Voice and Action: Can Communities Change Education Outcomes? Evidence from a Field Experiment in Pakistan

Salman Asim[¥]
Amina Riaz[*]
DRAFT: NOT FOR CITATION

Abstract

We present findings from a large-scale experiment in rural Sindh, Pakistan. We test three approaches to catalyze engagement of communities with schools: face-to-face dialogue at community meetings; dialogue via text messages; and elections and capacity-building of school committees, in a cross-over design. We find dialogue both in meetings and via text messages to be effective in amplifying communities' voice resulting in improvements in access to school, in particular school functioning and staffing. Text message-based dialogue additionally improved school enrollment, suggesting gains in demand for education. The addition of elections and capacity building achieved improvements in school infrastructure and the number of new admissions, suggesting increased effort by committee members. However, we find that elections also undermined community participation in dialogue and reduced impacts on school functioning and staffing. The findings suggest that even in conditions with high barriers to participation, community engagement, particularly through ICT, can improve access to schooling; but that support to official institutions of collective decision-making may undermine efforts to bolster community voice.

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

---

[*]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; Dhushyanth Raju, Nazmul Chaudhury, Ravinder Casley Gera and Syeda Shahbano Ijaz for substantive technical discussions and edits. 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 and Abhijit Banerjee provided outstanding research assistance at various stages of the project. The authors thank Alaka Holla, Juliea M. Trias, Deon Filmer, Owen Ozier, Achim Schmillen and David Evans for helpful discussions and insightful comments. This paper is dedicated to the living memory of Mr. Muhammad Nouman Bashir, Managing Partner of Weitek Group, the implementation support firm, who contracted a fatal virus during the rollout of the project. We thank discussant and participants at the 2016 RISE Annual Conference at the Blavatnik School of Government, University of Oxford. We gratefully acknowledge the financial support provided for this project by the Strategic Impact Evaluation Fund (SIEF), the Education Sector Development Program and the World Bank. 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.
¥ Corresponding author: Salman Asim, Economist, World Bank; email: sasim@worldbank.org;
* Graduate student at the University of Chicago.

1. Introduction

Despite widespread reforms and significant investment by governments in basic education, an estimated 61 million primary school age children do not attend school (World Bank, 2018). Out-of-school children and high rates of dropout reflect widespread problems in school service delivery, ranging from understaffing and lack of inputs to poor teacher knowledge and practices.

School Based Management (SBM) is a decentralization reform that has become increasingly popular to address management weaknesses in school systems (Bruns et al., 2011). In conditions where democratic politics fails to provide adequate incentives for improvement in services, direct engagement by users – citizens – can provide an alternative route for accountability (World Bank, 2004). In developed countries, SBM gives greater management control to teachers and parents over what goes on inside the school; in developing countries, the reforms are typically less ambitious, involving parents and community members in the decision-making process rather than placing them directly in charge of schools (Barrera-Osorio et al., 2009). SBM projects typically work through a School Management Committee (SMC) comprising of the head masters, teachers, parents and other community members. Such institutions are intended to empower communities and amplify voice in school decision-making processes, hold service providers accountable, and facilitate collective action by the community, including resource mobilization (Santibañez, 2006; Bruns et al., 2011).

SBM reforms typically have three aspects: information, voice and collective action. First, communities may receive information on service standards, users' rights and entitlements, or the current state of service delivery. By doing this, these interventions seek to address asymmetries of information between service providers and users which can disguise underperformance or prevent citizens' engagement (Keefer, 2007). Second, most mechanisms of community engagement also entail the facilitation of user voice through meetings, votes or representative bodies. Citizens may share their preferences for service delivery, helping to ensure that priorities of service providers are aligned with those of end-users. Citizens may also monitor use of resources and performance, drawing attention to problems of delivery which are not being attended to by providers. Third, SBM reforms may incorporate an element of collective action, directly by community members or through a representative committee; often contributing to decision-making regarding the use of resources (Mansuri and Rao, 2013). Through these channels, it is expected that engagement can increase the *allocative efficiency* of public service systems, the extent to which resource allocations reflect user preferences; and the *technical efficiency*, the extent to which resources are used effectively to improve intended outcomes (ibid.).

In practice, however, developing countries have found it challenging to engage communities in meaningful ways (Joshi, 2013). There are several barriers to participation in public services, including low awareness, high opportunity costs, and cultural and social restrictions (ibid.; Platteau and Abraham, 2002; De Grauwe, 2005). In addition, in contexts where trust between citizens and state is low, citizens may opt out of engagement rather than engage with institutions they perceive as lacking legitimacy (Croke et al., 2015). In education, SMCs have not always proven capable of organically emerging as credible channels of collective action and in many cases suffer from low engagement and elite capture (Barrera-Osorio et al., 2009; Bruns et al,, 2011).

A wide range of frameworks and projects have attempted to identify means through which institutions of engagement can be strengthened (Ostrom, 1990; Menocal and Sharma, 2008; Heller and Rao, 2015). These have typically taken two main forms. First, some effort is typically made to bolster the legal and fiscal authority of institutions, more deeply embedding them in wider institutional frameworks; in Sindh, as in many SMC systems, this involved providing fiscal resources to be controlled by SMCs as an incentive for engagement. Second, various means have been tried to directly boost citizen engagement with those institutions, typically through the provision of information to communities on how to fulfill the roles established for them by the institution in question (Pandey, 2009; Banerjee et al., 2010). In certain developing country contexts, more success has been achieved by efforts to facilitate community engagement through participatory means, for example through actively supporting communities to exercise voice or fulfill duties of collective decision-making (Barr et al., 2012; Pradhan et al., 2014).

The context for this study is Sindh, Pakistan's second-largest province. Although a dense network of public schools exists in the province, many lack basic facilities and adequate numbers of teachers and are prone to temporary closures, and a large number of children are out of school, while learning outcomes are poor and not improving (ASER 2015). Feudal landholding patterns and high levels of ethnic and caste fragmentation in Sindh create powerful barriers to participation (Miguel and Gugerty, 2005; Hussain et al., 2013). SMCs were introduced in all government schools in 2000, comprised of parents and non-parent community members; however, these remained largely non-functional until 2006, when they were given control of annual grants worth around $200 (PKR 22,000). The grants succeeded in reactivating dormant SMCs, but the institutional structures remained weak, with limited engagement and widespread elite capture. Community-wide consultations are rarely held, SMC members are seldom elected, and cases have been reported of collusion between head teachers and community members with strong ties to elites. As a result, SMC grant funds have been widely mis- and-under-utilized.

We test three strategies to address these challenges by facilitating community engagement in school management. First, we carry out externally-facilitated face-to-face meetings to catalyze engagement with SMCs through provision of information and moderated dialogue on education and school quality. Second, we pilot an SMS-based dialogue platform, intended to provide communities with an anonymous, low-cost, and convenient means of communicating preferences and coordinating action regarding their local school. Third, we facilitate transparent elections and capacity building support for SMCs to align committees with preferences of communities, address elite capture, and strengthen collective action, including utilization of grants. We employ a cross-over design, with the elections and capacity building intervention acting as an additional treatment in a subset of schools receiving each of the dialogue interventions.

The first intervention, MEET, was designed to catalyze dormant institutions of engagement through external facilitation. In a community-wide meeting, participants were informed about the roles and responsibilities of SMCs, given a space to discuss issues related to school performance through an externally facilitated discussion, and encouraged to arrange follow-up meetings and to contact SMC members.[1] Evidence from countries including Uganda and Indonesia suggests that

[1] Under the Sindh system, the wider school community is considered part of an SMC General Body, which has a deliberative and ratifying role; while a five-person elected Executive Body carries out the majority of duties. We use *SMC members* to refer to members of the Executive Body.

such facilitation can stimulate dialogue, lower barriers to further participation, and lead to improvements in school outcomes (Barr et al., 2012; Pradhan et al,, 2014; Cerdan-Infantes and Filmer, 2015).

However, evidence from India suggests that in contexts where barriers to participation are high and trust in government is low, the benefits of facilitation are contingent on institutional structure. 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, it had a significant impact on teachers' attendance and on student learning outcomes (Pandey et al., 2009). In the context of Sindh, where both governance failures and barriers to engagement are particularly severe, effectively boosting engagement may require new approaches which cut across power structures and significantly lower the costs of exercising voice and collective action.

Our second intervention, SMS, tests an ICT-driven approach, in the form of a text message-based community dialogue platform (CDP). The specially-developed platform is designed to act as a virtual space for communities to receive and deliberate on information and voice their preferences on the most salient issues that affect their local schools. The SMS intervention was intended to reduce the cost of participation, by enabling community members to participate virtually at a time of their convenience; and to reduce the risks associated with participation, by maintaining the anonymity of participants, enabling participation by individuals who were otherwise unable to voice their preferences within the prevalent power structures. At the same time, the intervention facilitated accountability by sending regular summary messages, keeping parents informed of their collective preferences and also of the issues that the SMCs subsequently chose to address. It thus created an ICT-based interface between the community and the SMC. While SMS is growing in popularity as a mechanism for collection of data on school service delivery by governments, its use as a tool for community engagement remains at an experimental stage (Abdul-Hamid, 2017); however, the intervention does build on experimental evidence from Indonesia, where Cerdan-Infantes and Filmer (2015) evaluate the impacts of informing communities about a school management community engagement program through a range of modalities. They found that while providing information through written materials was ineffective, engaging parents through village meetings and through text messages increased awareness of the program and participation significantly. The CDP approach takes this one step further by combining both information provision and elicitation of preferences within a single platform.

Our third intervention, ELECTIONS, shifts the emphasis from community dialogue to collective action and decision-making. First, it supports fresh and transparent elections of SMC members, to address elite capture and improve alignment between members and the community. Evidence from India suggests that increasing the representativeness of local decision-making bodies is effective in aligning the bodies' decision-making with the preferences of the community (Chattopadhyay and Duflo, 2004), while a study in Indonesia found that aligning SMCs and the community through elections, as well as building linkages with community leaders, led to greater engagement by education stakeholders and improved learning outcomes (Pradhan et al., 2014). Second, to address problems of capacity and experience, the intervention provides elected members with participatory capacity building to develop plans for school improvement. Experimental evidence from Uganda

finds that training SMCs to fulfill their roles has positive impacts on student test scores, as well as student and teacher absenteeism, only where information provision is combined with participatory capacity building approaches (Barr et al, 2012).

We find evidence of impact for the meeting-based facilitation intervention on access to schooling, and greater impact for the SMS platform. Schools participating in the meeting-based intervention were more likely to be open at endline than those in the control group, and less likely to have ceased functioning between baseline and endline. Schools participating in the SMS intervention were more likely than control schools to have reopened if closed at baseline, and had significantly more students and teachers. However, we do not observe significant impacts from either intervention on school infrastructure, in part as a result of limited power and the difficulty of measuring the wide range of potential impacts in a diverse population of schools.

We do find significant impacts of elections and capacity building of SMCs on school infrastructure, as well as early-grade enrollment, when carried out in combination with the SMS platform; we also observe small impacts on learning outcomes. Unexpectedly, the addition of the elections and capacity building intervention reduced both the impact of the meeting-based intervention on school functioning, and the impact of the SMS intervention on enrollment and teachers. The holding of elections appears to have reduced the level of community participation in meetings; the most likely explanation, given the low level of trust in government in the Sindh context, is that the presence of a government education officer, intended to legitimize elections, reduced the perceived independence of the interventions.

The results suggest that, in conditions where cultural and social barriers to participation are large, ICT-based approaches, and specifically SMS-based dialogue, can achieve greater participation and improvements in school access. Direct support to representative decision-making bodies can improve their fulfilment of official duties, including utilization of school funds in alignment with community preferences; however, in conditions of low trust in government, trade-offs may exist between maximizing community engagement and building linkages between these representative bodies and government. Trust is developed through repeated interactions involving reciprocal obligations among community members and with government officials; collective action through official channels is unlikely to emerge in conditions where communities have limited trust in government officials. The findings also reinforce experiences from similar interventions in other countries that suggest that interventions to bolster community engagement in school management may struggle to raise learning outcomes without directly supporting engagement between parents and teachers.

The rest of this paper is structured as follows. Section 2 describes the study design, including the context of rural Sindh, the sampling framework, treatments, and study timeline. Section 3 presents the data strategy, outcomes of interest, descriptive statistics, and estimation strategy. Section 4 presents findings, including balance checks, treatment fidelity and intent-to-treat impacts. Section 5 discusses the findings. Section 6 presents analysis of cost-effectiveness, conclusions, and directions for future research.

2. Experiment

2.1 Context

Sindh, Pakistan's second largest province, provides a particularly informative environment to test the impact of interventions to strengthen community engagement, owing to widespread service failures in education and a local political economy which creates large barriers to participation in community engagement institutions. Sindh has a largely feudal political economy in rural areas: 76% of rural Sindhi families are landless and work as tenants, while an elite constituting less than one percent of rural families own land holdings of 50 acres or more (Bengali, 2015). These large-scale landlords often dominate economic and political life, exercising control over land, irrigation and credit, and many families are bound to their landlord by generations of debt (Hussain et al., 2013). Sindh is also subject to a particularly large degree of ethnic and caste fragmentation. Villages function as the lowest administrative units but typically consist of a collection of spatially dispersed settlements, *paros*, each primarily representing a particular 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). Public primary schools typically feed 1-2 settlements, with a larger school often located in the main settlement, the traditional hub of the village. In **Figure 1**, we illustrate some typical characteristics of a sample village, Seri, in rural Sindh. Seri contains residents from 11 different castes living in 12 settlements. These settlements are served by five officially mandated public primary schools. Such a dense distribution of schools reflects, in part, the depth of social fragmentation: most settlements contain households of one or two castes, and the social barriers against sending children, particularly girls, to a school outside their home settlement are high (Jacoby and Mansuri, 2011).

[Figure 1 about here]

Despite this dense distribution of schools, however, access to schooling in Sindh remains low. In 2012, according to the official Annual School Census, 14 percent of Sindh's officially established primary schools were non-functional, having been closed for more than six months or having no teachers or students registered (Asim, 2013). Of the remaining schools, 54 percent had only one teacher, leading to multiple-grade classrooms and creating a high risk of temporary closure.

