0.17)	(0.04)
Observations	387	387	387	387
R-squared	0.33	0.74	0.61	0.03
Mean of Control	0.52	2.25	1.96	0.05
Equivalence Test:				
MEET = SMS	0.12	0.03	1.16	0.47
MEET-E = SMS-E	0.08	0.02	0.30	1.40

Notes: Standard errors clustered at village level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Regression estimates control for school size (enrollment), whether school was affected by flood and for district fixed effects. One-teacher school is a dummy which takes a value of 1 if the school is one-teacher at the end-line and 0 otherwise. Total Teachers shows number of teachers employed at school. Teachers present shows number of teachers present at unannounced or announced visit. Teacher shows share of employed teachers not present at announced or unannounced Visit. F-statistic is reported to test the equivalency (H0) for the two dialogue interventions.

Table 10: Impact on school Infrastructure

	Number of Classrooms	School has:			
		Toilets	Electricity	Drinking Water	Boundary Wall
MEET	-0.05	0.04	0.07	-0.14	0.05
	(0.13)	(0.06)	(0.06)	(0.08)	(0.06)
SMS	0.23	0.01	-0.00	-0.05	-0.04
	(0.15)	(0.06)	(0.07)	(0.09)	(0.05)
MEET-E	0.04	-0.04	0.11	0.06	0.01
	(0.13)	(0.06)	(0.07)	(0.09)	(0.05)
SMS-E	0.19	**0.10***	0.01	-0.08	0.04
	(0.12)	(0.05)	(0.06)	(0.09)	(0.06)
Observations	387	387	387	387	387
R-squared	0.67	0.23	0.31	0.11	0.44
Mean of Control	2.18	0.75	0.51	0.60	0.70
Equivalence Test:					
MEET = SMS	3.49	0.24	1.14	1.22	2.16
MEET-E = SMS-E	1.47	5.42	1.80	2.72	2.72

Notes: Standard errors clustered at village level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Regression estimates control for school size (enrollment), whether school was affected by flood and for district fixed effects. F-statistic is reported to test the equivalency (H0) for the two dialogue interventions.

Table 11: Impact on Children Test Scores

	English	Math
MEET	-0.06	-0.09
	(0.08)	(0.10)
SMS	-0.08	-0.08
	(0.08)	(0.11)
MEET-E	-0.05	-0.01
	(0.08)	(0.10)
SMS-E	-0.02	0.06
	(0.08)	(0.11)
Observations	2837	2476
R-squared	0.02	0.06
Control group mean	0.07	-0.02
Equivalence Test:		
MEET = SMS (F-stat)	0.04	0.00
MEET-E = SMS-E (F-stat)	0.16	0.45

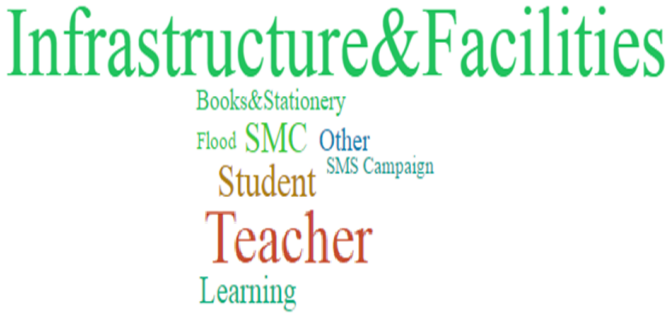
Notes: Household Survey Endline (2017). Observations at student level. Robust standard errors in parenthesis. ***p<0.01, **p<0.05, *p<0.10. English and Math test scores are scaled by fitting two-parameter logistic item response model normalized to μ=0 and ∂ =1 with respect to baseline test scores. Regression estimates control for district level fixed effects.

Table 12: Impact on Household Awareness

Household Head is aware of:	SMC exists	SMC has five members	SMC meets at least twice a year	SMC receives PKR 22,000 annually	School Improvement Plan (SIP)
MEET	0.02	-0.02	0.01	0.00	0.01
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
SMS	-0.01	-0.02	-0.02	-0.01	-0.03
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)
MEET-E	-0.00	-0.02	0.00	-0.01	-0.03
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)
SMS-E	0.03	0.01	0.01	-0.00	0.01
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Observations	160	160	160	160	160
R-squared	0.39	0.29	0.38	0.29	0.12
Mean of Control	0.22	0.18	0.19	0.20	0.11
Equivalence Test:					
MEET = SMS	1.58	0.00	1.85	0.69	**3.44***
MEET-E = SMS-E	1.31	2.33	0.04	0.30	**2.93***

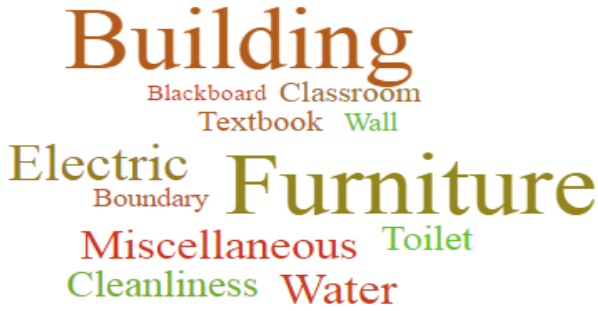
Notes: Standard errors clustered at village level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Regression estimates control for district fixed effects. Endline surveys limited to 160 random villages with 32 in each treatment arm.

Figure 5A. Elicitation of Preferences through Community Dialogue on SMS[41]



Weights per item mentioned in community dialogue	
Items	**Weights**
Books & Stationery	**2%**
Flood	**1%**
Infrastructure & Facilities	**38%**
Learning	**8%**
SMC	**10%**
SMS Campaign	**3%**
Student	**12%**
Teacher	**22%**
Other	**4%**

Figure 5B. Expenditure Items Identified by Elected Members during Training[42]

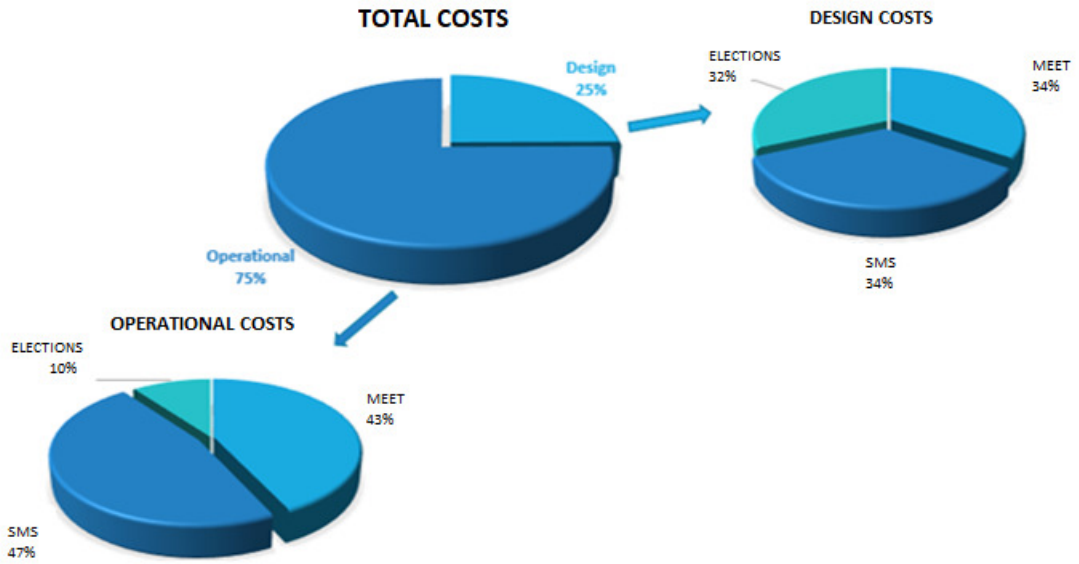


Weights per item mentioned in SIP	
Item	**Weights**
Blackboard	**1%**
Boundary Wall	**3%**
Building	**24%**
Classroom	**4%**
Cleanliness	**6%**
Electric	**12%**
Furniture	**24%**
Textbooks	**4%**
Toilet	**6%**
Water	**9%**
Miscellaneous	**8%**

