American Economic Journal: Applied Economics 2017, 9(1): 33–69
https://doi.org/10.1257/app.20150027

Experimental Evidence on the Long-Run Impact of Community-Based Monitoring[†]

By Martina Björkman Nyqvist, Damien de Walque, and Jakob Svensson*

We evaluate the longer run impact of a local accountability intervention in primary health care provision in Uganda. Short-run improvements in health care delivery and health outcomes remained in the longer run despite minimal follow-up. We find no impact on the quality of care or health outcomes of a lower cost intervention that focused on encouraging participation but did not provide information on staff performance. We provide suggestive evidence that informed beneficiaries are more likely to identify and challenge (mis)behavior by providers and, as a result, turn their focus to issues that they can manage locally. (*JEL* H75, I11, I18, O15, O18)

Lack of local control and inadequate accountability relationships are often highlighted as underlying determinants of poor quality public service provision in developing countries.[1] Weak central states, coupled with the fact that many of these countries adopted centralized government control over service delivery

---

* Björkman Nyqvist: Stockholm School of Economics, P.O. Box 6501, Sveavägen 65, 113 83 Stockholm, Sweden, and CEPR (e-mail: martina.bjorkman.nyqvist@hhs.se); de Walque: The World Bank, Development Research Group, 1818 H Street NW, Washington, DC 20433 (e-mail: ddewalque@worldbank.org); Svensson: Institute for International Economic Studies, Stockholm University, 106 91 Stockholm, Sweden, and CEPR (e-mail: jakob.svensson@iies.su.se). An earlier version of this paper has been circulated under the title “Information is Power: Experimental Evidence on the Long-Run Impact of Community Based Monitoring.” We are deeply indebted to Frances Nsonzi for her contributions at all stages of the project and acknowledge the important contributions of R. Reinikka, G. Kajubi, A. Ojoo, A. Wasswa, J. Kanyesigye, P. Okwero, R. Chase, C. Winter, I. Njosa, M. Bitekerezo, and the field and data staff with whom we have worked over the years. We thank Andrea Guariso for excellent research assistance. We thank the Uganda Ministry of Health, the World Bank’s Country Office in Uganda, and the Social Development Department, the World Bank, for their cooperation. We are grateful for comments by Tessa Bold, Adam Wagstaff, as well as seminar and conference participants at Dartmouth, University of Michigan, the World Bank, DIAL conference in Paris, and the NEUDC conference at Harvard. We also thank four anonymous referees for constructive suggestions. Financial support from the Bank Netherlands Partnership Program (BNPP), the World Bank Research Committee, the Gender Action Plan (GAP), the Strategic Impact Evaluation Fund (SIEF) at the World Bank, the Swedish Research Council (421-2009-2209), and the Program for Development Research, SIDA, (SWE-2009-053) is gratefully acknowledged. The trial was approved by the ethic committee of Fondazione IRCSS (IRB approval number: D229171). The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank, its executive directors, or the countries they represent.

[†] Go to https://doi.org/10.1257/app.20150027 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] Das, Hammer, and Leonard (2008) find that doctors in Tanzania completed less than 25 percent of the essential checklist for patients with malaria, a disease that is endemic in the country. Indian doctors asked an average of one question per patient (“What’s wrong with you?”). Chaudhury et al. (2006) find an average absence rate of 27 percent among primary school teachers and 37 percent of primary health center staff in Uganda. Bold et al. (2011) find that primary students in urban schools in Tanzania spend about one-quarter of the daily schedule in a classroom with a teacher present. Moreover, roughly half of the primary school teachers in Senegal fail to demonstrate mastery of the curriculum their students are supposed to master, and every other primary health clinician in Senegal is unable to detect a simple case of pneumonia and in total spends about half an hour per day counseling patients.

at independence, in turn, can help explain the existing institutional setup (Duflo, Dupas, and Kremer 2015). In response, policies to enhance community involvement as a way of strengthening demand-responsiveness and local accountability are becoming increasingly popular. Despite the enthusiasm for this approach, however, the evidence provides at best mixed results about its short-run impact. Whether these mixed findings are driven by differences in the details of the interventions or context is difficult to disentangle. Whether these community-driven projects at all manage to influence norms and collective actions, and thus systematically change local decision making in the longer run, is largely unknown and cannot be extrapolated from short-run results, since any intervention putting attention on the public providers might have some positive short-run impact, given the large pre-existing X-inefficiencies in public service provision in many low-income countries.