Those schools which are functional lack important resources and amenities. 17 percent of schools in 2012 had no building, with classes taking place outdoors, in a hut or temporary structure. Only 58 percent had a toilet, 51 percent had drinking water and 40 percent had electricity. A large share of schools, although functional, lack all basic facilities: 28 percent of functional schools, more than 10,000 schools, either lack a school building or have neither drinking water, electricity, a toilet nor a boundary wall. Of these schools, the majority have only one teacher.

These service failures reflect the role of schools 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 other purposes

such as acting as makeshift accommodations or communal nighttime gathering places, with little investment to convert them into functioning centers of learning. Teaching positions are frequently created at the behest of local influential figures to provide secure income to trusted political allies, who act as custodians to sustain the existing power dynamics in the community; these politically-appointed teachers are largely immune from discipline and rarely attend schools (Babur, 2016). The provincial government estimates that 60,000 teachers in Sindh are regularly absent from school (Naviwala, 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.

As a result of these functional weaknesses, despite the dense distribution of schools, enrollment in Sindh remains low. The net enrollment rate (NER) at primary level (ages 6 to 10) in rural Sindh is just 52 percent, compared to 63 percent in all of rural Pakistan. At middle-school level (ages 11 to 13), the NER for rural Sindh falls to just 36 percent, versus 55 percent for all of rural Pakistan (PSLM 2014-2015).

Furthermore, for children who do attend, learning outcomes are poor. At Grade 3, only 24 percent of students can read words while only 32 percent can perform basic subtraction; at Grade 5, 19 percent of students can read full sentences while 33 percent of Grade 5 students can perform division (ASER 2015). Gender disparities in learning are large and consistent: for both English and Math, boys outperform girls by 6 percentage points.

During the period of this study, the situation of basic education in Sindh was exacerbated by extensive flooding in both 2010 and 2011. The 2010 floods, which affected much of Pakistan, destroyed an estimated nine percent of Sindh's educational facilities and damaged a further 19 percent, as well as destroying the homes or severely disrupting the livelihoods of around 865,000 Sindh households (Asian Development Bank and World Bank, 2011). Further heavy rains in 2011, which occurred primarily in Sindh, destroyed or damaged a further ten percent of the province's educational facilities, and destroyed the homes or severely disrupted the livelihoods of approximately 810,000 households (Asian Development Bank and World Bank, 2012). Schools which were not physically damaged were frequently employed as shelters for displaced people, disrupting teaching for several weeks. Overall, sixty-six percent of schools in the villages sampled for this study reported being affected by the floods in some way at baseline.

2.1.1 School-based Management in Sindh

The weak state of the school system in Sindh provides an example of a context in which community engagement in schools is needed to strengthen both access to schooling and the quality of service delivery. However, the same feudal political economy that weakens service delivery also acts as a constraint on public participation, as tenants are unsurprisingly hesitant to voice concerns or challenge the preferences of landlords and other elites.

Previous reforms have attempted to bolster community engagement in schools, with limited success. Sindh introduced SMCs in the 2000s as part of an agenda of decentralization and community engagement in public services. However, these first SMCs had little formal control over school decision-making or expenditure and in practice remained largely dormant.

In 2006, supported by the World Bank, the Government of Sindh attempted to revitalize SMCs as a formal channel for local communities to engage with schools. Following the reforms, SMCs receive annual grants from the Government worth PKR 22,000 (approximately USD 200) to partake in school improvement activities. These grants were intended to motivate communities to engage with schools and participate in school improvement processes. The SMC is given control over its allocated funds and is empowered to withdraw these resources to implement activities, but activities are expected to be carried out in accordance with a School Improvement Plan (SIP) developed in consultation with the wider community. Each SMC consists of a five-person Executive Body, which has the overall responsibility for preparing the SIP; and a larger General Body, consisting of all parents of students, which approves SIPs on an annual basis. SMC Executive Body includes the school's head teacher, as well as four elected community members, two of which are parents and two non-parents. In addition to the preparation of the SIP, the SMC Executive Body also monitors implementation and procurements, and carries out repairs and maintenance of school facilities. They are also responsible for monitoring enrollments and ensuring that children in their village enter school at the appropriate age, as well as monitoring of teacher attendance.

The reforms were successful in ensuring that funds were transferred to schools annually; as of June 2012, over 81 percent of SMC bank accounts were active and receiving grants from the Government. However, actual engagement and activity remained low. Evidence from the baseline survey carried out in three representative districts of Sindh as part of this study revealed that in 2012, according to head teachers, while 91 percent of schools had an SMC fewer than half of these (45 percent) had held a Executive Body meeting in the previous six to twelve months, while fewer than a quarter (21 percent) had held a General Body meeting during the same time. Only 29 percent of head teachers said that Executive Body members at their school were chosen by election. Gender disparities in participation were large and persistent: only 15 percent of head teachers thought it was common to hear women speak at General Body meetings, while 66 percent thought the same for men. Only 33 percent of head teachers believed that women are allowed to join the SMC Executive Body.[2]

While the reforms ensured that SMC grant funds reached school bank accounts, a large proportion simply went unutilized: In two of the three study districts, Mirpurkas and Matiari, 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 unutilized funds.[3]

2.2 Potential impacts of community engagement

Effective community engagement in schools can affect a wide range of aspects of school functioning, capacity, and outcomes. In terms of school functioning, evidence from a range of

[2] We did not observe any positive and significant impacts on these outcomes from the interventions, and therefore endline findings for these are omitted for brevity.

[3] In 2013, as the result of the evidence generated from the baseline surveys for this study, the Government of Sindh revamped the scope of the use of SMC 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 SMC funds in delinquent accounts and only provide grants to schools with active SMC members and functional accounts.

evaluations suggests that systems of community monitoring can increase teacher effort. Such impacts can occur through two mechanisms: direct effects of community monitoring on teachers, who alter behavior in response to social sanction or the threat thereof; and indirect effects through better management of teachers by headteachers and officials in response to community pressure. Improvements in teacher effort have been observed in the wake of SBM reforms in a wide range of countries, including El Salvador (Sawada and Ragatz, 2005) and Guatemala (di Gropello, 2006). More recent studies have provided experimental evidence of impacts of community engagement on teacher attendance specifically: in India, the provision of information to communities on institutions of engagement in school management succeeded in raising teacher attendance (Pandey et al., 2009); similar effects were observed in Uganda and Gambia from programs which provided participatory capacity building for community members (Barr et al., 2012; Blimpo and Evans, 2011). In the context of Sindh, where a large proportion of schools have only one teacher and teacher absence is a key driver of school closure, it is likely that mere availability of teachers in schools could lead directly to improvements in the number of schools which are open and functioning.

There is also some evidence of impacts on school facilities. Although school conditions may be primarily limited by the availability of resources from government, evidence from a range of countries suggests that effective community monitoring and collective action can address these constraints through a range of channels: better targeting and utilization of existing government finance, such as Sindh's SMC grants; successful lobbying of government for additional resources; monitoring of existing finance to prevent loss or misappropriation; and resource mobilization from within the community. In Uganda, publication of information on school grants led to a large increase in the proportion of finance reaching schools (Reinikka and Svensson, 2005); in Indonesia, provision of information to communities on their roles and responsibilities in school management led to increased transparency by schools on the use of grant funds, while SMS-based provision of similar information led to an increase in parental contributions to schools (Cerdan-Infantes and Filmer, 2015).

Community engagement also has potential impacts on education demand, as well as supply. Engagement mechanisms can raise community awareness of the benefits of education, and improvements in school quality can lead to increased demand as a result of growing parental confidence in the value of schooling. In addition, in some cases (including in Sindh), monitoring and bringing out-of-school children to school, dropout and absence are explicitly included in the responsibilities of SMCs. As a result, SBM systems have been associated with improvements in student enrollment and reductions in dropout in countries including Brazil and Mexico (Paes de Barros and Mendonça, 1998; Gertler, Rubio-Codina, and Patrinos, 2006). Experimental evidence suggests that facilitation of community engagement has led to improvements in student enrollment and attendance in countries including Uganda and Gambia (Barr et al., 2012; Blimpo and Evans, 2011).

Ultimately, through these improvements in school functioning, teaching practices, and demand for education, effective community engagement can in some instances improve learning outcomes. Experimental evidence from Uganda, India and Indonesia suggests that facilitation of community engagement can have some positive impacts on test scores (Barr et al., 2012, Pandey et al., 2009, Pradhan et al., 2012). However, a large number of SBM reforms have failed to improve learning

outcomes, even in cases where impacts have been observed on intermediate outcomes such as school inputs (see Snilstveit et al., 2016, for a review). Although providing grants for discretionary spending by schools has typically been associated with increased enrollment, there is very little rigorous evidence of impacts on learning (Evans and Popova, 2015). This reflects a wider trend whereby, in general, improvements in quantity of schooling in developing countries have been found to inconsistently generate improvements in learning (Pritchett, 2013; Bold et al., 2018). In cases where impacts have been observed on learning, these have typically taken several years to emerge (Bruns et al., 2011).

A range of factors can prevent schooling from translating into learning, including low teacher subject knowledge, poor use of teaching time, and outdated and inadequately learner-centered teaching practices, particularly in a context such as Sindh where teacher placements act as a form of patronage and teachers frequently lack skills and motivation (World Bank, 2018). Interventions which focus on improving direct engagement between parents and teachers, as distinct from parental participation in school management, may be more effective in generating changes in teacher behavior and improving learning outcomes (Islam, 2017; Asim et al., 2017). However, improvements in learning can also be constrained by systemic factors outside of the control of schools, such as overambitious curricula which leave students struggling to learn advanced material who have yet to master basic concepts (Pritchett and Beatty, 2012).

2.3 Sample

Sampling was carried out according to three key goals: the districts should be representative of the overall education profile of rural Sindh; there should be sufficient number of schools in these districts to meet the power requirements of our analysis; and enumerators should be able to canvass these districts without excessive logistical problems and security risks. Using the Pakistan Social and Living Standards Measurement (PSLM) survey, we ranked all districts of Sindh by: i) proportion of adults that ever attended school, and ii) school attendance rates of primary-age children (5-12 years). Also, using administrative school census data, we ranked districts on 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 and one small district, ensuring representativeness of the selected districts along the two education measures,[4] as well as ensuring the safety of field teams. The selected districts were Matiari, Mirpurkhas and Sanghar.

In the absence of an updated population census, administrative data—the Sindh Annual School Census (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.[5] 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

[4] 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, district Matiari was in the middle, while Sanghar was among the highest. They were deemed relatively safe by the Government for field teams to visit.

[5] 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.

schools drawn from 377 villages. The administrative data appeared to have problems of noise and unreliability with regard to the exact location of schools; to address this, we mapped 300 of the 377 villages in a village-census exercise. The cap of 300 villages was reached using a randomized ranking procedure.[6] The process through which treatment schools were selected from the village mapping exercise is outlined in **Figure 2.**

[Figure 2 about here]

Following mapping, we conducted village-level focus group questionnaires to obtain qualitative information on village characteristics to assist the intervention firm in targeting the right nodes for organizing village level meetings, as well as determining the main settlement in the village, historically regarded as the central hub of village activity (typically the largest settlement). This was selected as the location of the village-level meeting in which we conducted introductory activities for the various interventions. 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 covering a total of 299 schools from 224 villages. In addition, we also covered 15 percent of schools, 54 schools, in settlements other than the main settlement within each village. In villages that did not have any school in the main settlement, a maximum of three schools were surveyed from other settlements based on their total enrollment. In total we added another 132 schools from 63 villages to our sample. Further, there were 9 villages that did not have either a functional or temporarily closed school in any of the village's settlements; for these, we included all schools, 16 such schools, even if they were closed for a period of more than one year. Four villages had to be dropped because no school was found during the village-level census mapping of primary schools. Our sample then consisted of 501 schools across 296 villages. The resulting sample is representative of both the three treatment districts, and of rural Sindh in general, across a range of key outcomes; the sample differs significantly from treatment districts in one indicator, number of female teachers per school, and from all of rural Sindh in two, school enrollment and drinking water.[7]

At the time of data collection, the baseline field team faced political resistance and was denied entry in four of our sample villages. This reduced our final baseline sample to 292 villages. Afterwards, extensive quality checks were performed on the baseline data – five villages that did not meet the benchmark due to poor quality or incomplete surveys were dropped from the study, bringing the overall sample to 287 villages, containing 489 schools. Three villages were later dropped from the analysis owing to issues of data quality at endline, leaving a final sample of 284 villages and 479 schools. Of these, 387 schools were consistently open at baseline and endline. For the majority of outcomes of interest, measurement was only possible at these consistently functioning schools and therefore analysis is based on a sample of 387 schools. We have also excluded schools from analysis that switched status from closed to open at endline, as the

[6] 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.

[7] See Table A2, in Appendix A.

characteristics of these school status are largely different from schools that were functional both at baseline and endline.[8]

2.4 Treatments

The evaluation follows a clustered randomized control trial (RCT) design in order to causally determine the impact of three interventions on school-level outcomes. We randomly assign villages to one of the treatment groups or to control, and then cluster all sampled schools within each village to the treatment groups. The unit of randomization is the village and the school is the unit of inference.

The first two interventions, MEET and SMS, test alternative modalities of facilitation of community dialogue and collective action around education, employing both face-to-face and ICT-based approaches. The third intervention, ELECTIONS, tests the additional impact of increasing linkages between communities and SMCs through the holding of elections for Executive Body members, and raising the capacity of SMCs for collective action by providing participatory capacity building for Executive Body members. In a cross-over design, the third intervention combines each of the dialogue interventions in a randomly selected subset of schools, creating four distinct treatment combinations. **Table 1** summarizes the structure of the interventions and treatment groups.

Table 1: Cross-over RCT design

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="3">Dimension 1: Community mobilization and facilitation of dialogue</td></tr>
<tr><td>No facilitation of dialogue</td><td>Meeting-based community dialogue (MEET)</td><td>SMS-based community dialogue (SMS)</td></tr>
<tr><td rowspan="2">Dimension 2: Elections and capacity building of SMC Executive Bodies</td><td>No Executive Body elections and capacity building</td><td rowspan="2">Control</td><td>MEET</td><td>SMS</td></tr>
<tr><td>Executive Body elections and capacity building (ELECTIONS)</td><td>MEET-E</td><td>SMS-E</td></tr>
</table>

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. 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 importance of education. All treatment groups' findings therefore include the impact of these common sensitization activities, which were not carried out in control villages.