[41] The CDP system included manual classification of incoming messages by topic based on word recognition. "Infrastructure and facilities" includes references to availability and condition of classrooms and other buildings, toilets, drinking water, and other amenities. "Teacher" includes issues around teacher numbers, absenteeism, and quality. "SMC" included queries and comments about the SMC or SIP funds, while "SMS Campaign" includes all comments about the CDP itself.

[42] Lists items targeted for investment by SMC members during preparation of SIP during participatory capacity building, according to monetary allocations. "Classrooms" refers to classroom construction, "buildings" refer to improvement in buildings including repairs and maintenance.

Figure 6. Project Costs



Overall intervention costs			
	MEET	**SMS**	**ELECTIONS-CAPACITY**
Design (USD)	27,083	26,737	24,891
Operational (USD)	101,976	111,632	24,294
Total (USD)	**129,059**	**138,369**	**49,185**

Notes: Cost data obtained from implementing team or from implementing partners: M3 Technologies Private Ltd., Weitek Group, J. Walter Thompson, and the Social Policy and Development Centre. Two-thirds of the cost of "Development of Audiotapes" was attributed to MEET and one-third to SMS. Conversion from PKR to USD, where required, carried out using Monthly Statistical Bulletin, Annual Report of SBP and International Financial Statistics (IFS). Monthly exchange rates used in most cases. Costs of World Bank staff calculated using 2012 annual rate (design phase) and 2013 annual rate (operational phase) and assigned equally to all three treatments. Estimates include opportunity costs for participants in village-wide meetings, estimated at half a day's wage per participant per meeting (two meetings for MEET, one for SMS/ELECTIONS-CAPACITY). The daily wage was approximated to be PKR 333, as per the Household Integrated Economic Survey (HIES) 2010-11. Conversion to USD was done using average PKR-to-USD exchange rates for 2012 and 2013.

Table 13. Cost Effectiveness (Enrollment)

	MEET	**SMS**
Enrollment impact on Schools found Open Endline	566.6	517.3
Enrollment impact in Functioning Schools	360.24	682.28
Cost per additional enrolled student (US$)	110.02	93.06
Additional students enrolled per US$100	0.91	1.07

Notes: Calculated from operational costs only to estimate scale-up costs. We estimate additional enrollment in schools found open at Endline by gains in enrollment in Treatment schools net of gains in control schools after scaling to the same number of schools as in MEET/SMS. Enrollment impact in continuously functioning schools computed from treatment estimands in Table 7 net of control schools adjusted to the number in the control group. Costs per additional student are estimated by dividing the total gains in enrollment with the operational cost of the intervention.

Appendix A

Table A1: Covariance Balance for Household Sample (baseline)

		Difference (Treatment-Control)				
	Control (1)	**MEET (2)**	**SMS (3)**	**MEET-E (5)**	**SMS-E (6)**	N
Household Characteristics						
Household is aware of:						
SMC	0.15	0.00	-0.01	-0.02	0.00	160
	(0.09)	(0.02)	(0.01)	(0.01)	(0.01)	
SMC has five members	0.07	0.00	0.00	-0.01	0.01	160
	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	
SMC meets at least twice a year	0.13	-0.01	-0.01	-0.03	0.00	160
	(0.09)	(0.02)	(0.01)	(0.01)	(0.02)	
SMC receives PKR 22,000 annually	0.13	-0.01	-0.02	-0.03	-0.01	160
	(0.09)	(0.02)	(0.01)	(0.01)	(0.01)	
School Improvement Plan (SIP)	0.06	0.01	0.01	-0.02	0.01	160
	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	
Test Score (IRT)	0.05	0.11	0.04	0.05	-0.12	155
	(0.45)	(0.09)	(0.09)	(0.09)	(0.08)	

Note: *** p < 0.01, ** p < 0.05, * p < 0.1. Variable used in this table are from the baseline data collection in 2012. Sample is restricted to 160 villages covered in Endline Household Survey. Col (1) reports mean and standard deviation for Control group at baseline. Col (2) - (7) reports on differences in means between Treatment and Control. Standard errors clustered at village level are reported in parenthesis. Test scores are scaled here using Item Response Theory Model.

Table A2: Representativeness of school sample in 300 villages

	Panel A: Sample compared to Rural Sindh			**Panel B: Sample compared to 3 Districts**		
Variable	All Districts	Sample	Difference in Means	3-Districts	Sample	Difference in Means
SMC Functional	0.95 (0.16)	0.95 (0.11)	-0.01	0.96 (0.14)	0.95 (0.11)	0.01
School Size (Enrollment)	69.03 (49.31)	56.69 (28.75)	**12.35*** **	59.98 (58.71)	56.69 (28.75)	3.29
Teachers Registered	1.85 (1.53)	1.94 (1.22)	-0.09	2.05 (1.89)	1.94 (1.23)	0.11
Teachers Registered (Male)	1.53 (1.32)	1.63 (1.02)	-0.10	1.58 (1.72)	1.63 (1.02)	-0.05
Teachers Registered (Female)	0.32 (0.93)	0.32 (0.98)	0.00	0.47 (1.17)	0.32 (0.98)	**0.16***
School Building	0.84 (0.26)	0.83 (0.23)	0.01	0.81 (0.30)	0.83 (0.23)	-0.02
Classrooms (Available)	1.92 (1.25)	1.84 (0.90)	0.08	1.94 (1.50)	1.84 (0.90)	0.11
Boundary Wall	0.58 (0.35)	0.59 (0.29)	-0.01	0.58 (0.37)	0.59 (0.29)	0.00
Toilet	0.58 (0.36)	0.61 (0.29)	-0.02	0.57 (0.39)	0.61 (0.29)	-0.04
Drinking Water	0.51 (0.40)	0.40 (0.31)	**0.11*** **	0.36 (0.38)	0.40 (0.31)	-0.04
Electricity	0.40 (0.37)	0.38 (0.32)	0.02	0.35 (0.37)	0.38 (0.32)	-0.03

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. In the columns "All District", "3-District" and "Sample", mean values given with standard deviations in parenthesis. In the columns "Difference in Means", difference calculated between our sample and overall population of schools in rural Sindh (Panel A), and population of schools in Matiari, Mirpurkhas and Sanghar (Panel B) using administrative data. Statistics provided at the village level (unit of randomization) and cover all schools that were open on the day of visit or were temporarily closed. School size is from Pre-primary to Grade 5. Teachers Registered include government teachers only. Infrastructure related variables (boundary wall, toilet, drinking water and electricity) are defined in terms of their availability at school.