This paper makes two contributions. First, we provide evidence of the longer run impact of a community driven local accountability project in primary health care provision in Uganda. The *information & participation* intervention combined traditional tools to enhance community participation with the dissemination of report card information on staff performance. Björkman and Svensson (2009) show that the intervention resulted in significant improvements in health care delivery (e.g., on utilization) and health outcomes (e.g., child mortality) after one year. We show here that these large treatment effects on health care delivery and health outcomes remain more than four years after the initial intervention despite minimal follow-up. Thus, properly designed, efforts to stimulate community participation and local control can result in large and sustained improvements in health service provision and health outcomes in both the short and longer run.

Second, we examine the role of information in these kinds of community driven accountability programs. The dissemination of report cards on performance was intended to address informational asymmetries between providers and beneficiaries and drive the reform agenda toward actions the community and the providers could address themselves. However, collecting such data using standard survey methods is a costly and, to some extent, technically complex endeavor and, ex ante, it is unclear if it is worth the cost. To that end we designed an additional intervention (the *participation* intervention), which replicated the participatory components of the original intervention but did not provide communities with baseline information on performance, and compared outcomes across trials

We find that the impact of the cheaper *participation* program after two years differed markedly from the impact of the *information & participation* program—both in the short and longer run. Without information, the process of stimulating participation and engagement had little impact on the health workers' behavior, health outcomes or the quality of health care. For several endpoints, including infant mortality, utilization, and intermediate outcomes, such as extent of monitoring and health treatment practices, albeit not all, we can reject the null hypothesis of equal treatment effects across the two trials.

We investigate why the provision of information appears to have played such a key role by using data from the implementation phases of the two interventions. A core component of both experiments was the agreement of a joint action plan outlining the community's and the providers' agreement on what needs be done, and by whom, in order to improve health care delivery. While the process of reaching an agreement looks similar on some observable measures in the two treatment groups—the same number of community members participated in the community meetings and, on average, the two groups identified the same number of actions to be addressed—the types of issues identified differed significantly. Specifically, in the *participation* group the health provider and the community identified issues that primarily required third-party actions; e.g., more financial and in-kind support from upper level authorities and NGOs. In the *information & participation* group, by contrast, the participants focused almost exclusively on local problems, which either the health workers or the users could address themselves, including absenteeism, opening hours, waiting time, and patient-clinician interactions. These results are consistent with the hypothesis that lack of information on performance makes it more difficult to identify and challenge (mis)behavior by the provider, and hence constrains the community's ability to hold providers to account. That is, with access to information, users are better able to distinguish between the actions of health workers and factors beyond their control and, as a result, turn their focus to issues that they can manage and work on locally.

The results from experimental studies on local accountability provide mixed evidence of impact. Olken (2007) finds that invitations to village-level meetings, where project officials documented how they spent project funds for local road construction in Indonesia, did not result in corruption becoming a salient issue in village meetings and did not lead to a reduction in measured corruption on average. Banerjee et al. (2010) show that a project in India where parents were provided with simple tools to gather information, later shared in village meetings, about their children's learning outcomes did not result in discussions about the teachers' performance or the learning environment and did not prompt increased teacher effort or improvement in educational outcomes. Andrabi, Das, and Khwaja (2015), on the other hand, find that providing households with report cards that displayed test results of their own children and test results from all schools catering to students in the village, raised test scores both in private schools and in public schools with limited market or administrative disciplining mechanisms—a result consistent with the hypothesis that information about relative performance can strengthen parents' ability to hold public providers to account. Duflo, Dupas, and Kremer (2015) show that a governance intervention in a contract teachers program in Kenya that provided parents (school management committees, SMC) specific training on how to monitor and assess teachers' performance, and gave them influence in the decision whether to renew contract teachers' contracts, resulted in significant improvements in learning. Barr et al. (2012) evaluate two community-based monitoring interventions where SMC members were provided with tools to monitor teacher performance. They show that the intervention including an explicit participatory component, combined with tools for assembling performance information, resulted in reduction in teacher absenteeism and increased student test scores—a result broadly consistent with our findings. Taken together, the findings from the two experiments reported here, in light of the existing work, suggest that information provision, or provision of tools to acquire information, is a necessary, albeit not sufficient, condition for these types of programs to have an effect in a poor information environment, but also that the details of the design, including what type of information is provided and how it is provided, matter. Identification of the key design components and mechanisms, and replication of successful interventions in other contexts, are ways to push the knowledge frontier on local accountability forward.