The content of the remainder of the meeting, and any follow-up activities, varied according to the intervention. The first intervention, MEET, provided information to communities on their rights

[8] See Table A3, in Appendix A.

and responsibilities with regard to school management, and catalyzed dialogue around school quality, through an externally facilitated village meeting. This design enables us to assess the potential of external facilitation of dialogue in Sindh employing an established institution, community-wide meetings.

Following the introductory activities described above, meeting attendants were played a second, unique audio clip which contained a 20-minute dramatized discussion on the purpose, structure and roles of the SMC, and highlighting specific actions community members could take to improve education outcomes in schools, including participating in SMC meetings, visiting the school regularly to check the presence and activity of teachers, and monitoring any need for repairs and maintenance. A moderated discussion was then held amongst meeting participants around key ideas communicated in the foundation as well as the unique audio clips. At the end of the discussion, existing SMC Executive Body members were introduced to the villagers and their names and phone numbers given to participants. Villagers were encouraged to independently organize and conduct a follow-up village meeting.

It was expected that the intervention would lead to improved outcomes in school practices through two channels. The primary expected effect was improved engagement by the community with the SMC Executive Body, helping to ensure SMC alignment with community priorities. Second, it was expected that greater dialogue around education could lead directly to improvement in some aspects of school management, for example if teachers or other staff provided extra effort in response to heightened awareness of education within the community. Third, it was expected that increased awareness of education issues among communities could lead to improvements in demand-driven outcomes such as student enrollment. However, the meeting-based modality was also expected to remain subject to various barriers to participation, including high opportunity cost and social and cultural restrictions, which could prevent or limit follow-up engagement by participants.

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, followed by a hands-on demonstration by field facilitators. As with the MEET intervention, information was given to participants containing the names and telephone numbers of all SMC Executive Body members.

The CDP contained two key elements. First, key information shared in the MEET intervention, regarding rights and responsibilities, was reiterated through SMS messages to the community members. Later, rounds of additional messages provided additional information on the use of the portal. In order to encourage participants to move from dialogue to collective action, this second round also included information on the functioning of the SMC and School Improvement Plans. Second, the CDP provided a platform for user voice on education issues. Participants were provided with a number to which to send comments and responses to questions regarding conditions at their local school. Over the next five months, participants received messages on a weekly basis summarizing the topics and content of the messages received by the system, enabling the community to be aware of the dialogue about local schools and of each other's primary concerns. The design of the system was intended to minimize barriers to participation:

contributions were anonymous, and airtime was provided to participants to defray the cost of sending messages. Volunteers, selected by communities, were appointed to support mobile-illiterate members in contributing to the system.

As with the MEET intervention it was expected that the SMS intervention could affect school practices both directly, through the raised awareness and concern around education engendered by the dialogue, and indirectly through improved alignment of the SMC Executive Body with the communities' priorities. However, the SMS-based approach was intended to enable users to participate in a flexible way, with minimum investment of time, and anonymously, removing cultural and social disincentives for participation and allowing a greater number of community members to exercise their voice. In addition, the CDP was implemented continuously for five months, providing a longer period of facilitated voice than the meeting-based approach. The SMS intervention therefore enables us to assess the effectiveness of an ICT-based approach in breaking down barriers to participation.

The third intervention, ELECTIONS, specifically targeted SMC Executive Bodies as the formal mechanism of community collective action and engagement in school management. As noted above, many of Sindh's SMC Executive Bodies were not fully and regularly elected, meaning that even if the MEET and SMS interventions increased the level of community dialogue, the Executive Body could not be expected to respond to the concerns and preferences expressed. At the same time, the low level of utilization of SMC funds, under the direct purview of Executive Bodies, suggested a lack of enthusiasm or capacity among members to fulfill their roles. Therefore, the intervention included two aspects: holding of fresh elections for Executive Bodies, and participatory capacity building for Executive Body 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 training, a script, and 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 SMC Executive Bodies to assume official authority and claim SMC grant funds.

Following the elections, the Executive Body 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 Executive Body and provided participatory training on how to develop a School Improvement Plan (SIP), with members completing a draft plan chart during the meeting, and how to present the plan to the SMC General Body for ratification, withdraw and manage grant funds, and oversee implementation. It was expected that the ELECTIONS intervention would lead to improvements in school inputs, primarily through improved utilization of SMC funds in greater alignment with the priorities of communities; as well as improved performance in other areas under the direct purview of SMC Executive Bodies, such as enrollment of children entering school age.

2.5 Randomization

Table 2 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 2 about here]

2.5.1 Power Calculations

The specific service challenges to be addressed by community engagement vary widely between schools, with different schools facing divergent needs for improvement in staffing, facilities and conditions. The primary outcome measures on which sample size was determined at the school level are the school's functionality, infrastructure (classrooms, toilets, electricity, drinking water and boundary wall), registered teachers and teacher presence, and student enrollment. We find that our study is powered to detect effect sizes of 25-29 percent[9] for these outcomes, a comparatively low level of power which increases the challenge of identification.

For student test scores at the school level, we conduct an ex-post power calculation using endline normalized Learning Assessment Test (LATs) data. Similar to the other outcomes, we find that our study is able to detect effect sizes of 28-29 percent of changes in student test scores at the school level. This power calculation is based on an individual random assignment design, with the school serving as the primary sampling unit (PSU).

2.6 Timeline

Figure 3 summarizes the experiment sampling frame and timeline.

[Figure 3 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 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. Household endline survey activities were conducted between January and June 2017.

3. Data and Estimation Strategy

3.1 Data

We collected data over a four-year period, through one census and two survey exercises. The first data collection round was a census of households and schools in the three selected districts (2011).

[9] Power calculations reflect significance levels of 0.05 and R^2 of 0.5. Tables available on request.

The second was a baseline survey (2012) conducted in schools and sampled households. The initial plan of action 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 2012 to January 2013.

An endline survey was conducted in all schools from January 2015 to March 2015. The quantitative 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. Following analysis of the school endline survey data, the endline household survey was conducted for a sub-sample of 160 randomly selected baseline villages. 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. Children between the ages of 7-13 at the household level were tested at both the baseline and endline survey. In total we were able to track 85 percent of households in the sub-sample villages at endline.

In both the baseline and endline, the school surveys collected detailed data on school-level variables such as enrollment, attendance, teachers-on-task, facilities, infrastructure, SMC functioning, funding and expenditure. The field team completed teacher rosters to retrieve data on the total number of teachers on staff and teacher presence. Information on enrollments, teacher and student presence, and SMCs was collected from school records in interviews with headteachers. In addition, household surveys for both baseline and endline gathered information on household demographic and socioeconomic characteristics, household schooling choices and engagement with SMCs from the male and female head of the household and a student questionnaire administered to sampled children. In addition, standardized tests were administered to school-age children in sample households. At endline, these tests were also conducted with randomly selected students from Grades 3 and 4 in schools.[10]

3.1.1 Student Assessments

Standardized tests for English, Math and Sindhi were administered to children ages 7-13 at the household level both for the baseline and endline. In addition, to more precisely estimate the impact of the various interventions on learning outcomes, the same test was administered to Standard 3 and 4 students in schools at endline. The instruments were designed, pre-tested and piloted based on students' textbooks and curriculum standards for Standard 1-5 of primary school. The test instructions were translated to local vernacular and enumerators were trained to read out the instructions and work with the children on 2-3 practice questions from each subject before starting the test. The assessments were about 30 minutes long with 25 to 30 multiple choice questions, covering a wide range of difficulty level. Some open-ended questions were included in the Sindhi test to make the assessment consistent with exercises from the student's textbooks.

We fitted a two-parameter item response theory (2pl-IRT) model to account for what students of similar knowledge and skills ought to know by distinguishing between difficulty level. The

[10] Student tests administered to teachers follow a Service Delivery Indicators (SDI) style knowledge assessment in which teachers were asked to mark a random student's completed exam.

majority of the items across assessments for all subjects at baseline and endline were the same and we used the IRT model to express results of baseline and endline students on the same scale.[11]

3.2 Outcome Variables

Our choice of outcome variables is guided by the objective to measure improvement in the delivery of school services, as well as improvement in demand of schooling in the community, with careful attention to the political economy and governance characteristics of Sindh. First, to assess improvement in the overall functionality of schools, we measure whether schools were open at the unannounced visit at baseline and endline. As described in section 2.2, improvements in school functionality could reflect improvements in teacher attendance, particularly in one-teacher schools, or simply a response by service providers to increased awareness of and demand for education within communities.

Second, we look at total enrollment in schools, as well as enrollment by grade and by gender. Community engagement can drive improvement in enrollment through a number of channels: increased demand for education, as a result of increased awareness of education issues engendered by community dialogue; or increased confidence in schools, as a result of other improvements engendered by engagement (for example, in teacher attendance). In addition, in Sindh in particular, ensuring that children of school age are enrolled in school is a specific role of SMC Executive Bodies.

Third, in order to assess the impact of the interventions on the availability of and utilization of school inputs, we measure the condition of school staffing and infrastructure. Specifically, in terms of staffing, we measure the total number of teachers working at a school, and the number of teachers present at both baseline and endline. Improvements in the number of teachers working at a school are likely to reflect successful lobbying by communities for increased resource investment into schools, while improved teacher attendance could reflect either greater official monitoring of attendance as a result of community engagement, or a direct response by teachers to communities and the threat of social sanction. For infrastructure, we look at the impact on the availability of five key facilities that are important for a safe, secure and comfortable school environment: classrooms, toilets, drinking water, electricity, and a boundary wall. As discussed in section 2.2, improvements to school infrastructure could be achieved through improved utilization and targeting of SMC funds, improved resource mobilization from the community, or other channels.[12]

Fourth, to test for any overall impact on learning outcomes, we analyze students' test scores in Mathematics, English and Sindhi. Learning outcomes may improve as a result of changes in school service delivery, such as those described above; or from increased demand for education, leading to improved student attendance and effort. These outcomes enable us to assess whether the treatments enhanced the physical and functional capacity of the school; whether the collective demand for schooling increased with the community bringing out-of-school children to schools; and whether the effectiveness of schools increased, learning to improved learning.

[11] Appendix B describes all test instruments in detail.

[12] Gazdar (2000) notes that in many cases in Sindh, schools lack electricity or drinking water even in villages where they are widely available, suggesting a lack of collective action by communities to make these facilities available in schools.

Finally, in order to illuminate mechanisms of impact, we also measure the impact of the various interventions on numerous measures of community awareness of the SMC, its role and structure, and the wider institutional framework of school-based management.

3.3 Descriptive Statistics

Table 3 shows the observed values at endline in the control group villages and the various treatment groups, for a range of indicators.

[Insert Table 3 about here]

The findings suggest that, absent the intervention, the problems of the school system in Sindh have not considerably improved. Only 80 percent of schools were found to be open at the unannounced visit at endline. While ten percent of schools which had been closed at baseline in the control group were found to be open at endline, a greater proportion, 13 percent, had changed from open to closed. The average school had 71 students, including all grades from *katchi* (pre-primary) to Grade 5, suggesting the continuation of both Sindh's general pattern of multiple large schools. 46 percent of students at control schools at endline were female, suggesting continuing gender disparities in enrollment.

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 2.18 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. In terms of other indicators of infrastructure, only 51 percent of control schools had electricity, 60 percent drinking water, and 70 percent a boundary wall.

The control endline data suggests an ongoing crisis of staffing and teacher effort. The average control school had 2.25 teachers; however, an average five percent of teachers were absent even during announced visits. Fifty-two percent of control schools remained one-teacher schools at endline. The average school had only 16 percent female teachers. In terms of teacher effort, the average teacher spent only 4.23 hours teaching each day. The endline data also suggests ongoing problems of teacher knowledge. Asked to take a knowledge assessment based on the same curriculum as those undertaken by students, teachers scored an average 72 percent for English and 84 percent for Math, suggesting incomplete familiarity with even the very basic Standard 1 through Standard 5 material.

Unsurprisingly, given the poor condition of schools and teaching in the control group at endline, test scores remained low. School-age children undergoing learning assessment in household surveys at endline in control villages scored, on average, 22 percent across subjects.

3.4 Estimation Strategy

3.4.1 School-level Effects

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

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

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 this 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 also include point estimates of pooled treatment effects of MEET-SMS and Pooled ELECTIONS on outcome variable. We test the sensitivity of our estimates to inclusion of treatment interaction with participation rates at the village meeting.

ANCOVA is the preferred estimator as standard difference-in-difference (DiD) estimator would require twice as much sample to yield the same power (McKenzie, 2012). We are underpowered to detect small effect sizes for a standard DiD estimator.

3.4.2 Student level Effects

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

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

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 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.

4. Results

This section presents findings, beginning with the level of participation and proceeding to intention-to-treat (ITT) estimates of impact of the interventions on the outcomes of interest. For the ITT estimates, we first discuss the impact of the MEET and SMS interventions, then the impact of the addition of the ELECTIONS intervention.

4.1 Treatment Fidelity

Table 4 illustrates fidelity to treatment for the various interventions. All three interventions were implemented through community-wide meetings, which provided the forum for facilitated

discussion under MEET, introduction of the CDP under SMS, and the electing of SMC Executive Bodies under ELECTIONS; therefore, we first measure fidelity to treatment through the rate of attendance at these meetings (Panel A).[13] 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.[14] Overall, on average the participation rate was 63% among treatment villages. Mirpurkhas had the highest participation rate with 75% households, followed by 57% in Sanghar and 51% 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%, compared to 58% for MEET-E and SMS-E. The differences in attendance are particularly stark for Matiari where average participation rates were only 38% in groups receiving ELECTIONS, and 63% for MEET and SMS treatments alone. Given the low quality of relations between the public and government in Sindh, a likely explanation is the presence of the Taluka Education Officer, a representative of government who was only present at the meeting in villages receiving the ELECTIONS intervention. Uniquely among the three experiment districts, Matiari is a new district created in 2005; newly created districts have been associated with low trust in government as a result of high and unmet expectations for rapid improvements in service delivery.