Table A3: Functional School Sample and Temporarily Closed Schools

Variable	Functional Schools (Open on both visits)	Excluded Schools (Switch Status Schools)	Difference in Means
School Size (Enrollment)	78.42 (61.59)	57.60 (40.08)	**20.82****
Enrollment (Girl)	23.80 (27.74)	15.32 (17.86)	**8.48****
Enrollment (Boy)	54.61 (54.90)	42.27 (36.21)	**12.34***
Total Teachers (Available)	2.51 (2.16)	1.73 (1.53)	**0.78****
Share of Teachers Absent	0.10 (0.20)	0.14 (0.27)	-0.04
Number of Classrooms	2.33 (1.44)	2.00 (1.14)	0.33
Toilets available	0.78 (0.42)	0.75 (0.44)	0.03
Schools with Electricity	0.53 (0.50)	0.55 (0.50)	-0.01
Schools with Drinking Water Source	0.56 (0.50)	0.55 (0.50)	0.02
Schools with Boundary Wall	0.73 (0.45)	0.64 (0.49)	0.09

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. In the columns "Functional Schools" and "Temporarily Closed Schools", mean values given with standard deviations in parenthesis. In the column "Difference in means", difference calculated between 387 schools that were open on the day of visit (announced round) and that completed all modules of the survey at both baseline and endline. and 44 schools found closed in baseline but functional in endline.

Table A4-1: Analysis of Differential Attrition—MEET

	Attritor C mean/sd	Attritor MEET mean/sd	Non-Attritor C mean/sd	Non-Attritor MEET mean/sd	Diff-in-diffs coefficient/se
School Size (Enrollment)	85.31	71.38	81.71	87.45	-8.54
	(39.28)	(21.68)	(54.22)	(62.08)	(14.83)
Enrollment (boy)	55.99	45.94	52.57	57.62	-7.07
	(42.78)	(15.76)	(51.61)	(48.44)	(12.98)
One Teacher School	0.58	0.61	0.51	0.45	0.00
	(0.33)	(0.24)	(0.50)	(0.50)	(0.14)
Total Teachers	1.94	1.92	2.26	2.43	-0.25
	(0.76)	(0.57)	(1.91)	(2.28)	(0.53)
Schools affected by flood	0.68	0.47	0.66	0.76	-0.26
	(0.48)	(0.52)	(0.48)	(0.43)	(0.14)
Teacher Absence Rate	0.95	0.96	0.91	0.93	0.04
	(0.10)	(0.04)	(0.19)	(0.15)	(0.05)
Number of classrooms	1.20	0.53	2.07	2.26	-0.51
	(1.53)	(0.83)	(1.47)	(2.07)	(0.45)
Toilets available	0.36	0.27	0.64	0.64	-0.01
	(0.49)	(0.46)	(0.48)	(0.48)	(0.14)
Schools with boundary wall	0.52	0.53	0.66	0.74	-0.07
	(0.51)	(0.52)	(0.48)	(0.44)	(0.14)
Schools with drinking water source	0.44	0.47	0.52	0.46	0.07
	(0.39)	(0.35)	(0.50)	(0.50)	(0.14)
Schools with electricity	0.19	0.36	0.22	0.23	0.10
	(0.32)	(0.41)	(0.42)	(0.42)	(0.12)
Correct percent of English	0.21	0.26	0.25	0.22	0.05
	(0.12)	(0.12)	(0.19)	(0.17)	(0.05)
Correct percent of Math	0.23	0.26	0.27	0.24	0.02
	(0.14)	(0.09)	(0.20)	(0.18)	(0.05)
N	25	15	73	78	191
Joint Chi-square					8.17

Notes: All variables measured at baseline. Columns 1-4 report the means for the Control Group Attritors, the MEET-Group Attritors, the Control and the MEET observations with complete data (surveyed). Standard deviations in parentheses. Last Column is the difference between the treatment group mean and the control group mean among the attritors minus the difference between the treatment group mean and the control group mean among those schools that reported data. Differences in means are computed by OLS controlling for district fixed effects. Standard errors in parentheses are clustered at the village level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The chi-square is the result of a test for individual coefficients to be jointly equal to zero using seemingly unrelated estimation.

Table A4-2: Analysis of Differential Attrition—MEET-E

	C-Attritor mean/sd	MEET_E-Attritor mean/sd	C-Non-attritor mean/sd	MEET_E-Non-attritor mean/sd	Diff-in-diffs coefficient/se
School Size (Enrollment)	85.31	56.29	81.71	83.55	-17.57
	(39.28)	(23.15)	(54.22)	(54.43)	(15.56)
Enrollment (boy)	55.99	37.42	52.57	53.61	-10.02
	(42.78)	(16.21)	(51.61)	(40.21)	(13.62)
One Teacher School	0.58	0.67	0.51	0.45	0.06
	(0.33)	(0.27)	(0.50)	(0.50)	(0.15)
Total Teachers	1.94	1.77	2.26	2.29	-0.05
	(0.76)	(0.63)	(1.91)	(1.77)	(0.56)
Schools affected by flood	0.68	0.69	0.66	0.59	0.24
	(0.48)	(0.48)	(0.48)	(0.50)	(0.15)
Teacher Absence Rate	0.95	0.91	0.91	0.91	0.00
	(0.10)	(0.28)	(0.19)	(0.17)	(0.06)
Number of classrooms	1.20	0.54	2.07	1.90	0.08
	(1.53)	(0.88)	(1.47)	(1.25)	(0.47)
Toilets available	0.36	0.15	0.64	0.58	-0.02
	(0.49)	(0.38)	(0.48)	(0.50)	(0.15)
Schools with boundary wall	0.52	0.46	0.66	0.59	0.07
	(0.51)	(0.52)	(0.48)	(0.50)	(0.15)
Schools with drinking water source	0.44	0.39	0.52	0.52	-0.03
	(0.39)	(0.30)	(0.50)	(0.50)	(0.15)
Schools with electricity	0.19	0.20	0.22	0.22	-0.05
	(0.32)	(0.27)	(0.42)	(0.41)	(0.13)
Correct percent of English	0.21	0.23	0.25	0.24	0.01
	(0.12)	(0.13)	(0.19)	(0.20)	(0.05)
Correct percent of Math	0.23	0.26	0.27	0.27	0.03
	(0.14)	(0.18)	(0.20)	(0.20)	(0.06)
N	25	13	73	78	189
Joint Chi-square					6.29

Notes: All variables measured at baseline. Columns 1-4 report the means for the Control Group Attritors, the MEET-E-Group Attritors, the Control and the MEET-E observations with complete data (surveyed). Standard deviations in parentheses. Last Column is the difference between the treatment group mean and the control group mean among the attritors minus the difference between the treatment group mean and the control group mean among those schools that reported data. Differences in means are computed by OLS controlling for district fixed effects. Standard errors in parentheses are clustered at the village level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The chi-square is the result of a test for individual coefficients to be jointly equal to zero using seemingly unrelated estimation.