Our findings speak directly to the discussion of the Community Driven Development (CDD) agenda promoted by many donor agencies. A core objective in these programs is to increase community participation in local decision making. Over the past decade, the World Bank alone has allocated close to $85 billion to local participatory development programs. A comprehensive review of those programs finds that few programs are systematically evaluated and that the enthusiasm for participatory initiatives is driven more by ideology and optimism than by analysis, either theoretical or empirical (Mansuri and Rao 2013). We find that the intervention focusing on increasing participation, but not reducing possible informational asymmetries; i.e., the *participation* intervention, had little impact. As this intervention, from a process perspective, mimics many CDD-type projects, our results provide less encouraging news for those promoting greater community participation, especially if the intention is to increase local control of public services. On the other hand, the technically more demanding and more expensive *information & participation* intervention resulted in a more engaged community and in large and long-run improvements in both health service provision and health outcomes. Our results therefore suggest that assembling and disseminating relevant information is crucial in programs designed to strengthen local control and oversight as it enables citizens to focus on actionable tasks. The long-run outcomes, with a reduction in longer run under-five and infant mortality by 23 percent and 28 percent, respectively, further suggest that such an investment in data collection and dissemination, while expensive, may be worth the cost.

Our results also contribute to the literature on institutions in developing countries; a literature that so far has primarily focused on property rights institutions. Here we focus on local institutions for human capital formation. A line of recent research has identified the strength of common interests as a fundamental determinant of institutional quality (Besley and Persson 2011). Although there is little evidence on what creates or destroys common interest, it is conceivable that common interest thrives in a society where norms and institutions constrain rent-seeking. Community participation interventions, primarily meant to (informally) influence norms and collective actions and thereby local decision making, can thus be viewed as one attempt to enhance common interests, albeit starting at a small scale.

Finally, the paper links to a growing empirical literature on the relationship between information dissemination and accountability (Strömberg 2015; Olken and Pande 2012). We depart from most of the literature by focusing on mechanisms through which citizens can make providers, rather than politicians, accountable. Thus, we do not study the design or allocation of public resources across communities, but rather how these resources are utilized.

The next section describes the institutional setting for our study. Section II details our evaluation design, the features of the two interventions, and the data used to evaluate them. The long-run impact evaluation results of the *information & participation* intervention are presented in Section III, while the evaluation results of the *participation* intervention are discussed in Section IV. Section V presents some suggestive findings in support of the key mechanism, relates our findings to the mixed results in the literature, and discusses the policy recommendations of our findings.

I. Institutional Setting

The experiments were implemented in 75 rural communities served by a public primary health facility (or dispensary) in 9 districts covering all 4 regions of Uganda. Dispensaries are in the lowest tier of the health system where a professional interaction between users and providers takes place. Most dispensaries are staffed by six to ten workers—an in-charge or clinical officer (a trained medical worker), nurses, nursing aids and other assistants—and according to the government health sector strategic plan, the standard for dispensaries includes preventive, promotional, outpatient care, maternity, general ward, and laboratory services (Republic of Uganda 2000). Health services should be provided for free.

The health sector in Uganda is decentralized and a number of agents are responsible for supervision and control of the dispensaries. At the lowest tier, the Health Unit Management Committee (HUMC) is supposed to be the main link between the community and the facility. The HUMC should monitor the day-to-day running of the facility but it has no authority to sanction workers. Baseline data reveal, however, that the HUMC is not actively involved in the supervision or support of primary health care providers. The next level in the institutional hierarchy is the Health Sub-district. The Health Sub-district monitors funds, drugs, and service delivery at the dispensary. Supervision meetings by the Health Sub-district are supposed to appear quarterly but, in practice, monitoring is infrequent. The Health Sub-district has the authority to reprimand, but not dismiss, staff for indiscipline. Thus in severe cases of indiscipline, the errand will be referred to the Chief Administrative Officer of the District and the District Service Commission, which are the appointing authorities for the district. They have the authority to suspend or dismiss staff.