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% of village meeting attendees registered with a valid cellphone for the CDP. The average registration rate in non-ELECTIONS villages (65%) was slightly higher than that in ELECTIONS villages (61%). A total of 4,981 unique users registered on the CDP portal, of which 28 percent sent at least one non-junk message into the portal during the course of the project. All registered users received SMS summaries of messages sent into the portal. Participation in the portal appears to have been sustained over time: although approximately one-third of the messages received by the platform were sent in the first ten days following the introductory meeting, an average of two messages per village were sent during the last ten days of implementation five months later.

For the first key aspect of the ELECTIONS intervention, the holding of elections for SMC Executive Bodies, 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 SMC Executive Bodies, we also measure the extent to which the membership of these Bodies changed following elections (Panel B). Executive Bodies 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 exactly three new members; all four elected members

[13] The project team were instructed to revisit any village for which attendance was under 20% and hold a second meeting. Reported attendance is for second meeting in these cases.
[14] As discussed in Section 2.3, sample includes all main settlement schools and 15% of schools outside main settlements of sampled villages.

were replaced in slightly more than one-third of schools (36 percent). The extent of composition change varied slightly between the two ELECTIONS treatment groups, from an average 74 percent in MEET-E villages to 71 percent in SMS-E villages.

The second key aspect of the ELECTIONS intervention was capacity building for SMC Executive Body members, carried out in the month following elections. Panel C shows the attendance rate at these sessions. The mean attendance for the three meetings, across MEET-E and SMS-E, was 96 percent. Scanned copies of School Improvement Plans prepared by the Executive Body during the capacity building were received for all schools in ELECTIONS groups.

Overall, fidelity to treatment was sufficiently consistent across treatment groups to enable accurate comparison of effectiveness based on ITT estimates. The variance in take-up between districts does pose a potential threat to validity; this is addressed by the inclusion of district fixed effects and effect sizes are insensitive to village participation rates for most outcomes, with small positive effects on school functioning.[15]

4.2 Average ITT effects: Community mobilization and facilitation of dialogue

This section presents and compares the impact of the first two interventions, MEET and SMS, both of which focused on information provision and facilitation of community dialogue. If in line with the findings of Cerdan-Infantes and Filmer (2015), SMS-based interventions can achieve a higher intensity of engagement than meeting-based interventions, we would expect to see greater impacts in the SMS group than in MEET across all outcomes of interest. These results are shown in the first three estimations in each of **Tables 5-9.**

4.2.1 School functioning

We first present estimates of the impact of the interventions on school functioning. As described in Section 3.2, improvements in school functionality may reflect a response by service providers to increased demand for education, as well as improvements in teacher attendance (particularly in one-teacher schools).

Comparison of means at endline between the MEET and SMS groups suggests that the MEET intervention had significant impacts on the overall number of functioning schools **[Table 3]**. To more fully analyze changes in school functioning, we next present transition probability analysis, measuring the likelihood of a change in school-level outcomes **[Table 5, Panel A]**. 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 difference in transition percentages between each treatment group and control represents the additional likelihood of a change in outcomes in the stated direction. For school functioning, we present analysis of the likelihood of the school transitioning from closed to open between baseline and endline; to capture any potential reduction in school closures, we also capture the likelihood of schools transitioning from open to closed.

[Table 5 about here]

[15] See Table A4, in Appendix A.

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, a difference in percentage of -9.0; 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.

For SMS, the picture is more mixed. The intervention does not appear to have significantly reduced the likelihood of a school closing between baseline and endline; however, the intervention does appear to have led to schools which were closed reopening. For SMS schools, the chances of transitioning from closed to open were more than double those for control schools (22.3 percent versus 10.2 percent, a percentage difference of 12.1 percent).

Next, we present the regression estimations for school functioning **[Table 6]**. 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; as in the transition analysis, they were also significantly 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 more likely to have reopened by endline.

[Table 6 about here]

Overall, the findings suggest that the MEET intervention had a primarily defensive impact, preventing the closure of a significant number of schools; while the SMS intervention was more effective in causing closed schools to reopen.

4.2.2 Inputs

We next turn to the impact of the interventions on key school inputs. Comparison of means **[Table 3]** suggests significant impacts from MEET and SMS on the proportion of schools with only one teacher. **Panel B** of **Table 5** presents transition probability analysis for inputs; in terms of staffing, the results suggest significant impacts from both MEET and SMS: MEET schools were almost twice as likely as control schools to gain at least one additional teacher (35.4 percent of MEET schools versus 17.8 percent of control, a difference of 17.6 percent); 29.7 percent of SMS schools gained at least one teacher, a difference from control of 11.9 percent, meaning that SMS schools were 66 percent more likely to gain staff than control schools.

A similar pattern is observed in the regression point estimates **[Table 7, Panel A]**; MEET schools had 0.26 additional teachers at endline compared to control schools, while SMS schools had an additional 0.3 teachers, equivalent to a 13 percent improvement in staffing. However, in the case of MEET, the finding is not statistically significant, reflecting limitations in power as described in Section 2.5.1. 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, suggesting that the improvements in staffing did not translate to improvements in the actual presence of teachers in school.

[Table 7 about here]

Comparison of means **[Table 3]** provides some clarification, revealing that the rate of teacher absence actually increased in treated schools. This is in line with findings from other similar interventions, which found that increased staffing at schools is often counteracted by a reduction in effort and presence by other teachers (e.g. Duflo, Dupas and Kremer, 2015).

We do observe a substantial positive impact on teacher presence from SMS, with treated schools having an average 0.24 additional teachers present at endline compared to control, equivalent to as 12 percent increase in teacher presence; however, the standard error is large and the result does not attain statistical significance. Comparison of means reveals some increase in teacher absence at SMS schools, but this again is not significant.

Turning to physical infrastructure, **Panel B** of **Table 5** 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 point estimates **[Table 7]**, we include additional 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; we similarly observe no impacts in the comparison of means **[Table 3]**.

Overall, the findings suggest that the MEET and SMS interventions did have an impact on the overall level of staffing at schools; however, no impact is observed on physical infrastructure.

4.2.3 Enrollment

Next, we turn to the impact of the interventions on enrollment. Improvements to enrollment could be achieved through increased awareness in communities of the importance of education, engendered through the village meetings in both the MEET and SMS interventions, and additionally through participation in the CDP under the SMS intervention. Increased enrollment could also reflect increased confidence by communities in the quality of schooling as a result of other improvements, or increased effort by SMC Executive Bodies to register out-of-school students.

The MEET intervention does not appear to have significantly improved enrollments in either comparison of means **[Table 3]** or regression point estimates **[Table 8]**. However, in regressions 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, equivalent to additional enrollment of 21 percent. The effect was distributed between Grades 2, 3 and 4, suggesting reductions in dropout throughout the school cycle as a result of the intervention.[16] However, no significant impact was observed on girls' enrollment.

[Table 8 about here]

[16] It is worth noting that, even in the case of the SMS intervention, endline data collection took place several months after the intervention ended, suggesting lasting change in enrollment patterns.

4.2.4 Learning Outcomes

Finally, in order to assess the overall impact of the interventions on learning, we present findings on the impact of the interventions on learning outcomes, according to our standardized tests tools. Difference of means suggests no significant impacts from either MEET or SMS. **Table 9** shows regression point estimates: **Panel A** shows findings from LATs completed in schools; in order to capture spillover effects on out-of-school children, **Panel B** presents findings from LATs completed during household surveys. We find no significant impact on learning outcomes for either the MEET or SMS interventions.

[Table 9 about here]

4.3 Average ITT effects: Support to elections and capacity of SMC Executive Bodies

We next turn to the additional impacts of the ELECTIONS intervention. As described in Section 2.4, the intervention, which consisted of holding of elections and provision of participatory capacity building for SMC Executive Bodies, was conducted in a cross-over design with the MEET and SMS interventions. The fourth, fifth and sixth estimations in each of Tables 4 and 6-10 present the impact of the MEET and SMS interventions combined with the ELECTIONS intervention. We present results for the two combined treatments – MEET-E and SMS-E – as well as a pooled impact reflecting both groups.

4.3.1 School functioning

In the case of school functioning, given that the reopening of a school is likely to reflect general community concern with, and elite support for, education rather than the use of SMC grants explicitly, we would not expect significantly greater impact on school functioning from the ELECTIONS intervention than from MEET and SMS alone. In fact, however, the addition of the intervention appears to have had a negative impact which counteracts the impact of the MEET and SMS interventions. In the comparison of means **[Table 3]** and transition probability analysis **[Table 5]**, the impact of MEET in reducing the likelihood of school closures is reduced 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, in which MEET-E and SMS-E schools are not significantly more likely to be open at endline, or to have reopened, or less likely to have closed.

4.3.2 Inputs

In the case of inputs, the picture is more mixed **[Table 7]**. In the comparison of means, the impact of MEET and SMS on the likelihood of a school having only one teacher is reduced in MEET-E and SMS-E, and no longer significant in the case of MEET-E. Similarly, 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 MEET-E and SMS-E, although a significant impact is still observable in the pooled estimation, with 27.8 percent of ELECTIONS schools overall gaining at least one teacher versus 17.8 of control schools. Similarly, the impact of SMS on teacher numbers in the regression analysis disappears with the addition of ELECTIONS.

However, the addition of ELECTIONS does appear to have positively impacted physical infrastructure when carried out in combination with the SMS intervention. In the transition probability analysis, SMS-E schools were significantly more likely to gain an additional classroom than either control schools or SMS schools. More than one in four SMS-E schools – 27.7 percent – gained an additional classroom, versus 17.6 percent in SMS and 15.1 percent in control. Furthermore, SMS-E had a significant and positive impact on the availability of toilets which was not observed for SMS alone, observable in both the transition analysis **[Table 5]** and point estimates **[Table 7]**, though not in the comparison of means **[Table 3]**.

4.3.3 Enrollment

A similar pattern is observed for measure of enrollment **[Table 8]**. As with school functioning, given that enrollment is likely to respond primarily to community-level awareness and enthusiasm for education rather than SMC investment specifically, we would not generally expect significant additional enrollment as a result of the ELECTIONS intervention. In fact, as with schools reopening, the impacts observed from the SMS intervention on enrollment are reduced and become marginally insignificant in SMS-E. However, SMS-E does outperform SMS in enrollment of pre-primary (*katchi)* and Grade 1 boys, a key responsibility of the SMC Executive Body. The average SMS-E school had four more students in these grades than in the control group, an impact of 21 percent.

4.3.4 Learning Outcomes

Positive impacts for ELECTIONS are observed with regard to learning outcomes, where point estimates demonstrate a significant impact for SMS-E which is not observed for SMS alone. The SMS-E treatment raised student overall LAT scores by 15 points. A look at the subject-specific effects shows this impact to be principally driven by improved scores in Math, though a large, if statistically insignificant, impact was also observed for English. By contrast, no similar effect is observed for MEET-E.

5. Discussion

Our findings suggest that the impacts of the interventions varied considerably between outcomes. Looking first at the interventions focused on community mobilization and dialogue, we find that both MEET and SMS had significant impacts on school functioning and numbers of registered teachers, suggesting improved support to schools from government and local elites in response to community mobilization. Given the large proportion of schools in Sindh with only one or two teachers, in which teacher absence can lead directly to school closure, it is likely that these two impacts are closely related, with the provision of additional teachers preventing school closures and enabling closed schools to reopen.

Specifically, while MEET had a predominantly defensive impact on school functioning, preventing closures, SMS was more effective in causing closed schools to reopen, and also had significant impacts on enrollment which were not found under MEET. This suggests that SMS

was more effective at creating and sustaining improvements in demand for education over time. These likely stems from the ongoing nature of the SMS intervention, in which the CDP operated for a period of five months, enabling sustained improvements in awareness of, and demand for, education in treatment communities. It is noteworthy that, in contrast to previous similar ICT-based interventions (for example, Cerdan-Infantes and Filmer, 2015), the CDP not only provided information to users, but provided a platform for exercise of voice and expression of preferences, and provided summary feedback to users to keep them informed of each other's collective preferences. This process of two-way dialogue may have been more effective in sustaining long-term participation than a more one-way information dissemination approach taken by other ICT interventions.

The impacts of SMS on enrollment did not extend to girls. This may simply suggest that the SMS intervention, although effective, was not powerful enough to overcome cultural barriers to girls' enrollment. However, an alternative explanation is differential access to the CDP by female users. 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.

Neither MEET or SMS significantly improved physical infrastructure at the school level. Given that the primary mechanism of improvements in physical infrastructure is SMC funds, this suggests the voice-focused interventions were not effective in improving the activity level or responsiveness to the community of SMC Executive Bodies, at least with regard to utilization of SMC funds. Overall, the impacts of these voice-focused interventions appear to have primarily reflected increased demand for education from communities, and a degree of responsiveness by government and elites to this demand, rather than improvements in collective decision-making through the institution of SMC Executive Bodies.

In order to more clearly assess the impact of the interventions on community engagement with collective decision-making institutions, we measure the awareness of households of a school's SMC Executive Body, its institutional arrangements and functions **[Table 10]**. The findings demonstrate the difficulty interventions face in patrimonial contexts in improving community engagement with formal institutions of engagement: despite information about SMCs being a key aspect of both the MEET and SMS interventions, this information appears to have failed to raise awareness by communities of the structure or functioning of Executive Bodies, or of the school's entitlement to SMC grant funds.

[Table 10 about here]

Although surprising, these findings reflect those from other countries suggesting that the provision of information alone does not significantly improve community engagement (e.g. Banerjee et al., 2010). The findings suggest that those improvements which were observed did not stem from increased community engagement with the SMC Executive Body as originally envisioned, but may more likely reflect the direct impacts of greater community engagement with education and informal collective action by the wider community. However, given the significant time delay (more than four years) between the interventions and endline household survey, an additional potential explanation is that awareness did increase, but these effects were not sustained over time.