Table A4-3: Analysis of Differential Attrition—SMS

	C-Attritor mean/sd	SMS-Attritor mean/sd	C-Non-attritor mean/sd	SMS-Non-attritor mean/sd	Diff-in-diffs coefficient/se
School Size (Enrollment)	85.31	69.63	81.71	82.87	-3.66
	(39.28)	(47.31)	(54.22)	(50.15)	(13.43)
Enrollment (boy)	55.99	48.30	52.57	52.11	3.54
	(42.78)	(45.19)	(51.61)	(41.16)	(11.75)
One Teacher School	0.58	0.57	0.51	0.40	0.02
	(0.33)	(0.21)	(0.50)	(0.49)	(0.13)
Total Teachers	1.94	2.01	2.26	2.36	0.02
	(0.76)	(0.50)	(1.91)	(2.04)	(0.48)
Schools affected by flood	0.68	0.52	0.66	0.63	-0.01
	(0.48)	(0.51)	(0.48)	(0.49)	(0.13)
Teacher Absence Rate	0.95	0.88	0.91	0.93	-0.06
	(0.10)	(0.22)	(0.19)	(0.14)	(0.05)
Number of classrooms	1.20	0.43	2.07	2.36	-0.75
	(1.53)	(1.12)	(1.47)	(1.70)	(0.40)
Toilets available	0.36	0.14	0.64	0.64	-0.16
	(0.49)	(0.36)	(0.48)	(0.48)	(0.13)
Schools with boundary wall	0.52	0.57	0.66	0.64	0.12
	(0.51)	(0.51)	(0.48)	(0.48)	(0.13)
Schools with drinking water source	0.44	0.50	0.52	0.54	0.05
	(0.39)	(0.27)	(0.50)	(0.50)	(0.13)
Schools with electricity	0.19	0.27	0.22	0.24	0.00
	(0.32)	(0.32)	(0.42)	(0.42)	(0.11)
Correct percent of English	0.21	0.24	0.25	0.25	0.01
	(0.12)	(0.12)	(0.19)	(0.18)	(0.05)
Correct percent of Math	0.23	0.26	0.27	0.27	0.01
	(0.14)	(0.11)	(0.20)	(0.19)	(0.05)
N	25	21	73	73	192
Joint Chi-square					14.00

Notes: All variables measured at baseline. Columns 1-4 report the means for the Control Group Attritors, the SMS-Group Attritors, the Control and the SMS observations with complete data (surveyed). Standard deviations in parentheses. Last Column is the difference between the treatment group mean and the control group mean among the attritors minus the difference between the treatment group mean and the control group mean among those schools that reported data. Differences in means are computed by OLS controlling for district fixed effects. Standard errors in parentheses are clustered at the village level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The chi-square is the result of a test for individual coefficients to be jointly equal to zero using seemingly unrelated estimation.

Table A4-4: Analysis of Differential Attrition—SMS-E

	C-Attritor mean/sd	SMS_E-Attritor mean/sd	C-Non-attritor mean/sd	SMS_E-Non-attritor mean/sd	Diff-in-diffs coefficient/se
School Size (Enrollment)	85.31	65.19	81.71	76.59	-2.58
	(39.28)	(27.02)	(54.22)	(35.87)	(13.87)
Enrollment (boy)	55.99	38.72	52.57	51.51	-8.45
	(42.78)	(22.62)	(51.61)	(39.21)	(12.14)
One Teacher School	0.58	0.59	0.51	0.39	0.07
	(0.33)	(0.23)	(0.50)	(0.48)	(0.13)
Total Teachers	1.94	1.98	2.26	2.31	-0.06
	(0.76)	(0.54)	(1.91)	(1.95)	(0.50)
Schools affected by flood	0.68	0.50	0.66	0.72	-0.14
	(0.48)	(0.51)	(0.48)	(0.45)	(0.13)
Teacher Absence Rate	0.95	0.90	0.91	0.92	-0.03
	(0.10)	(0.24)	(0.19)	(0.19)	(0.05)
Number of classrooms	1.20	0.89	2.07	2.05	0.17
	(1.53)	(1.13)	(1.47)	(1.19)	(0.42)
Toilets available	0.36	0.28	0.64	0.66	-0.02
	(0.49)	(0.46)	(0.48)	(0.48)	(0.13)
Schools with boundary wall	0.52	0.44	0.66	0.75	-0.20
	(0.51)	(0.51)	(0.48)	(0.43)	(0.13)
Schools with drinking water source	0.44	0.39	0.52	0.56	-0.15
	(0.39)	(0.32)	(0.50)	(0.49)	(0.13)
Schools with electricity	0.19	0.26	0.22	0.20	0.01
	(0.32)	(0.36)	(0.42)	(0.40)	(0.11)
Correct percent of English	0.21	0.23	0.25	0.27	-0.03
	(0.12)	(0.10)	(0.19)	(0.18)	(0.05)
Correct percent of Math	0.23	0.25	0.27	0.29	-0.03
	(0.14)	(0.09)	(0.20)	(0.20)	(0.05)
N	25	18	73	85	201
Joint Chi-square					8.29

Notes: All variables measured at baseline. Columns 1-4 report the means for the Control Group Attritors, the SMS-Group Attritors, the Control and the SMS observations with complete data (surveyed). Standard deviations in parentheses. Last Column is the difference between the treatment group mean and the control group mean among the attritors minus the difference between the treatment group mean and the control group mean among those schools that reported data. Differences in means are computed by OLS controlling for district fixed effects. Standard errors in parentheses are clustered at the village level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The chi-square is the result of a test for individual coefficients to be jointly equal to zero using seemingly unrelated estimation

Table A5: Impact on Students' Test Scores (Endline)

	English	Math
MEET	-0.16	0.01
	(0.109)	(0.10)
SMS	0.02	0.10
	(0.12)	(0.12)
MEET-E	-0.11	-0.07
	(0.13)	(0.11)
SMS-E	0.01	0.12
	(0.13)	(0.11)
Observations	4527	4527
R-squared	0.11	0.10
Control group mean	0.06	0.16
Equivalence Test:		
MEET = SMS (F-stat)	2.67	0.66
MEET-E = SMS-E (F-stat)	0.47	1.87

Sources: School Survey Endline (2015). Observations at student level. Robust standard errors in parenthesis. ***p<0.01, **p<0.05, *p<0.10. English and Math test scores are scaled by fitting two-parameter logistic item response model. Specialized tests were administered to Standard 3 and 4 students in class-room setting at the endline.

Table A6: Robustness: Impact on Teachers, Student Enrollment, and Infrastructure

	Total Teachers (Available)	Total Teachers (Present)	One Teacher School	Enrolment [Overall]	Enrolment [Boys]	Total Number of Classrooms	Total Number of Toilets
MEET	0.18	-0.04	**-0.13***	4.81	1.69	-0.08	0.04
	(0.17)	(0.17)	(0.07)	(5.86)	(4.15)	(0.13)	(0.14)
MEET-E	0.03	0.08	-0.06	4.13	4.74	0.20	0.12
	(0.15)	(0.16)	(0.07)	(5.06)	(3.96)	(0.14)	(0.17)
SMS	**0.31***	0.24	**-0.14****	**10.49***	**10.38****	-0.03	-0.03
	(0.17)	(0.19)	(0.07)	(5.59)	(4.73)	(0.13)	(0.16)
SMS-E	0.12	-0.03	-0.10	6.54	**7.26***	0.16	**0.31****
	(0.17)	(0.17)	(0.07)	(4.89)	(3.86)	(0.13)	(0.15)
Observations	367	367	367	367	367	367	364
R-squared	0.73	0.61	0.32	0.67	0.70	0.67	0.23
Mean of control	2.23	1.95	0.54	69.07	45.68	2.17	1.42

Notes: Standard errors clustered at village level in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Sample is restricted upon treatment schools located at least 1 kilometer away from control schools and covers all such schools that were open on the day of visit (announced round) and that completed all modules of the survey at both baseline and endline. Regression estimates control for district fixed effects and a flood dummy.