The setting for our experimental study—rural Uganda—is characterized by poor public health service provision (large X-inefficiencies). For example, roughly 50 percent of the staff are absent from the clinic on a typical day (based on observational data from unannounced visits); the average waiting time is more than two hours; and only four out of ten patients report that any equipment was used the last time the respondent (or the respondent's child) visited the clinic.

Uganda's under-five mortality rate is lower than the mean in sub-Saharan Africa. However, Uganda is still ranked as one of the 40 countries with the highest rate of child mortality in the world (World Bank 2015).

II. Experimental Design and Data

A. *Overview*

The research project was initiated in 2004 and extended in 2007. In 2004, 50 public dispensaries, and health care users in the corresponding catchment areas, in nine districts covering all four regions in Uganda were identified. These 50 facilities/communities constitute the experimental sample for the *information & participation* trial. In 2007, an additional 25 public dispensaries from the same nine districts were identified. These 25 facilities/communities constitute the experimental sample for the *participation* trial. All 75 project facilities were located in rural areas.

The catchment area or community for each dispensary was defined as the households residing in the five-kilometer radius around the facility. A community in our sample has on average 2,500 households residing within the five-kilometer radius of the clinic, of which 350 live within a one-kilometer radius.

For the *information & participation* experiment, the units (facility/community) were first stratified by location (districts) and then by population size. From each block, half of the units were randomly assigned to the treatment group (25 units) and the remaining health facilities were assigned to the control group. A similar procedure was initiated in 2007 when the project was extended with the *participation* intervention; i.e., after stratifying on location and population size, the 25 new facilities were randomly assigned to a treatment group (13 units) and a control group (12 units).

The end of trial evaluation surveys for both experiments were implemented in the first half of 2009.

Trial sizes were set to detect effects on utilization and child mortality (under-five and infant mortality). The trial sizes were also influenced by logistical and cost constraints and the anticipation of smaller long-run treatment effects. In each community we surveyed approximately 100–110 households (and collected birth and death statistics from approximately 100 under five-year-old children). Thus, the sample for the *information & participation* experiment consists of 50 communities/health facilities and approximately 5,000 households. The sample for the *participation* experiment consists of 25 communities/health facilities and approximately 2,500 households. Mortality and utilization data were available for 2005 and these data were used to estimate study power (Björkman and Svensson 2009). The estimated overall under five-year-old mortality rate (the number of child deaths per 1,000 child-year observations) in the control group in 2005 was 34.1. Assuming 300 child-years of observations in each cluster (three years, 2006–2008; and 100 child observations per year), 50 clusters and an estimated coefficient of variation of 0.15, the *information & participation* intervention has 80 percent [60 percent] power of detecting significant differences at the 0.05 significance level if the intervention reduced mortality by 25 percent [20 percent]. Assuming 200 child-years of observations in each cluster (two years, 2007–2008, and 100 child observations per year), 25 clusters, and a coefficient of variation of 0.15, the *participation* intervention has 80 percent [60 percent] power of detecting a significant difference at the 0.05 significance level if the intervention reduced mortality by 40 percent [32 percent].

The mean utilization in the control group in 2005 was 660 patients visiting the facility per month for outpatient care, with a standard deviation of 175. The *information & participation* intervention therefore has 80 percent [60 percent] power to detect a significant difference at the 0.05 significance level if the intervention increased utilization by 21 percent [17 percent]. The *participation* intervention has

80 percent [60 percent] power to detect a significant difference at the 0.05 significance level if the intervention increased utilization by 29 percent [23 percent].

As a reference, Björkman and Svensson's (2009) evaluation of the short-run impact of the *information & participation* intervention document a 26 percent reduction in under five-year-old mortality, a 32 percent reduction in infant mortality, and a 20 percent increase in outpatient served in the intervention relative the control group after one year.

Due to logistical and budget constraints, the experiments were not designed to pick up differential treatment effects across the two trials.

B. *Interventions*

The aim of the research project was two-fold. First, to evaluate the long-run impact of the community monitoring intervention initiated in 2004 (*information & participation* intervention). Second, to evaluate the impact of a cheaper community monitoring intervention, without report cards (*participation* intervention).