Turning to the additional intervention supporting SMC Executive Bodies, we find a mixed picture. The ELECTIONS intervention does appear to have driven improvements in school infrastructure, particularly in combination with SMS. This suggests that the combination of elections to Executive Bodies, which led to the replacement of an average three out of four elected members, and participatory capacity building for Executive Body members, succeeded in improving utilization of SMC funds in alignment with community priorities. Additional evidence for this mechanism of impact comes from a comparison of community preferences for the use of SMC funds, as expressed through the CDP under the SMS intervention, with investment priority decisions made by SMC Executive Body members during the capacity building exercise. **Figure 4A** presents a summary of community preferences in CDP, demonstrating a clear preference by communities for investment in physical infrastructure (the single largest item at 38 percent). **Figure 4B** presents the expenditure priorities identified by SMC Executive Body members during the participatory capacity building exercise; the findings show a similar level of interest in physical infrastructure, with 'building', 'classrooms' and 'toilets' accounting for a total of 34 percent. This suggests that the capacity building was successful in aligning SMC priorities with those expressed by communities. The improvement in Grade 1 enrollment, a particular area of responsibility for Executive Bodies, under SMS-E provides further evidence that ELECTIONS did succeed in raising the capacity of Executive Bodies to fulfil their mandated roles.

The holding of elections does not appear to have driven increased awareness of the functioning of SMCs and Executive Bodies. Households in villages exposed to the ELECTIONS intervention did not experience significant gains in awareness of the SMC Executive Body functions **[Table 10]**. This suggests that the impact of ELECTIONS stemmed from improvements in effort and attention to community priorities by Executive Body members themselves, rather than enhanced expression of preferences by communities.

Put more succinctly, the findings suggest that while enhancing the capacity of communities to express preferences may lead to improvements in demand for education and resulting improvements in school functioning, direct support to the capacity of representative bodies is required to obtain improvement in their capacity and alignment with community priorities, and resulting improvements in the utilization of funds under their control.

One particularly surprising finding is that the addition of ELECTIONS appears to have undermined the positive impacts of MEET and SMS on school functioning and availability of teachers, and of SMS on enrollment (in grades other than Grade 1). As discussed in Section 4.1, the most likely explanation is the reduced level of participation in the introductory village meetings in ELECTIONS villages; given that the MEET intervention was based entirely in this meeting, reduced participation is likely to significantly reduce the impact of the intervention, while for SMS, the negative impacts are mitigated by the fact that the intervention continued for several months following the initial meeting.

As described above, the reduced participation in ELECTIONS villages likely reflects an adverse community reaction to the presence of a government official, the Taluka Education Officer, at the meetings. The finding suggests that in conditions of low trust in government, trade-offs may exist between the dual goals of bolstering community engagement and supporting elections to

representative 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). An additional potential explanation is that the replacement of existing Executive Body members, often appointed by elites, with elected members damaged patronage networks which were effective in delivering some degree of inputs to schools.

We do not observe significant impacts from MEET and SMS on learning outcomes. As described in Section 2, a range of barriers can prevent improvements in school inputs from translating into improvements in learning outcomes. In contrast to other community engagement interventions, the approaches trialled in Sindh did not include structured engagement directly between communities and teachers, a factor common to many of the interventions which have achieved improvements in learning outcomes (Asim et al., 2017). Given the limited motivation and capacity of the teaching workforce in Sindh, as evidenced by low attendance and poor scores on curricular testing **[Table 3]**, it is perhaps not surprising that increased teacher recruitment did not translate into increased learning.

The lack of impact on learning outcomes may also reflect a side-effect of the increased enrollment and reduced dropout engendered by the SMS intervention. Across a wide range of countries, dropout tends to be highest among children from poorer backgrounds, owing to greater opportunity costs in terms of informal fees and lost labor (World Bank, 2018). Students from poorer households also typically exhibit lower learning outcomes; more generally, students with lower academic performance are typically most likely to drop out as a result of lower enthusiasm, parental doubts over the value of education. Reductions in dropout are therefore likely to reduce the average ability of the student population by undoing these selection effects. It is therefore possible that, while students did experience some benefit in learning, these effects were counteracted by the impacts of increased enrollment. It is notable that in schools which received both the SMS and ELECTIONS schools, where enrollment did not significantly increase, we do observe improvements in learning outcomes; this may suggest that learning improvements are being engendered in SMS schools, but that these are counteracted by increased enrollment.

More generally, the findings are in line with those conclusion of Bruns et al. (2011) that improvements in schools as a result of SBM typically take several years to reflect in learning outcomes.

6. Conclusion

We report the findings of a randomized control trial of three interlinked interventions designed to bolster community engagement in schools in a cross-over design: external facilitation of face-to-face dialogue at community meetings; dialogue via text messages, on a specially-designed high-frequency, anonymous and low-cost text message platform; and elections and training of SMC Executive Body members. The setting is rural Sindh, Pakistan, a province in which feudal landholding patterns and social fragmentation have historically undermined engagement.

We find both meeting- and SMS-based dialogue to be effective in improving school staffing and the number of schools functioning. The MEET intervention reduced school closures by more than half, and reduced the number of schools with only one teacher by 19 percent; the SMS intervention achieved a 13 percent improvement in staffing, a 17 percent reduction in the number of one-teacher schools, and led to significantly more closed schools reopening. Despite improving school staffing, the interventions failed to improve the number of teachers actually present at school; in contrast to other similar interventions (for example, Barr et al., 2012), teacher absence appears to have increased. The most likely explanation is that existing teachers responded to improvements in school staffing by reducing effort.

In combination with SMS, elections and capacity building for SMC Executive Bodies achieved improvements in effort by SMC members, observed through improvements in school infrastructure and in enrollment of students entering school age. The combination of SMS and support to executive bodies also achieved improvements in learning outcomes. However, the elections intervention also reduced community participation in engagement meetings and, as a result, largely eliminated the positive impacts of the dialogue interventions.

Our findings suggest that, even in contexts where significant barriers to participation exist, mechanisms of community engagement can generate large improvements in school outcomes, particularly with regard to availability of education. The improvements in school functioning and enrollment engendered by the voice interventions represent a significant potential gain in access to schooling in Sindh: the SMS intervention generated a twenty percent improvement in boys' enrollment; given the Sindh rural boys' NER of 62 percent, this would constitute a major improvement in access if replicated across the province. In addition, both MEET and SMS improved the number of schools functioning, with clear implications for access. Cost-effectiveness analysis suggests that these gains could be scaled up at relatively low cost: including gains from increased enrollment and additional school functioning, estimates suggest that a one percent increase in student enrollment can be obtained for $9,271 through MEET and $10,148 through SMS **[Table 11]**.

The success of the SMS intervention in improving enrollment, in comparison to the failure of MEET in this area, suggests that ICT-based interventions can meaningfully drive community engagement on an ongoing basis even in conditions where large barriers exist to face-to-face participation. It is likely that the interactive nature of the SMS platform, which included not only provision of information but dialogue and elicitation of community preferences, helped sustain participation and improve demand for education, with benefits to enrollment.

Given the prevalence of mobile phones in developing countries, governments should consider using SMS-based platforms to dispense information and solicit feedback. In response to emerging evidence of the effectiveness of ICT-based platforms, the Pakistan Education Secretary has introduced an SMS-based initiative to create linkages between teachers, students and parents (Ilmi SMS). However, the lack of impact on girls' enrollment does suggest limitations to the ability of ICT-based engagement solutions to alter entrenched social restrictions.

Despite the impacts of the dialogue interventions, our findings also suggest that, in conditions of feudal or highly patrimonial politics, mechanisms of voice are more likely to improve demand for

education than to generate improvements in collective decision-making through formal institutions. A key aspect of both the MEET and SMS interventions was provision of information to communities of information on SMC Executive Bodies to improve engagement by communities with members and ensure better utilization of SMC funds. However, neither intervention appears to have improved the awareness of communities of SMC Executive Body functions or purpose, or to have led to improvements in the utilization of SMC funds as measured through improvements in school physical infrastructure. Elections and support to Executive Bodies led to improvements in effort by members, but with negative effects on participation in dialogue and, as a result, on impacts on school functioning and overall enrollment. The findings suggest that in the context of low trust in government, mechanisms of community representative decision-making are likely to be engender the greatest engagement and impacts when kept at arm's length from government.

Finally, the study affirms findings from a range of developing countries that improvements to school inputs from community engagement do not always translate into improvements in learning outcomes. A potential explanation is that benefits to learning did take place, but were counteracted at the school level by the gains in enrollment, which reduced the average academic ability of the student population.

The study is subject to limitations, particularly with regard to power. The widespread rate of school closure in Sindh, with around one-fourth of sample schools closed at either baseline or endline, limits the number of observations for key outcomes of interest, resulting in a high minimum detectable effect for the majority of outcomes. It is likely that a larger-scale trial would detect a wider range of significant impacts. In addition, with regard to infrastructure, the wide range of potential areas of investment for school funds limits the ability to detect impacts by tracking the condition of a small number of key infrastructure indicators; future studies may benefit from directly tracking utilization of school funds and measuring a wider range of school-level investments.

Several potential areas and questions emerge for future research. The first is the role that village elites play in impacting service delivery outcomes. As noted above, an additional potential explanation for the negative impacts of Executive Body elections may be that replacing established members established channels of patronage. In highly neopatrimonial conditions where the flow of resources to schools is heavily controlled by elites, efforts to improve the representativeness of SMCs with communities through elections may have negative consequences for resource flow to schools. Further research is required to more fully investigate the twin dynamics of patronage and community collective action in influencing school resources.

Second, we have demonstrated that even with stronger linkages between school officials, SMCs and community members, parents often do not know what is happening inside the schools. Gains in staffing as a result of the SMS intervention were not mirrored in increases in the number of teachers present in schools, likely reflecting decreased effort by other teachers; there is scope for further development of comprehensive community-level monitoring and accountability, an area that can be explored in future research.

References

Abdul-Hamid, H. 2017. *Data for Learning: Building a Smart Education Data System*. Washington, D.C.: World Bank Publications.

Andrews, M., Pritchett, L., and Woolcock, M. 2012. "Escaping Capability Traps through Problem-Driven Iterative Adaptation (PDIA)." *Working Paper 299. Washington, DC: Center for Global Development.*

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., Raju, D. 2014. "Achievement in Government Schools in Punjab, Pakistan: Evidence from large scale administrative data." *Education Sector Review.* Washington DC.: World Bank Publications.

Asim, S., Chase, R., Dar, A. and Schmillen, A. 2017 "Improving Learning Outcomes in South Asia: Findings from a Decade of Impact Evaluations." *The World Bank Research Observer* 32(1):75-106

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.

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 Publication.

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

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

Bold, T., Filmer, D., Martin, G., Molina, E., Stacy, B., Rockmore, C., Svensson, J. and Wane, W. 2017. "Enrollment without Learning: Teacher Effort, Knowledge, and Skill in Primary Schools in Africa." *Journal of Economic Perspectives* 31(4):185–204.

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

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.

Chattopadhyay, R., and Duflo, E. 2004. "Women as policy makers: Evidence from a Randomized Policy Experiment in India." *Econometrica*, 72(5), 1409-1443.

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. 2006. "A Comparative Analysis of School-Based Management in Central America." Working Paper 72. 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.

Evans, D. and Popova, A. "What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews." World Bank Policy Research Working Paper 7203. Washington, D.C.: World Bank.

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, Paul J., Marta Rubio-Codina, and Harry A. Patrinos. 2006. "Empowering Parents to Improve Education: Evidence from Rural Mexico." Policy Research Working Paper 3935. Washington, D.C.: World Bank.

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.

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

Islam, A. 2017. “Parent-Teacher Meetings and Student Outcomes: Evidence from a Developing Country.”

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.

Keefer, P. 2007. "Seeing and believing: political obstacles to better service delivery." In: S. Devarajan and I. Widlund, eds., *The politics of service delivery in democracies: Better access for the poor.*

King, Elizabeth M., and Berk Özler. 1998. “What’s Decentralization Got to Do with Learning? The Case of Nicaragua’s School Autonomy Reform.” Mimeo. Washington, D.C.: Development Research Group, World Bank.

Lopez-Calva, Felipe, L., and Espinosa, L. 2006. “Efectos diferenciales de los programas compensatorios del CONAFE en el aprovechamiento escolar.” In *Efectos del Impulso a la Participación de los Padres de Familia en la Escuela* (Effects of Parental Participation in the School), ed. Antonio Blanco Lerín. Mexico City: Consejo Nacional de Fomento Educativo.

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

Ostrom, E. 1990. "Governing the commons: The evolution of institutions for collective action." Cambridge, New York: Cambridge University Press.

Paes de Barros, R., and Mendonça, R. 1998. “The Impact of Three Institutional Innovations in Brazilian Education.” In *Organization Matters: Agency Problems in Health and Education in*

Latin America, ed. William D. Savedoff, 75–130. Washington, DC: Inter-American Development Bank.

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.

Pritchett, L. and Beatty, A. 2012. "The Negative Consequences of Overambitious Curricula in Developing Countries." John F. Kennedy School of Government Faculty Research Working Paper Series. Cambridge, MA: Harvard University.

Pritchett, L. 2013. "The Rebirth of Education: Schooling Ain't Learning." Washington, D.C.: Center for Global Development.

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

Santibañez, L. 2006. "School-Based Management Effects of Educational Outcomes: A Literature Review and Assessment of the Evidence Base." Lomas de Santa Fe: Centro de Investigación y Docencia Económicas.

Sawada, Yasuyuki, and Ragatz, A. 2005. "Decentralization of Education, Teacher Behavior and Outcomes: The Case of El Salvador's EDUCO Program." In *Incentives to Improve Teaching: Lessons from Latin America*, ed. Emiliana Vegas, 255–306. Washington, D.C.: World Bank Publications.

Snilstveit, B., Stevenson, J., Menon, R., Phillips, D., Gallagher, E., Geleen, M., Jobse, H., Schmidht, T. and Jimenez, E. 2016. "The impact of education programmes on learning and school participation in low- and middle-income countries." 3ie Systematic Review Summary 7. London: International Initiative for Impact Evaluation (3ie)

Thompson, M. 2008. "ICT and Development Studies: Towards Development 2.0." *Journal of International Development* 20(6), 821–35.