Figure A1: Intervention Design

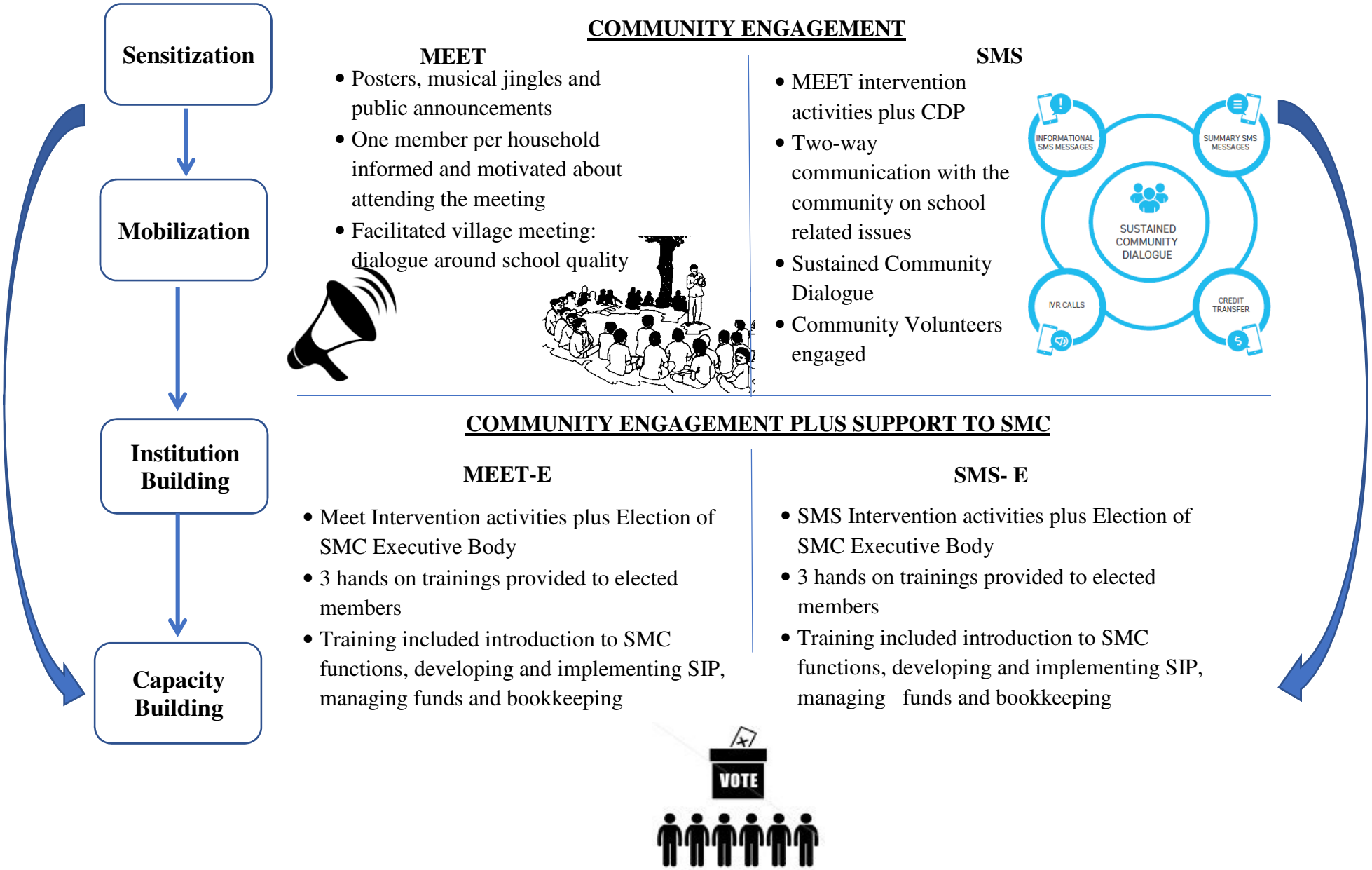
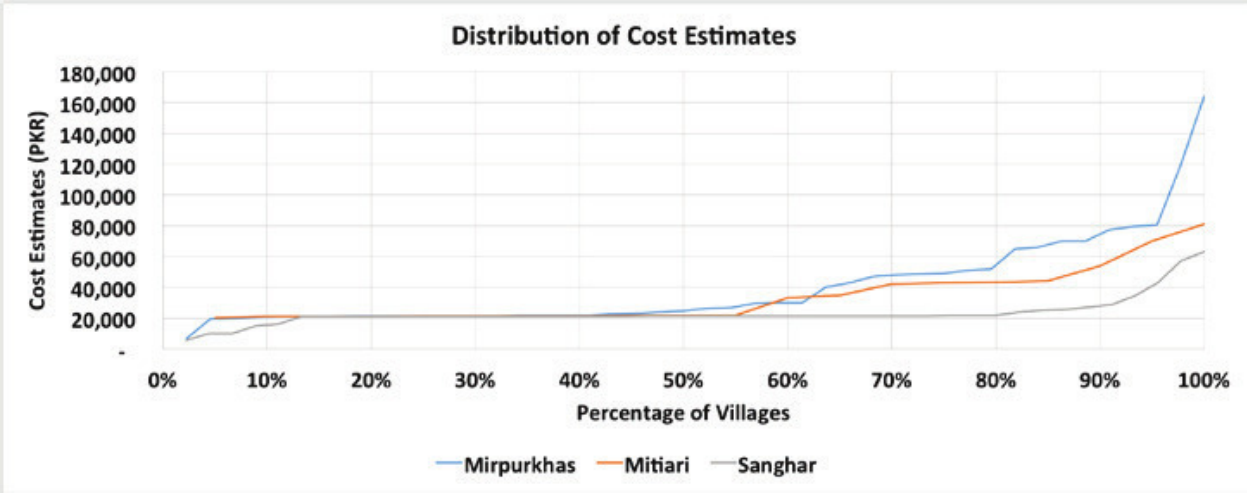