Efforts like the CDD approach operate on the principles of local empowerment and participatory governance as mechanisms to strengthen demand responsiveness and local accountability. The core of the strategy is the process through which problems and constraints are identified and how (local) decisions are made and executed. While there are variations across projects, in practice community driven development is achieved though facilitated meetings. Both the *participation* intervention and the *information & participation* intervention largely followed this approach, with the key difference being that the *information & participation* intervention also included the dissemination of report cards on the health clinic's performance in various dimensions.

The research design allows us to estimate and compare three treatment effects, holding the context; i.e., health care provision in rural Uganda, constant:

- the short-run treatment effect of the *information & participation* intervention (reported in Björkman and Svensson 2009);
- the longer-run, i.e., four years after the initial intervention, treatment effect of the *information & participation* intervention; and
- the short-run treatment effect of the *participation* intervention.

The design is illustrated in Figures 1 and 2. Note that while the long-run evaluation covers the period 2005–2009, the short-run treatment effects are estimated over two consecutive time periods: 2005 for the *information & participation* intervention, and 2007–2009 for the *participation* intervention.

The Participation Intervention.—The *participation* intervention involved three types of meetings facilitated by staff from local community-based organizations (CBOs): (i) a community meeting—a two-day afternoon meeting with on average more than 150 participants from all spectra of the catchment area per day and per community attending; (ii) a health facility meeting—a half-day event with all staff attending; and (iii) an interface meeting—a half-day event with representatives from

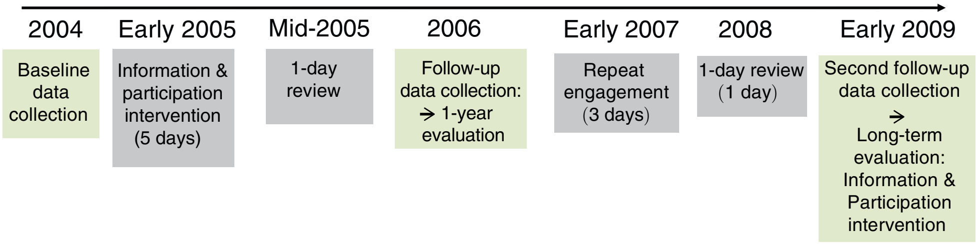


Figure 1. Timeline for the Information & Participation Intervention

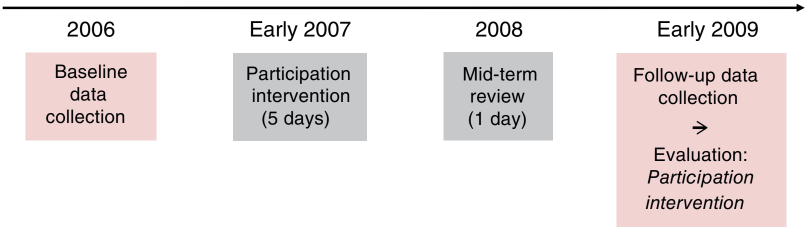


Figure 2. Timeline for the Participation Intervention

the community and the staff attending.[2] For the community meeting, and to avoid elite capture, the facilitators from the CBOs invited about 100 participants representing different spectra of society (i.e., young, old, disabled, women, mothers, leaders).

The main objective of these meetings was to encourage community members and health facility staff to develop a shared view of how to improve service delivery and monitor health provision in the community; i.e., to agree on a joint action plan or a community contract. In practice, the process of reaching an agreement was achieved in two steps. First, in the community and health facility meetings, using various participatory methods (e.g., focus group discussions, community score cards, and role plays), the community and the health staff were asked independently to identify and prioritize the key problems and issues they viewed as the most important to address and how to address them within the current resource envelope.[3] Second, in the facilitated interface meeting, representatives from the community and health facility staff presented and discussed their suggestions and agreed on a set of issues that were viewed as most critical to address. These issues were put into a joint action

---

[2] Altogether, 18 CBOs, each receiving 10 days of training, implemented the two interventions.

[3] Focus group discussions were held with subgroups (young, women, etc.) in the community meetings. The intention was to let each group voice its concerns so as to reduce the risk of elite capture. Community score cards are used as a method to identify issues to be addressed. In the community meeting, members scored the services provided by the facility on a scale from 0 to 100. In the health facility meeting, a similar exercise was implemented to identify key problems and constraints as viewed by the health staff, including a self-assessment of their performance. Role plays were primarily used in the interface meeting as a method to both illustrate and desensitize issues for which the provider and the community had differential views.

plan. The action plan outlines the community's and the providers' joint agreement on what needs be done to improve health care delivery, how, when, and by whom.