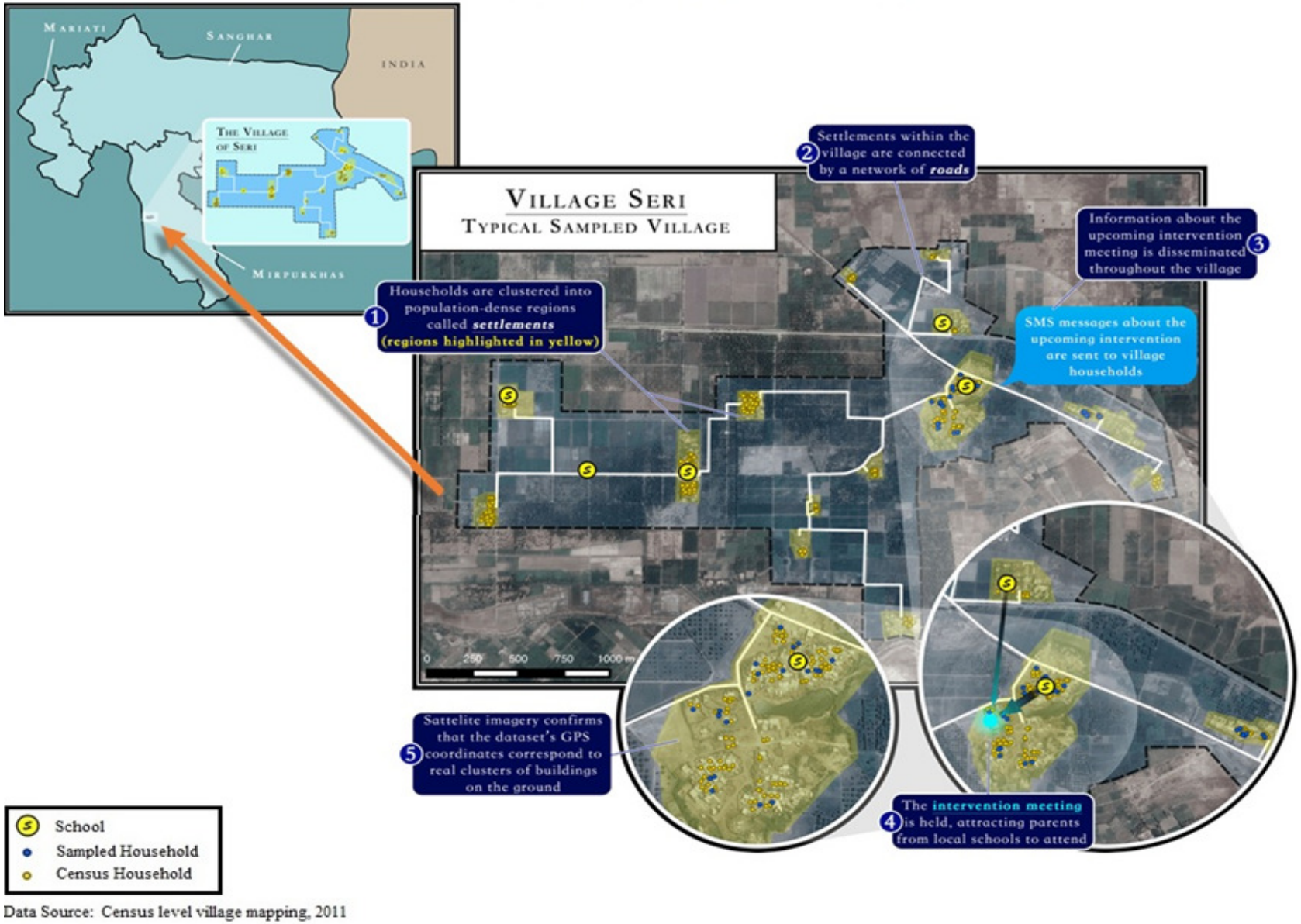
Wade, R. 2002. "Bridging the Digital Divide: Route to Development or New Form of Dependency?" *Global Governance* 8(4), 443–66

World Bank. 2004. “World Development Report 2004: Making Services Work for the Poor.” 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.

Tables and Figures

Figure 1: Sampled Village, Seri in Mirpurkhas District



Data Source: Census level village mapping, 2011

Figure 2: Population and sampling profile

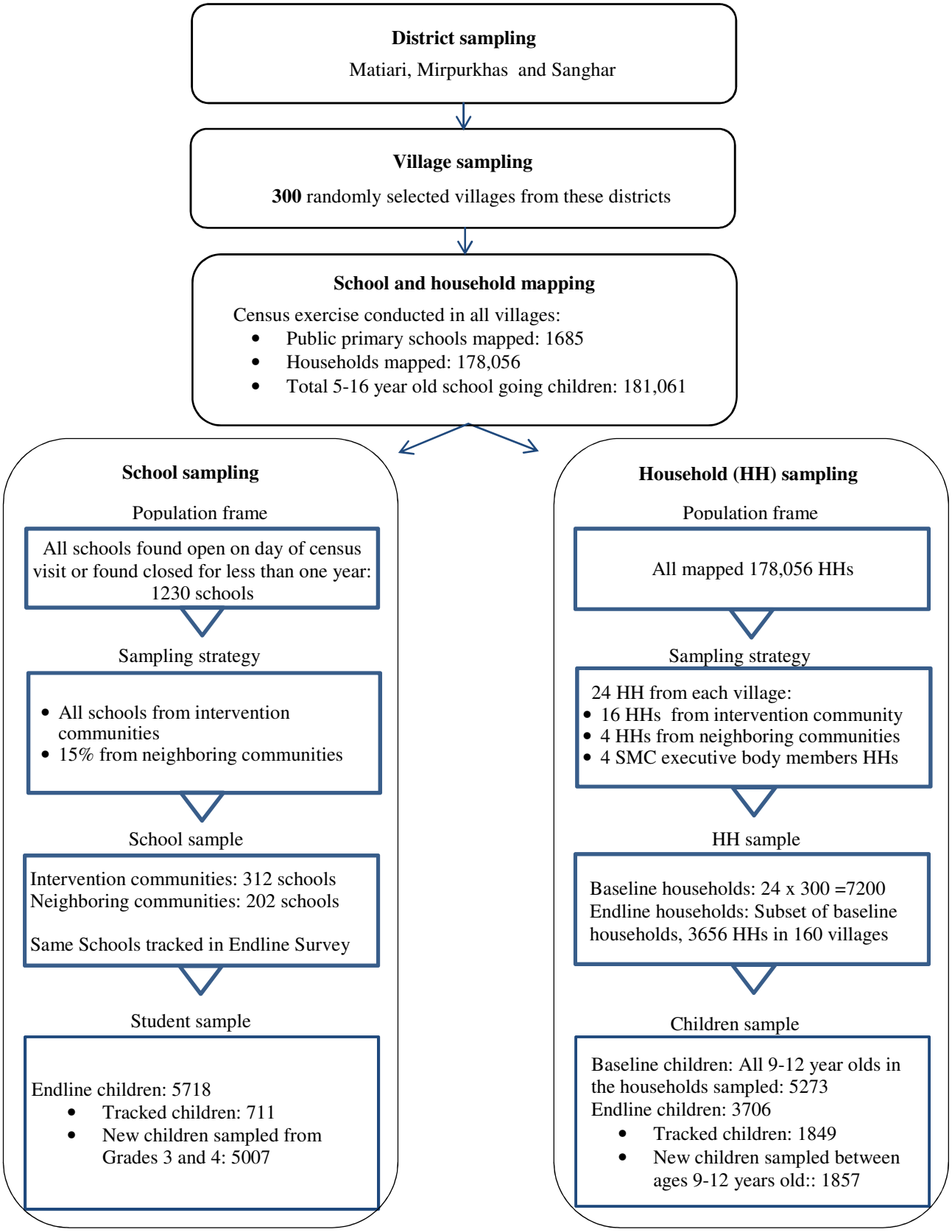


Table 2: Mean Comparison of Baseline Characteristics

		Difference (Treatment-Control)						
	Control (1)	MEET (2)	SMS (3)	POOLED (4)	MEET-E (5)	SMS-E (6)	POOLED-E (7)	N
Village characteristics								
Number of Primary Schools	5.54	-0.27	-0.49	-0.38	0.11	0.58	0.34	284
	(3.36)	(0.41)	(0.39)	(0.28)	(0.44)	(0.43)	(0.31)	
Share of Schools open on Unannounced visit	0.83	-0.09	-0.09	-0.09	0.04	-0.07	-0.02	284
	(0.34)	(0.06)	(0.05)	(0.04)	(0.04)	(0.05)	(0.03)	
Village affected by Flood	0.72	0.02	-0.01	0.01	-0.05	-0.06	-0.06	284
	(0.45)	(0.06)	(0.06)	(0.04)	(0.06)	(0.06)	(0.05)	
Share of Households with Cell Phones	0.76	-0.01	-0.04	-0.03	-0.03	-0.03	-0.03	284
	(0.22)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	
English Test Scores	0.06	0.07	-0.03	0.02	0.06	0.04	0.05	274
	(0.41)	(0.05)	(0.06)	(0.04)	(0.07)	(0.06)	(0.04)	
Math Test Scores	0.05	0.08	-0.03	0.02	0.08	-0.05	0.01	274
	(0.44)	(0.07)	(0.07)	(0.05)	(0.07)	(0.06)	(0.05)	
Sindhi Test Scores	0.14	-0.02	-0.04	-0.03	0.02	-0.12	-0.05	274
	(0.44)	(0.06)	(0.06)	(0.04)	(0.06)	(0.06)	(0.04)	
Household is aware of:								
SMC	0.13	0.02	0.02	0.02	0.02	0.02	0.02	284
	(0.08)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
SMC has five members	0.06	0.01	0.00	0.01	0.00	0.01	0.01	284
	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
SMC meets at least twice a year	0.11	0.01	0.02	0.01	0.01	0.02	0.02	284
	(0.08)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
SMC receives PKR 22,000 annually	0.11	0.01	0.01	0.01	0.00	0.01	0.01	284
	(0.08)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
School Improvement Plan (SIP)	0.06	0.01	0.02	0.02	0.00	0.02	0.01	284
	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
School Characteristics								
School Size (Enrollment)	73.89	11.81	2.32	7.36	2.73	1.08	1.93	249
	(43.81)	(7.01)	(6.59)	(4.84)	(6.03)	(5.49)	(4.07)	
Total Girls	25.63	2.43	0.04	1.30	-2.21	-3.63	-2.90	249
	(23.74)	(3.70)	(2.47)	(2.27)	(3.03)	(2.37)	(1.93)	
Total Boys	47.71	7.73	2.28	5.18	4.55	4.75	4.65	249
	(36.38)	(5.07)	(5.14)	(3.61)	(4.66)	(4.99)	(3.39)	
Total Teachers (Available)	2.29	0.02	-0.01	0.01	-0.03	0.09	0.03	249
	(1.78)	(0.25)	(0.22)	(0.17)	(0.22)	(0.26)	(0.17)	
Share of Teachers Absent	0.09	-0.03	-0.01	-0.02	-0.01	-0.01	-0.01	249
	(0.18)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	
Number of Classrooms	2.13	0.06	0.22	0.13	-0.25	0.02	-0.12	249
	(1.23)	(0.22)	(0.23)	(0.16)	(0.15)	(0.15)	(0.10)	
Toilets available	0.68	-0.07	-0.04	-0.06	-0.07	-0.01	-0.04	249
	(0.42)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)	(0.04)	
Schools with Electricity	0.51	-0.07	-0.08	-0.08	-0.07	-0.07	-0.07	249
	(0.46)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)	(0.04)	
Schools with Drinking Water Source	0.44	-0.04	-0.02	-0.03	0.07	0.11	0.09	249
	(0.45)	(0.06)	(0.06)	(0.04)	(0.06)	(0.06)	(0.04)	
Schools with Boundary Wall	0.69	0.01	-0.09	-0.04	-0.12	0.10	-0.01	249
	(0.40)	(0.06)	(0.07)	(0.04)	(0.06)	(0.05)	(0.04)	
F-Statistic for Joint Orthogonality		0.72	0.83	0.66	0.65	0.72	0.88	

Note: *** p < 0.01, ** p < 0.05, * p < 0.1. Variable used in this table are from the baseline data collection in 2012. Treatment and Control here refer to groups of villages who 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). 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 in English, Mathematics and Sindhi are scaled here using Item Response Theory Model. F-stat for joint orthogonality of each treatment against control is reported in the last row. Grade wise enrollment is also balanced, but not reported for space.