Figure A2: Sample CDP SMS messages

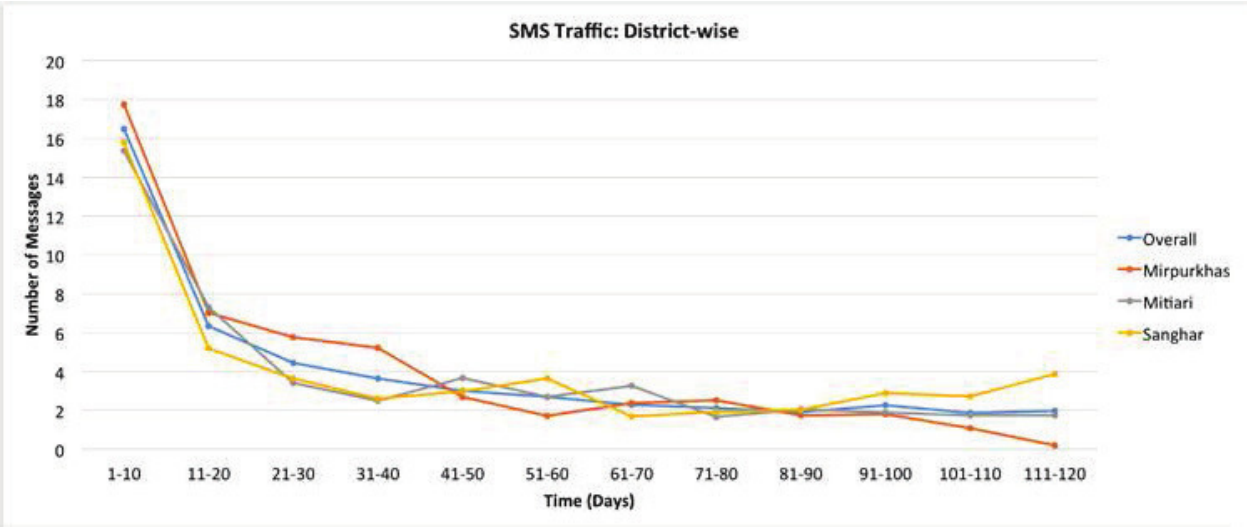
Round 1 messages: information on School-Based Management and SMC
Every year, the Government of Sindh provides all books free of charge to each student in primary school. These do not include notebooks.
Every year the School Management Committee of each primary school receives PKR 22,000, which goes into the Committee's account. This account can be used with the signatures of the headmaster or the chairperson. A School Management Committee's budget can be spent on three things: 1) facilities and maintenance of the school; 2) wages for workers to clean the school; and 3) transportation of students to and from school.
The Government of Sindh would like to inform you that there are two ways to vote in School Management Committee elections: by ballot or by a show of hands. All members of the General Body will vote for their preferred candidate. The candidate with the most votes for a seat will be elected to that seat. No member can be a close relative of another member. Keep in mind that in a girls school, at least three of the members should be female.
If there is a dearth of anything at the school, then it should be addressed through the following ways: take part in School Management Committee meetings; visit schools; and stay in regular contact with the members of the Executive Body."
The Government of Sindh would like to inform you that you should visit schools and check the following: Do the teachers arrive at school punctually and teach properly? Do the students know how to read and write according to their grade level? Are the school facilities usable or not? Are approved plans for improving the school being implemented properly? If you feel anything is lacking, then discuss it in the General Body meeting or contact the Executive Body.
You have been provided names and mobile numbers of all committee members. If you cannot meet them, then talk to them on the phone. Stay in regular contact with these members.
Round 2 messages: information on action steps to improve school
These steps can lead to an improvement in the school: • Exchange views via the SMS project. • Get involved in discussions with the Executive Body and teachers. • Take part in the planning of improvements in the School Improvement Plan or school. • Make sure that the expenditure of the School Management Committee's budget is according to approved School Improvement Plan.
The SMS project is a method for improving the school. Love of children, a call to knowledge.
You should send your opinions and recommendations via text message to the villagers.
A summary of the villagers' opinion is sent on a weekly basis to the villagers. The summary does not include anyone's name and your anonymity is guaranteed.
If you do not know how to read or write text messages, then meet with your village representatives [representatives' names]. These are their contact numbers: [representatives' contact information]. If these representatives do not help you, let us know.

Figure A3: Accumulated reserves of SIP Grant Funds, per village



Note: We report cost estimates from School Improvement Plans prepared by SMC members during capacity building under ELECTIONS-CAPACITY; participants were instructed to prepare SIPs to spend all accumulated reserves.

Figure A4: Average non-junk messages received by CDP, per village



Appendix B. Survey Instruments

In both the baseline and endline, the school surveys collected detailed data on school-level variables such as enrollment, attendance, teacher on-task, facilities, infrastructure, SMCs, funding and expenditure. In addition, student tests were administered to randomly selected students and teachers.[43]

The school observation questionnaire and the teacher roster questionnaires collected information on the school's functionality and conditions. The head teacher, teacher, and student questionnaire provided insights into school, teacher and student characteristics and perceptions.

In the school observation questionnaire, enumerators recorded their observations of school conditions, functionality and infrastructure[44]. This instrument gathered information on the school building, facilities and amenities, hygiene conditions inside and outside the school, on-going classroom practices and teacher activities.

The teacher roster gathered basic information on all teachers serving at the school. This data included information on the teachers' subject areas, salaries and their presence at school. The endline teacher roster collected information on all teachers registered with the school at the time of the baseline as well as all teachers registered at the time of the endline survey.

The head teacher questionnaire gathered information on the head teacher's personal and professional background, as well as his or her knowledge of students, facilities and SMCs. The second part of the instrument collected information from official school records. This information included details on the School Improvement Plan (SIP), attendance, fees, and SIP funds and expenditures. The teacher questionnaire was similar to the head teacher questionnaire – with the omission of the school records section – and was directed at other teachers besides the head teacher. It also collected personal and professional information on teachers as well as their perceptions about student learning and SMC functionality.

B.1 Test Development Process

The Learning Assessments was norm-referenced test designed to be administered to primary grade students. Each test had three components, Mathematics, English and Sindhi (the vernacular), aimed at measuring the child's performance across competencies. The test covered general range of content taught in primary grades in these subjects. The questions included anchor items, selected directly from standardized international tests, and questions adapted from textbooks taught in Sindh government schools.

The design of the tests was a five layered process:

1. Selection of Subjects and Content Domains
2. Compilation of Item bank
3. Compilation of tests A, B and C
4. Pilot of tests

[43] Student tests administered to teachers follow an SDI style knowledge assessment in which teachers are asked to mark a random student's completed exam.
[44] Observations include whether the school was open at the time of visit, the number of classrooms in use and whether the school had access to boundary wall, electricity and/or drinking water.

5. Compilation of final test

The description of these stages is as follows:

Stage 1: What to test

Selection of Subjects

The test aimed to deduce the literacy and numeracy skills gained by students in Government Primary Schools. For this purpose Mathematics, Sindhi and English were chosen as the testing subjects. The curriculum of these subjects is not only relatively standardized than other subjects across government schools but also introduces the students to varied logical and analytical thinking. Thus selection of these subjects did not only make the test comparable across schools but also allowed to test a range of cognitive abilities. Inclusion of Mathematics and English also makes the test comparable with international tests.

Selection of Content Domains: Similarity between National Curriculum and Sindh Textbook Board Books

The National Curriculum lists standards and related competency benchmarks that inform the learning expectations across grades for all subjects. These benchmarks are translated into content domains and learning outcomes for each grade. These domains then form the basis of textbooks.

An in depth analysis of the association between the National Curriculum (NC) and Sindh Textbook Board (STB) STB books for primary classes (Grades 1-5) revealed that the correlation between the two may not be as innate as was expected. The content in the STB books deviated from the learning outcomes outlined in the NC.