In total, the process of reaching an agreement took five days. The meetings were intended to kick-start a community-led process of monitoring. Thus, after the initial meetings, the communities themselves had the responsibility to monitor the implementation of the issues outlined in the joint action plan. In 2008, about a year after the initial intervention, there was a one-day repeat engagement on a smaller scale facilitated by the CBOs during which health workers and community members discussed suggestions for sustaining or improving progress on the issues outlined in the joint action plan.

The Information & Participation Intervention.—The *information & participation* intervention mirrored the participation intervention with one important exception. At the start of the community and the health facility meetings, the facilitators provided the participants with easily accessible quantitative data on the performance of the health provider. These data were collected from facility and household surveys implemented prior to the intervention. A unique report card was established for each facility summarizing information that was identified from the baseline data as key areas subject to improvement, including utilization, access, absenteeism, and patient-clinician interaction. The report cards also included comparisons vis-à-vis other health facilities and with the national standard for primary health care provision. The report cards were translated into the languages spoken in the community and posters were designed to help the nonliterate process the provided information. The aim of the report card was to address informational asymmetries between the providers and beneficiaries and help guide a discussion focused on issues that potentially could be dealt with locally. Following the dissemination of the report cards, the meetings had the same content and format as the meetings organized in the *participation* group.

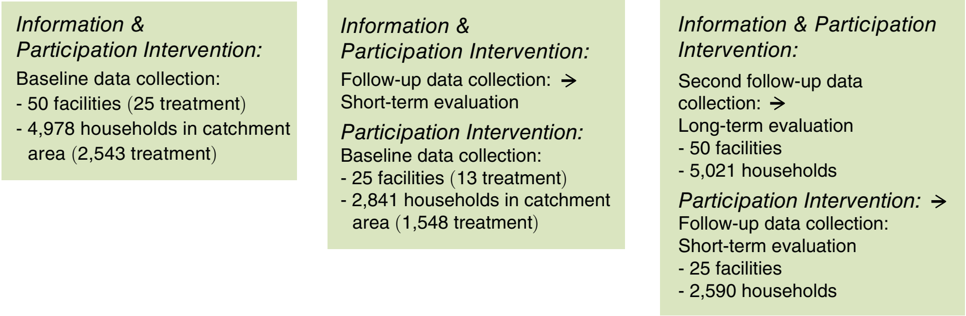
The *information & participation* intervention was initiated at the beginning of 2005 and was followed by small scale repeat (one day) engagements in mid-2005, 2007, and 2008 as illustrated in Figure 1. In between this fairly "minimal" intervention (a total of ten days in four years), the communities were left to themselves to monitor the providers as outlined in the agreed upon action plan.[4]

C. *Data*

Data collection was governed by two objectives. The first objective was to create report cards for the *information & participation* intervention on staff performance

---

[4] In 2007, the CBOs facilitated a community meeting, a health facility meeting, and an interface meeting (each for a few hours in the afternoon for three consecutive days) to discuss the joint action plan. No new data was provided and the CBOs were not asked to assist in revising the plan. Instead the health facility staff and community members jointly discussed suggestions for sustaining or improving progress based on the old plan, or whatever plan the actors has agreed upon at the time. A year after in 2008, as well as midway through 2005, an interface meeting was organized (a couple of hours in the afternoon). Again, no new data was provided and the CBOs facilitated a meeting around whatever joint plan that existed. All in all over the period 2006–2009, there were four afternoon meetings, so about four to five hours of administered meetings per year.



Information & Participation Intervention:
Baseline data collection:
- 50 facilities (25 treatment)
- 4,978 households in catchment area (2,543 treatment)

Information & Participation Intervention:
Follow-up data collection: → Short-term evaluation

Participation Intervention:
Baseline data collection:
- 25 facilities (13 treatment)
- 2,841 households in catchment area (1,548 treatment)

Information & Participation Intervention:
Second follow-up data collection: →
Long-term evaluation
- 50 facilities
- 5,021 households

Participation Intervention: →
Follow-up data collection:
Short-term evaluation
- 25 facilities
- 2,590 households

Figure 3. Data Collection

and how the community views the quality and efficacy of service delivery. The second objective was to rigorously evaluate the short- and long-run impacts.