Figure 3A: Experiment profile and timeline

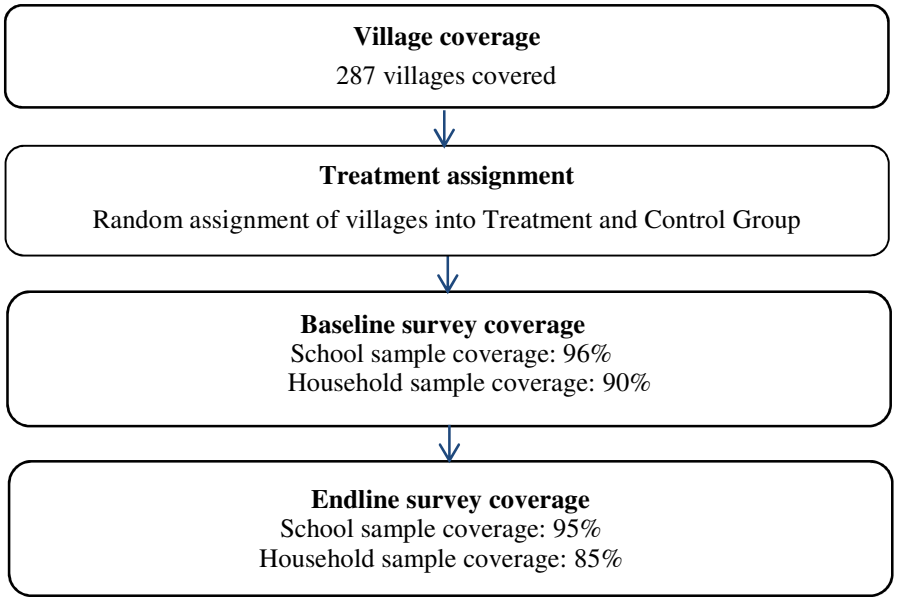


Figure 3B: Sample coverage by treatment

	Control	MEET	MEET-E	SMS	SMS-E	
Baseline	Villages: 57 Schools: 99 Households: 1309 Children: 1070	Villages: 59 Schools: 103 Households: 1415 Children: 1131	Villages: 57 Schools: 92 Households: 1280 Children: 1001	Villages: 57 Schools: 96 Households: 1220 Children: 975	Villages: 57 Schools: 102 Households: 1262 Children: 1091	287 492 6486 5268
Endline	Villages: 57 Schools: 99 Households: 643 Children: 762	Villages: 59 Schools: 102 Households: 625 Children: 738	Villages: 57 Schools: 92 Households: 605 Children: 700	Villages: 57 Schools: 94 Households: 616 Children: 759	Villages: 57 Schools: 102 Households: 635 Children: 747	287 489 3124 3706

Figure 3C: Experiment Timeline

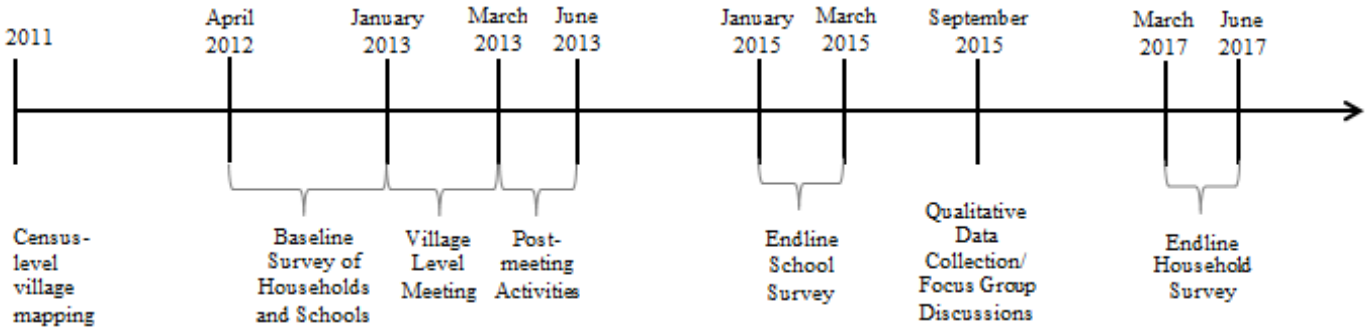


Table 3: Descriptive Statistics at Endline

	Control (1)	Difference (Treatment-Control) MEET (2)	SMS (3)	POOLED (4)	MEET-E (5)	SMS-E (6)	POOLED-E (7)	N
School Characteristics								
Open on unannounced visit	0.80	**0.09***	0.02	0.06	-0.01	0.06	0.03	479
	(0.41)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.03)	
Schools with electricity	0.51	0.05	-0.01	0.02	0.11	0.01	0.05	387
	(0.50)	(0.06)	(0.07)	(0.04)	(0.06)	(0.06)	(0.05)	
Schools with boundary wall	0.70	0.09	-0.04	0.03	-0.03	0.12	0.04	387
	(0.46)	(0.04)	(0.06)	(0.04)	(0.06)	(0.04)	(0.04)	
Schools with drinking water source	0.60	-0.14	-0.07	-0.11	0.08	-0.07	0.00	387
	(0.49)	(0.05)	(0.06)	(0.04)	(0.06)	(0.07)	(0.05)	
Schools with only one teacher	0.52	**-0.19****	**-0.17***	**-0.19****	-0.10	**-0.15***	-0.12	387
	(0.50)	(0.05)	(0.05)	(0.04)	(0.06)	(0.05)	(0.04)	
Total Teachers (Available)	2.25	0.59	0.42	0.51	0.18	0.10	0.14	387
	(2.01)	(0.27)	(0.27)	(0.19)	(0.23)	(0.20)	(0.15)	
Share of Teachers Absent	0.05	**0.06****	0.04	**0.05****	0.03	**0.08****	**0.06****	387
	(0.14)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	
School Size (Enrollment)	70.93	15.99	11.56	13.96	5.22	4.89	5.05	387
	(57.5)	(9.43)	(8.16)	(6.31)	(5.55)	(6.49)	(4.26)	
Student Teacher Ratio	37.25	-2.79	-1.95	-2.40	-1.56	-0.71	-1.13	387
	(17.37)	(2.15)	(2.18)	(1.53)	(1.84)	(2.14)	(1.41)	
Student Classroom Ratio	30.28	3.68	1.09	2.49	3.39	1.92	2.64	387
	(19.21)	(2.17)	(2.48)	(1.63)	(1.90)	(1.89)	(1.34)	
Student Toilet Ratio	58.93	19.74	12.16	16.27	1.90	11.14	6.61	387
	(64.35)	(10.17)	(9.19)	(6.91)	(6.57)	(7.33)	(4.92)	
Test Score (Percent)	0.36	0.00	0.04	0.02	0.00	0.03	0.01	372
	(0.19)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	
Teacher Characteristics								
Female	0.16	0.05	-0.06	0.00	-0.05	0.02	-0.01	387
	(0.19)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	
Education (Mean)	13.13	0.07	-0.19	-0.05	-0.22	0.02	-0.10	387
	(1.22)	(0.11)	(0.16)	(0.09)	(0.15)	(0.13)	(0.10)	
Tenure (Mean)	10.90	0.05	-0.59	-0.25	-0.30	-0.07	-0.18	387
	(7.08)	(0.73)	(0.88)	(0.57)	(0.84)	(0.84)	(0.59)	
Experience (Mean)	21.84	-1.43	-0.98	-1.23	-0.28	-0.76	-0.53	387
	(7.08)	(0.70)	(0.83)	(0.53)	(0.81)	(0.94)	(0.62)	
Teaching Time per day (Mean)	4.23	0.15	-0.02	0.07	0.17	0.06	0.11	387
	(1.63)	(0.20)	(0.23)	(0.15)	(0.16)	(0.18)	(0.12)	
Knowledge Score English %	72.27	1.90	-3.63	-0.46	-3.93	0.52	-1.55	320
	(20.01)	(2.13)	(3.28)	(1.86)	(3.50)	(2.87)	(2.23)	
Knowledge Score Math %	83.51	2.10	-3.07	-0.10	-6.63	1.42	-2.33	320
	(17.38)	(2.04)	(3.34)	(1.85)	(4.17)	(2.14)	(2.30)	
Children Characteristics								
Female	0.46	-0.01	0.01	0.00	-0.03	-0.02	-0.02	157
	(0.10)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	
Age	12.44	0.04	0.02	0.03	-0.17	0.05	-0.06	157
	(0.82)	(0.21)	(0.15)	(0.13)	(0.22)	(0.14)	(0.13)	
Native Language: Sindhi	0.59	-0.03	0.06	0.02	-0.12	-0.02	-0.07	157
	(0.32)	(0.06)	(0.06)	(0.04)	(0.06)	(0.05)	(0.04)	
Enrolled in School	0.50	0.02	0.01	0.02	-0.01	0.02	0.01	157
	(0.16)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	
Years of Schooling Completed	4.37	-0.09	-0.09	-0.09	0.04	-0.19	-0.08	157
	(0.98)	(0.20)	(0.19)	(0.14)	(0.18)	(0.21)	(0.14)	
Test Score (Percent)	0.22	0.00	0.00	0.00	0.00	0.02	0.02	157
	(0.95)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	

Note: *** p < 0.01, ** p < 0.05, * p < 0.1. Variable used in this table are from the endline data collection in 2015 (Schools) and 2017 (Households). Col (1) reports mean and standard deviation for Control group at endline. Col (2) - (7) reports on differences in means between Treatment and Controls. Standard errors clustered at village level are reported in parenthesis. School and Teacher characteristics reported at School level, Children characteristics reported at village level.

Table 4: Treatment Fidelity

	Total	Matiari	Mirpurkhas	Sanghar
Panel A: Participation Rates in Community Meetings				
MEET	67%	59%	77%	58%
SMS	68%	68%	77%	62%
POOLED	67%	63%	79%	60%
MEET-E	58%	42%	69%	54%
SMS-E	58%	34%	76%	52%
POOLED-E	58%	38%	73%	53%
Panel B: SMC Composition Changes				
MEET-E	74%	73%	74%	75%
SMS-E	71%	73%	75%	65%
POOLED-E	73%	73%	75%	71%
Panel C: Average Attendance in Training for SMC				
MEET-E	96%	86%	97%	100%
SMS-E	96%	88%	96%	100%
POOLED-E	96%	87%	97%	100%

Note: Participation rate is defined as proportion of households represented at village meetings out of total number of households in the intervention settlements, as given in the project census data. SMC Composition Changes reflect changes in Executive Body Members only and is calculated out of a total of 4 members eligible for election. Average attendance in election groups is the average number of Executive Body members who attended capacity building meetings out of a total of 5 members which includes head teacher.

Table 5: Transition Probabilities by Treatment Status

Outcome	Control	Difference in transition percentage (Treatment -Control)						Total Transitions
		MEET (1)	SMS (2)	POOLED (3)	MEET-E (4)	SMS-E (5)	POOLED-E (6)	
Panel A: School functioning								
Open to Closed	13.27	**-8.97****	-1.5	-5.25	3.21	-3.56	-0.38	53
Closed to Open	10.20	10.23	**12.14****	**11.19****	-2.51	7.28	2.69	75
Panel B: Inputs								
Added a Teacher	17.81	**17.63****	**11.92***	**14.87****	10.40	11.11	**10.76***	109
Added a Classroom	15.07	7.71	2.50	5.19	6.72	**12.64****	**9.77****	82
Added a Toilet	16.44	5.08	-0.22	2.51	1.51	**5.45***	4.06	74

Note: Sample includes all schools covered by the enumerator during the unannounced visit at the baseline and endline. Out of these, 36 schools were covered in baseline and endline on announced visit. Panel B only reports positive improvements in outcomes.

Table 6: Intention-to-Treat (ITT) estimates for school functioning

	School is open on unannounced visit	School changes from closed to open	School changes from open to closed
	(1)	(2)	(3)
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)
POOLED	0.08	**0.11****	-0.05
	(0.05)	(0.05)	(0.04)
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)
POOLED-E	0.03	0.03	-0.00
	(0.05)	(0.04)	(0.05)
Observations	479	479	479
Adjusted R-squared	0.04	0.07	0.02
Mean of Control	0.79	0.10	0.13

Note: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Sample includes all schools covered by the enumerator during the unannounced visit at the baseline and endline. Estimates control for district fixed effects and flood. Standard errors are clustered at village level.

Table 7: ITT estimates on school inputs (teachers and infrastructure)

	Panel A: Teachers		Panel B: Infrastructure				
	Teachers		Classrooms	Toilets	Electricity	Drinking Water Supply	Boundary Wall
Variables	Registered (1)	Present (2)	(3)	(4)	(5)	(6)	(7)
MEET	0.26	-0.00	-0.05	0.04	0.07	-0.14	0.05
	(0.18)	(0.17)	(0.13)	(0.06)	(0.06)	(0.08)	(0.06)
SMS	**0.30***	0.24	0.23	0.02	-0.00	-0.05	-0.04
	(0.18)	(0.20)	(0.15)	(0.07)	(0.07)	(0.09)	(0.05)
POOLED	**0.28***	0.12	0.09	0.02	0.03	-0.10	0.01
	(0.15)	(0.15)	(0.12)	(0.05)	(0.06)	(0.08)	(0.05)
MEET-E	0.13	0.10	0.04	-0.06	0.11	0.06	0.01
	(0.16)	(0.15)	(0.13)	(0.07)	(0.07)	(0.09)	(0.05)
SMS-E	0.15	0.00	0.19	**0.11***	0.01	-0.08	0.04
	(0.17)	(0.17)	(0.12)	(0.06)	(0.06)	(0.09)	(0.06)
POOLED-E	0.14	0.05	0.12	0.03	0.06	-0.01	0.03
	(0.14)	(0.14)	(0.11)	(0.05)	(0.06)	(0.08)	(0.05)
Observations	387	387	387	387	387	387	387
Adjusted R-squared	0.74	0.61	0.66	0.07	0.31	0.11	0.44
Mean of Control	2.25	1.96	2.18	0.75	0.51	0.60	0.70

Note: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Sample covers all schools that were open on the day of visit (announced round) and that completed all modules of the survey at both baseline and endline. Drinking Water Supply indicates that school has piped water or bore hole or a hand pump. Estimates control for district fixed effects, flood and school size. Standard errors are clustered at village level.

Table 8: ITT estimates on enrollment

			Boys				
	Girls	Boys	Katchi and Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MEET	3.80	0.36	1.85	-1.49	-0.13	0.57	0.75
	(3.11)	(4.39)	(2.89)	(1.18)	(0.94)	(0.75)	(0.80)
SMS	-1.06	**10.06****	3.35	**2.60***	**2.11****	**1.77****	0.51
	(2.53)	(4.83)	(2.60)	(1.47)	(1.03)	(0.89)	(0.66)
POOLED MEET-SMS	1.44	5.09	2.58	0.50	0.96	**1.15***	0.64
	(2.26)	(3.66)	(2.13)	(1.18)	(0.82)	(0.68)	(0.61)
MEET-E	-1.39	3.39	1.65	1.41	-0.66	**1.59***	0.08
	(2.23)	(3.82)	(2.04)	(1.22)	(1.02)	(0.81)	(0.65)
SMS-E	-0.52	5.84	**4.05***	1.54	0.32	0.92	-0.63
	(1.91)	(3.87)	(2.10)	(1.16)	(0.82)	(0.72)	(0.66)
POOLED ELECTIONS	-0.94	4.67	2.89	1.49	-0.15	**1.25***	-0.29
	(1.79)	(3.17)	(1.77)	(1.04)	(0.77)	(0.65)	(0.56)
Observations	387	387	387	387	387	387	387
Adjusted R-squared	0.64	0.69	0.53	0.51	0.53	0.49	0.43
Mean of Control	23.92	47.04	18.93	8.74	7.96	6.01	5.39

Note: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Sample covers all schools that were open on the day of visit (announced visit) and that completed all modules of the survey at both baseline and endline. Estimates control for district fixed effects and flood. Standard errors are clustered at village level.