The deviations were most significant for the English language. The similarity between the learning outcomes of the NC and STB increased with increasing grades. This was largely because STB books delayed the introduction of concepts outlined in the NC till later grades and so while textbooks for lower grades might not have included concepts outlined in the NC, textbooks for grade 4 and 5 were largely similar to the curriculum. The similarity between the concepts introduced in the NC and STB books also differed across competencies. For example the learning outcomes overlapped to significant levels for concepts relating to Oral practice, however, there was no direct reference to the competency "Appropriate ethical and social development" in the STB books. The "Reading skills" remained somewhat similar across all grades while concepts for "language practice" and "writing skills" were only similar in books for grades 3 and onwards.

The account was drastically different for Math textbooks where the content not only conformed to the National Curriculum but also followed the same pattern of competencies. There were certain circumstances where the STB textbooks excluded the introduction of difficult concepts within competencies and the lag was maintained throughout subsequent grades. As a result of this, there were several concepts that remain left out throughout the primary grades such as percentages, distance and time. On the other hand some topics were introduced earlier in the STB books than required by the NC, such as currency, introducing the students to the very basic concepts before they are formally introduced to the core competency of the topic, e.g. adding and subtracting prices of objects using currency notes and coins. Both these shifts allowed for a smooth learning curve across grades unlike that projected by the English textbooks.

The topics chosen for the test and their grade wise distribution were adjusted according to the findings of NC and STBB comparison. The test only included those topics that were successfully mirrored in the

textbooks developed by the Sindh Textbook Board. Competencies that could not be tested in the field like use of protractor to measure an angle in Math and Oral communication and Ethical and Social Development in English and Sindhi were excluded from the test.

Stage 2: Item bank

Compilation

This stage included selection of Primary and Comparable questions for each selected topic.

The first preference for the Primary questions for Math and English was standardized International tests: TIMMS for Math, PIRLS for Language and another standardized test for Math and English developed and administered by Education Initiatives in India. All these tests are also primary level English and Math tests and are marked using a common achievement scale. However, these tests could not be directly replicated in Pakistan, as even though the National Curriculum does mirror most of the competencies required of grade 4 and grade 5 students compared to International Standards (competencies required in TIMMS and PIRLS) there are deviations from the NC in the STBB, as outlined in the previous stage.

Thus, an in-depth comparison of international test items with those in local textbooks was carried out: while an overlap in competencies over a large range of test items in standardized tests was found, this overlap was misleading. A small pre-test of some of the international items revealed that the approach used to ask questions in the international tests is very different from the way they are asked in students' textbooks. For this reason all international test items were categorized on levels of similarity with the questions in the textbooks and the items that were extremely similar or somewhat similar were chosen for the compilation of the tests while items that were not similar were disregarded.

The Item Bank further categorized all test items as either "primary questions" or "comparable questions". It consisted of clusters of questions for each selected competency with one primary question for each cluster and at least 2 comparable questions. All primary questions belonged to the "extremely similar" category of international test items (also referred to as anchor items) or in the absence of the same, they were taken from textbooks as is. On the other hand comparable questions are either "somewhat similar" questions from international tests or modified questions from STBB books.

For Sindhi section, primary questions were borrowed from The Learning and Educational Achievement in Punjab Schools (LEAPs) test. Similar pattern as that for compilation of English and Math item bank was followed. For topics not covered in LEAPs test, STBB questions were used.

The challenge that some competencies are introduced at a higher level without necessary building blocks in earlier grades was addressed through a weighting function with the number of questions in the item bank corresponding closely to the emphasis placed on that competency in the textbook across grades.

The item bank for all subjects included several questions belonging to same topic to allow for random selection at a later stage. The wide range of questions ensured coverage of the content taught across all chosen subjects. It was also made certain that these questions were of varying level of difficulty and tested different abilities. The types of questions included in the item bank were multiple-choice questions, matching pair questions, comprehension and short answer questions.

Testing

A small pre-pilot of the compiled item bank was then conducted in selected schools in one of the study districts of Sindh. The aim was to test student's familiarity with the topics included in the item bank,

type/level of assistance needed to understand the questions, test question format and to check if the questions were easy to implement in the field.

Using the results of the pre-pilot phase, the topics so far selected in the item bank were further refined. Any questions which were difficult to implement in the field (e.g. drawing a bar graph) or required too much assistance/explanation were dropped. The results of the per-pilot were also helpful in deciding the length of the final tests.

Stage 3: Compilation of tests: A, B and C

After the pre-pilot phase, all the remaining questions in the item bank were reviewed thoroughly and the modifications were incorporated (if needed) according to the observations from the field. It was ensured that the item bank had at least three questions for each topic so as to form three tests. These questions were then randomly assigned to Test A, B or C.

Based on the filed observations, it was also decided that the instructions for all the questions of Math and English will be translated into Sindhi and Urdu as well. An example was also to be inserted alongside each question of the test.

The distribution of sections for all tests was such that Math and English section contained 25 questions each and Sindhi contained 39 questions (counting all the parts) for the baseline test. It was decided to keep more questions in Sindhi section so as ensure complete coverage of curriculum topics.

Stage 4: Pilot of tests and IRT

The above steps enabled the design team to create tests based on questions that an average student in rural Sindh should be able to attempt. However, the success of questions of these tests rested on two assumptions: i) students are taught using textbooks in school and ii) students are able to follow the textbooks completely. The next stage in test development was to test these assumptions in the field and adjust the difficulty range of items based on test-runs of the items in pilot tests.

During the pilots it was found that although all students were being taught from Government textbooks but the learning level was much lower than expected from the mastery of the concepts in textbooks. Basically, in the field the second assumption broke down and hence the anticipated ability distribution for the group of students for whom the test was intended was not the same as reflected in the textbooks.

These tests were piloted in the study districts of Sindh and item indices were constructed to map the ability distribution of the intended population. Three parameter model for estimating the item response functions for our pilot tests was used. One desirable characteristic of test item was that high-ability candidates would answer it correctly more often than the low ability candidates—discriminatory power of the test item could easily be detected from the slope of the Item Characteristic Curve (ICC) fitted to the test data using item response functions—steeper the slope of an ICC, higher an item's discrimination value. The other important parameter was the difficulty level of the item: the more difficult the item is the higher the student's ability must be in order to answer it correctly. The ability-level of the examinee at the point where ICC curve was the steepest gave us the difficulty parameter for the test items. Items selected for our test were neither too easy not too hard given the ability distribution of the students. Mostly items that more than 80% of the examinees attempted correctly or less than 20% got right were the ones dropped out from the item bank(less than the guessing parameter). Further analysis using item response parameters to construct test information functions was desirable but could not be performed due to small sample sizes of pilot runs.

Stage 5: Compilation of Baseline test

As a result of the processes described above, the final test incorporated the following desirable features; (1) it covered competencies that were common across the National curriculum, international tests and most of all STBB books, (2) it reflected the weightage assigned to competencies across primary textbooks, (3) it had been adjusted to reflect actual learning levels of students in the study districts and (4) it incorporated sufficient anchor items to allow for a regional/international comparison of assessment results.

The final test was formed through random selection of items from either test A, B or C for each of the shortlisted topics. The final baseline test comprised of 25 English questions, 25 Mathematics questions and 39 Sindhi questions. The endline test contained 30 English questions, 30 Mathematics questions and 39 Sindhi questions. The number of Mathematics and English items was increased in endline test so as to allow for inclusion of additional anchor items to increase the comparability of the test with standardized International tests.