Table 9: ITT estimates on student learning outcomes

	Panel A: Children tested (school)				Panel B: Children tested (household)			
	Total	English	Math	Sindhi	Total	English	Math	Sindhi
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MEET	0.09	-0.06	0.09	0.11	0.02	-0.00	0.03	-0.05
	(0.16)	(0.13)	(0.13)	(0.15)	(0.10)	(0.10)	(0.09)	(0.08)
SMS	0.15	-0.00	0.06	0.20	-0.09	-0.04	-0.04	-0.13
	(0.17)	(0.15)	(0.14)	(0.16)	(0.10)	(0.09)	(0.09)	(0.09)
POOLED	0.12	-0.03	0.08	0.15	-0.03	-0.02	-0.00	-0.09
	(0.15)	(0.13)	(0.12)	(0.13)	(0.08)	(0.09)	(0.08)	(0.07)
MEET-E	-0.07	-0.22	-0.17	0.11	0.00	-0.04	0.02	-0.01
	(0.18)	(0.16)	(0.14)	(0.16)	(0.08)	(0.09)	(0.08)	(0.08)
SMS-E	0.08	0.03	0.04	0.10	**0.15***	0.12	**0.18****	0.08
	(0.17)	(0.15)	(0.14)	(0.15)	(0.08)	(0.09)	(0.09)	(0.08)
POOLED-E	0.01	-0.09	-0.06	0.11	0.08	0.04	0.11	0.04
	(0.15)	(0.14)	(0.12)	(0.13)	(0.07)	(0.08)	(0.07)	(0.07)
Observations	239	239	239	239	157	157	157	157
Adjusted R-squared	0.06	0.08	0.09	0.03	0.13	0.22	0.19	0.13
Mean of Control	-0.10	0.01	-0.07	-0.13	0.15	0.24	0.14	0.060

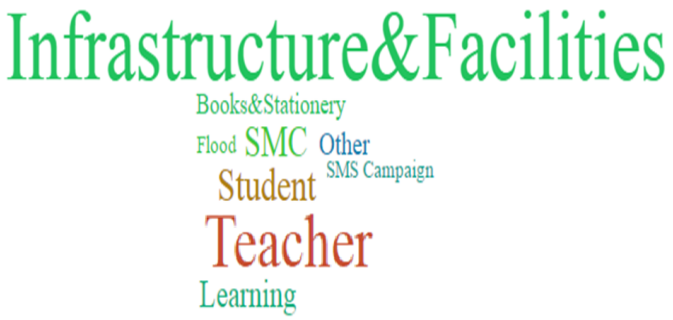
Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Test scores estimated using IRT scale. 24 questions for English, 23 questions for Mathematics and 30 questions for Sindhi considered for this analysis. Panel A covers all villages where students were tested in Endline School Survey and where schools were found open and that completed all modules of the survey at both baseline and endline; totaling to 239 villages. Panel B covers all villages where children were tested in Endline Household Survey. Estimates control for district fixed effects, flood and school size

Table 10: ITT estimates on SMC awareness

	Household is aware of:				
	SMC	SMC has five members	SMC meets at least twice a year	SMC receives PKR 22,000 annually	School Improvement Plan (SIP)
	(1)	(2)	(3)	(4)	(5)
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)
POOLED	0.00	-0.02	-0.00	-0.00	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(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)
POOLED-E	0.01	-0.01	0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	160	160	160	160	160
Adjusted R-squared	0.39	0.29	0.38	0.29	0.12
Mean of Control	0.215	0.178	0.187	0.197	0.107

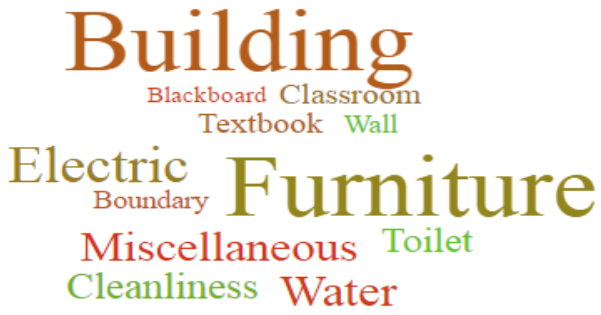
Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Table reports different measures of SMC awareness. Sample includes all villages covered in Endline Household Survey, totaling to 160. Estimates control for district fixed effects. Standard errors are clustered at village level.

Figure 4A: Elicitation of preferences through community dialogue on SMS[17]



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 4B: Expenditure items identified by elected members during training[18]

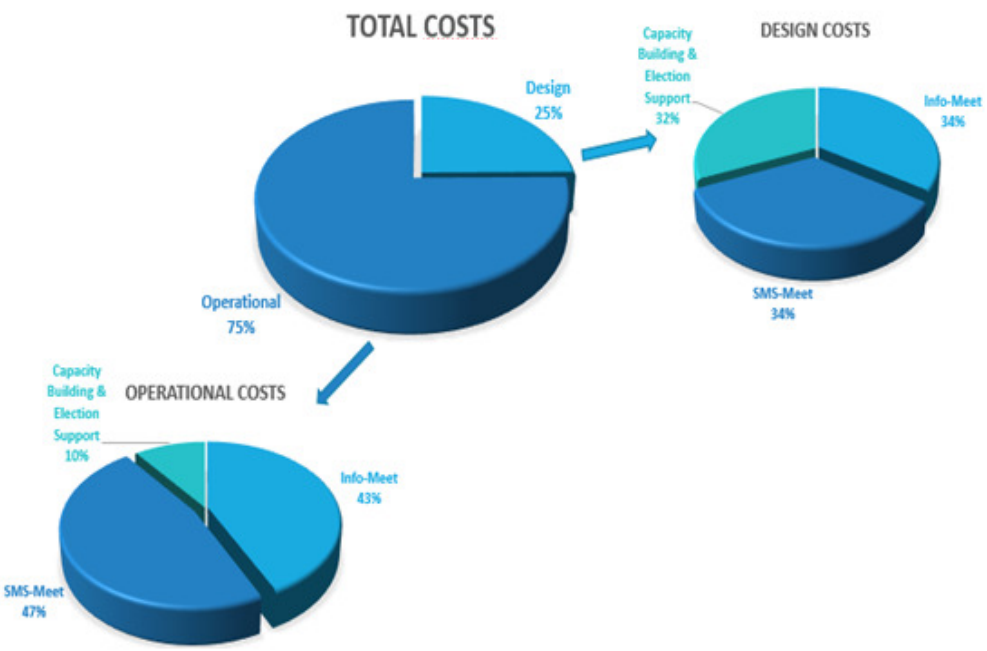


Weights per item mentioned in SIP	
Item	**Weights**
Blackboard	1%
Boundary Wall	3%
Building	24%
Classroom	4%
Cleanliness	6%
Electric	12%
Furniture	24%
Textbooks	4%
Toilet	6%
Water	9%
Miscellaneous	8%

[17] 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 SMC funds, while "SMS Campaign" includes all comments about the CDP itself.

[18] Lists items targeted for investment by EMC Executive Body 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 5: Project Costs



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

Table 11: Cost effectiveness (enrollment)

	MEET	**SMS**
Enrollment in additional open schools (%)	11%	5%
Additional enrollment in already open schools (%)	4%	9%
Cost per 1% increase in enrollment	$9,271	$10,148
Cost per additional enrolled student	$79	$109

Notes: Calculated from operational costs only to estimate scale-up and ongoing cost. Estimates enrollment in additional open schools based on per-group average enrollment at baseline. Proportional increase in enrollment compared to total group enrollment at baseline.

Appendix A: Additional Tables

Table A1: Covariance Balance for restricted Household Sample (baseline)

		Difference (Treatment-Control)						
	Control (1)	MEET (2)	SMS (3)	POOLED (4)	MEET-E (5)	SMS-E (6)	POOLED-E (7)	N
Household Characteristics								
Household is aware of:								
SMC	0.15	0.00	-0.01	0.00	-0.02	0.00	-0.01	160
	(0.09)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
SMC has five members	0.07	0.00	0.00	0.00	-0.01	0.01	0.00	160
	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
SMC meets at least twice a year	0.13	-0.01	-0.01	-0.01	-0.03	0.00	-0.01	160
	(0.09)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	
SMC receives PKR 22,000 annually	0.13	-0.01	-0.02	-0.02	-0.03	-0.01	-0.02	160
	(0.09)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
School Improvement Plan (SIP)	0.06	0.01	0.01	0.01	-0.02	0.01	0.00	160
	(0.07)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	
Test Score (IRT)	0.05	0.11	0.04	0.08	0.05	-0.12	-0.04	155
	(0.45)	(0.09)	(0.09)	(0.06)	(0.09)	(0.08)	(0.06)	

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	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 Mitiari, 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: Comparison of School Characteristics between 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****
Total Girls	23.80 (27.74)	15.32 (17.86)	**8.48****
Total Boys	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

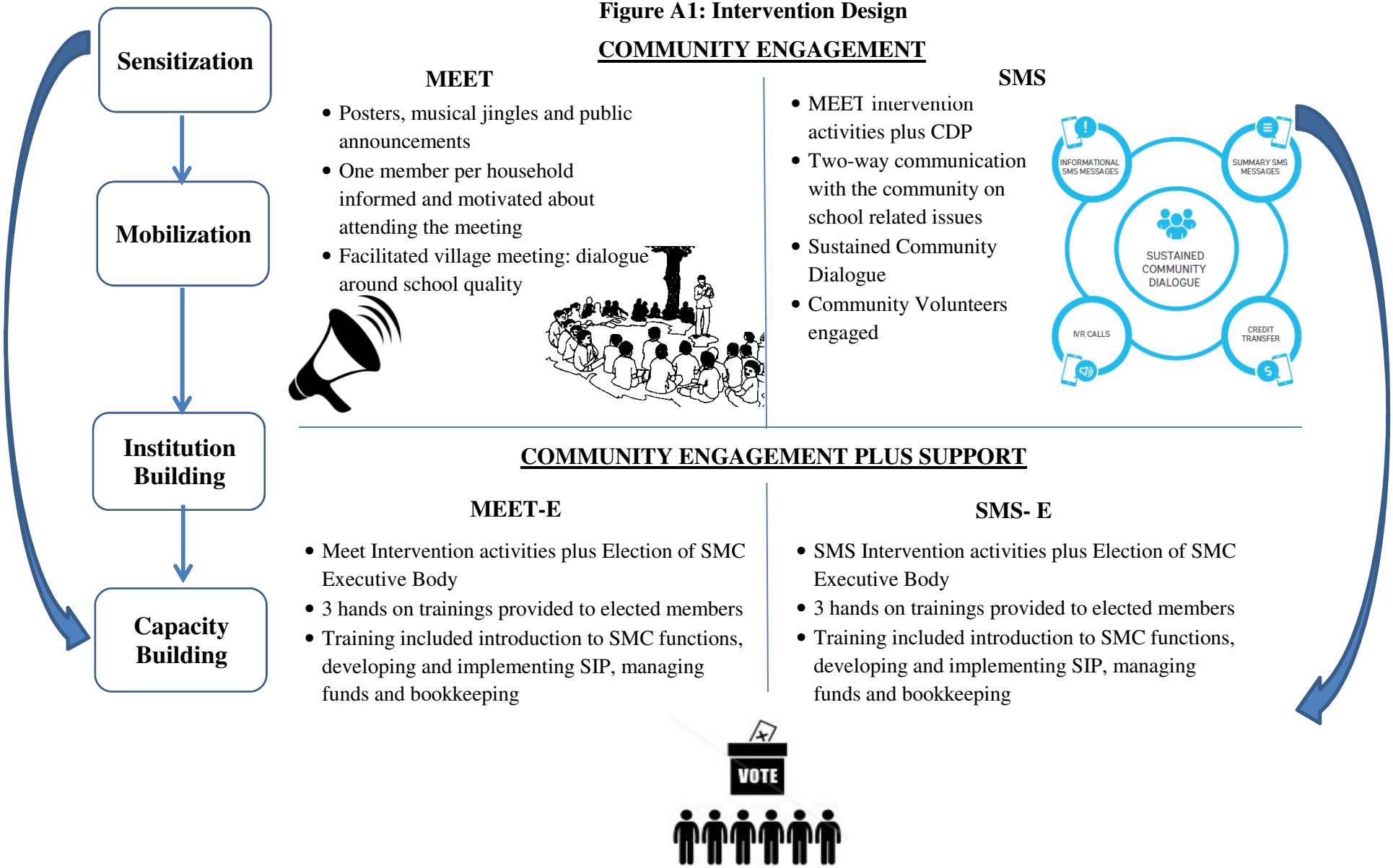
Notes: *** p < 0.01, ** p < 0.05, * p < 0.1. In the columns "Functional Schools" and "Temporarily Closed Schools", mean values given with standard deviations in parenthesis. In the column "Difference in means", difference calculated between 387 schools that were open on the day of visit (announced round) 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: Heterogeneity in Treatment Effect by Community Participation Rate

	School open on unannounced visit	Close to Open	Open to Close	Teachers registered	Teachers present	Classrooms available	Toilets available	Boundary wall exists	Water Source available	Electricity available
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
MEET*P/10	0.02**	-0.01	-0.01**	-0.01	0.00	-0.05**	0.01**	0.01	-0.02	0.01
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
SMS*P/10	0.01	0.01	-0.01	0.00	0.04	-0.01	0.01	-0.01	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
POOLED*P/10	0.01**	-0.00	-0.09*	-0.00	0.02	-0.03*	0.01	-0.00	-0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
MEET-E*P/10	0.01	-0.02**	0.00	-0.01	0.02	-0.03*	-0.00	-0.00	-0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
SMS E *P/10	0.02*	-0.01	-0.01	-0.02	-0.00	-0.02	0.02**	0.00	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
POOLED- E*P/10	0.01	-0.02*	-0.00	-0.02	0.01	-0.03*	0.01	0.00	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	381	381	381	314	314	314	314	314	314	314
Adjusted R-squared	0.03	0.08	0.01	0.72	0.57	0.66	0.20	0.44	0.07	0.04

Note: Robust Standard Errors in Parenthesis. ***p< 0.01, **p<0.05, *p<0.1. The dummy variable for each Treatment is multiplied by participation rate in community-wide meeting and divided by 10. Estimates control for district fixed effects, school size and flood.

Figure A1: Intervention Design



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.[19]

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[20]. 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 SMC 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
5. Compilation of final test

[19] Student tests administered to teachers follow an SDI style knowledge assessment in which teachers are asked to mark a random student's completed exam.

[20] 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.

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