The grade and content weightage for the final tests was as follows:

Table B1: Distribution of Baseline Test Questions by Grade and Competency

Difficulty Level	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
Mathematics Competencies						
Number	1	0	2	0	0	3
Algebraic Operations	1	7	0	1	0	9
Measurement	1	1	3	0	0	5
Geometry	0	0	1	1	1	3
Information Handling	0	1	3	0	1	5
Total	**3**	**9**	**9**	**2**	**2**	**25**
English Competencies						
Reading Skills	0	0	5	0	0	5
Language Skills	2	4	4	1	3	14
Writing Skills	4	0		2	0	6
Total	**6**	**4**	**9**	**3**	**3**	**25**
Sindhi Competencies						
Reading Skills	0	0	0	4	0	4
Language Skills	1	17	6	2	4	30
Writing Skills	3	0	2	0	0	5
Total	**4**	**17**	**8**	**6**	**4**	**39**

Table B2: Distribution of Endline Test Questions by Grade and Competency

Difficulty Level	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Total
Mathematics Competencies						
Number	1	0	2	0	0	3
Algebraic Operations	2	7	1	3	0	13
Measurement	1	1	3	0	0	5
Geometry	1	0	1	1	1	4
Information Handling	0	1	3	0	1	5
Total	**5**	**9**	**10**	**4**	**3**	**30**
English Competencies						
Reading Skills	0	0	5	0	0	5
Language Skills	2	6	4	2	3	17
Writing Skills	4	2	0	2	0	8
Total	**6**	**8**	**9**	**4**	**3**	**30**
Sindhi Competencies						
Reading Skills	0	0	0	4	0	4
Language Skills	1	17	6	2	4	30
Writing Skills	3	0	2	0	0	5
Total	**4**	**17**	**8**	**6**	**4**	**39**

Table B3: Learning Assessments: School-Level Attempt Rates and Scores for English and Math Test Items

		Baseline			**Endline**		
		Attempt Rates	**Mean Score**	**Std Dev**	**Attempt Rates**	**Mean Score**	**Std Dev**
	English Items						
1	Missing Alphabets	0.49	0.83	0.37	0.87	0.80	0.40
2	Missing Alphabets	0.50	0.76	0.43	0.81	0.75	0.43
3	Capital and Small Letters	0.48	0.72	0.45	0.79	0.66	0.47
4	Capital and Small Letters	0.51	0.72	0.45	0.74	0.54	0.50
5	Concept of Vowels: using a an	0.44	0.62	0.49	0.71	0.71	0.46
6	Concept of Vowels: using a an	0.49	0.63	0.48	0.81	0.69	0.46
7	Verb: to be	-	-	-	0.71	0.45	0.50
8	Prepositions	0.48	0.59	0.49	0.82	0.43	0.50
9	Adjectives	0.49	0.51	0.50	0.78	0.42	0.49
10	Pronouns	0.44	0.67	0.47	0.80	0.59	0.49
11	Pronouns	0.43	0.52	0.50	0.77	0.56	0.50
12	Verb: to have	0.44	0.46	0.50	0.67	0.69	0.46
13	Question Words	0.43	0.23	0.42	0.63	0.28	0.45
14	Spelling words	-	-	-	0.56	0.83	0.38
15	Spelling words	-	-	-	0.52	0.80	0.40
16	Prepositions	-	-	-	0.60	0.55	0.50
17	Antonyms	0.45	0.56	0.50	0.71	0.47	0.50
18	Past tense	0.39	0.39	0.49	0.69	0.43	0.50
19	Irregular Plurals	0.45	0.66	0.47	0.65	0.51	0.50
20	Irregular Plurals	0.42	0.55	0.50	0.62	0.49	0.50
21	Days of the week	0.46	0.68	0.47	0.67	0.60	0.49
22	Picture recognition	0.39	0.68	0.47	0.54	0.57	0.50
23	Sentence formation	0.15	0.47	0.50	0.21	0.46	0.50
24	Spelling words	0.38	0.41	0.49	0.65	0.39	0.49
25	Adjectives	-	-	-	0.62	0.44	0.50

26	Comprehension	0.47	0.59	0.49	0.65	0.52	0.50
27	Comprehension	0.42	0.72	0.45	0.61	0.48	0.50
28	Comprehension	0.43	0.58	0.49	0.59	0.45	0.50
29	Comprehension	0.40	0.45	0.50	0.54	0.41	0.49
30	Comprehension	0.42	0.67	0.47	0.55	0.45	0.50
	Math Items						
1	Numbers: Before and After	0.49	0.69	0.46	0.78	0.79	0.40
2	Place Value	0.42	0.34	0.47	0.63	0.36	0.48
3	Subtraction: 1 Digit	0.56	0.83	0.38	0.86	0.82	0.39
4	Subtraction: 2 Digit	0.54	0.63	0.48	0.84	0.73	0.44
5	Addition: 2 Digit	0.52	0.68	0.47	0.82	0.69	0.46
6	Subtraction: 3 Digit	-	-	-	0.80	0.64	0.48
7	Smallest and Greatest 4 digit Numbers	0.43	0.51	0.50	0.76	0.62	0.49
8	Addition/Subtraction of Units of Mass	0.42	0.56	0.50	0.74	0.65	0.48
9	Multiplication	0.46	0.73	0.44	0.78	0.68	0.47
10	Division	0.42	0.59	0.49	0.75	0.46	0.50
11	Division	-	-	-	0.70	0.51	0.50
12	Geometric shapes	-	-	-	0.63	0.62	0.49
13	Fractions: One third and Two thirds	0.45	0.63	0.48	0.72	0.54	0.50
14	Fractions: One third and Two thirds	0.46	0.62	0.49	0.70	0.65	0.48
15	Fractions: Halves	0.42	0.58	0.49	0.72	0.44	0.50
16	Multiples	0.37	0.62	0.49	0.66	0.63	0.48
17	Mutual Conversion of Units of Weight	0.43	0.59	0.49	0.69	0.52	0.50
18	Mutual Conversion of Units of Time	0.45	0.46	0.50	0.69	0.59	0.49
19	Measuring Line Segment	0.47	0.67	0.47	0.68	0.60	0.49
20	Calculating Perimeter	0.39	0.55	0.50	0.63	0.63	0.48
21	Parallel and Perpendicular Lines	0.44	0.63	0.48	0.67	0.74	0.44
22	Calendar months	0.47	0.76	0.43	0.64	0.65	0.48
23	Addition/Subtraction of Units of Time	0.09	0.07	0.25	0.29	0.77	0.42
24	Addition/Subtraction of Units of Currency	0.61	0.78	0.42	0.72	0.83	0.37
25	Reading a Bar Graph	0.49	0.27	0.45	0.73	0.62	0.48
26	Unitary method	0.44	0.35	0.48	0.69	0.73	0.45
27	Fractions: Subtraction	-	-	-	0.48	0.36	0.48
28	Decimals: Addition	-	-	-	0.41	0.42	0.49
29	Reading a Bar Graph	0.45	0.60	0.49	0.60	0.39	0.49
30	Reading a Bar Graph	0.44	0.33	0.47	0.58	0.52	0.50