To meet these objectives, two types of surveys were implemented: a survey of health care providers and a household survey in the catchment areas of the facilities. A quantitative service delivery survey was used to collect data from the providers. Because health providers may have a strong incentive to misreport key data, the data were obtained directly from the records kept by facilities for their own need (i.e., daily patient registers, stock cards, etc.) rather than from administrative records. The former, often available in a highly disaggregate format, were considered to suffer the least from any incentive problems in record keeping. Data were also collected through visual checks by enumerators, including measuring absenteeism using unannounced visits.

Figure 3 provides details on the timing and scope of the data collection effort. The posttreatment survey collected data from 75 health facilities and roughly 7,500 households. While all surveys included a core set of modules, including socio-demographic characteristics, households' health outcomes and health facility performance as experienced by the household in the household survey, additional modules were added in the later rounds. For example, the household survey in 2006 included modules on child mortality and anthropometric measurements (height and weight of infants). The household survey in 2009 also included an additional module on women's health, including prenatal and antenatal care, and collected detailed birth and death statistics for children from 2006 and onward.

We use information referring to the end of the trial period, or the last year of the trial period, for most outcomes. For mortality outcomes, however, we use information referring to the whole period of the experiment (2006–2009 for the *information & participation* intervention, and 2007–2009 for the *information & participation* intervention). In addition to facility and household surveys, we have information from the action plans for the two intervention groups: in 2005 and 2007 (updated) for the *information & participation* intervention, and in 2007 for the *participation* intervention.

D. *Outcomes and Statistical Framework*

We divide our empirical investigation into two parts. We start by analyzing the impact on the main outcome of interest; i.e., whether the intervention resulted in improved health outcomes. We have four sets of health measures: child mortality, pregnancy, birth, and anthropometric measurements of children. We then turn to the quantity of health care. To measure utilization, we use data collected at the health facility (daily patient registers) as well as utilization data from the household survey.

The second part focuses on the channels through which the intervention may have influenced health outcomes and the demand and supply of health care. That is, we assess changes in all steps in the accountability chain; i.e., whether there is evidence of increased monitoring activities by the community, using information derived from direct observations at the facility and data collected at the household level, and whether there is evidence that the health facility staff responded by improving treatment practices and overall management of the health clinics.

We use the same set of outcome measures as those used in the short-run evaluation of the *information & participation* intervention (see Björkman and Svensson 2009). In addition, we report the findings on the extent to which clinical guidelines for the examination of patients were followed, with a focus on antenatal and postnatal care.

We start by assessing the causal effect of each intervention; i.e., we compare mean outcomes after accounting for stratification, by estimating

$$y_{ijd} = \alpha + \beta T_{jd} + \theta_d + \varepsilon_{ijd}, \tag{1}$$

where y_{ijd} is the outcome of household i (when applicable), in community/health facility j, in district d; T_{jd} is an indicator variable for assignment to treatment; θ_d are district fixed effects,;and ε_{ijd} is an error term.

For some outcomes, we have a group of related outcome measures. To assess the impact of the intervention on a set of K related outcomes, we follow Kling et al. (2004) and estimate a seemingly unrelated regression system, and derive average standardized treatment effects, $\tilde{\beta} = \frac{1}{K}\sum_{k=1}^{K}\frac{\hat{\beta}_k}{\hat{\sigma}_k}$, where $\hat{\beta}_k$ is the point estimate on the treatment indicator in the kth outcome regression and $\hat{\sigma}_k$ is the standard deviation of the control group for outcome k (see Duflo, Glennerster, and Kremer 2007).

The experiments were not designed to pick up differential treatment effects across the two trials. However, it is still interesting, although such a test is underpowered and the test results should be interpreted accordingly, to compare outcomes across the trials. To do so we employ a simple difference-in-differences approach; i.e., we estimate

$$\begin{aligned} DD = {} & \big[E(y|T = 1, I \,\&\, P) - E(y|T = 0, I \,\&\, P)\big] \\ & - \big[E(y|T = 1, P) - E(y|T = 0, P)\big], \end{aligned} \tag{2}$$

where I & P denotes *information & participation* experiment and P denotes the *participation* experiment, and T is, as before, an indicator variable for assignment to treatment (in any of the two experiments). The DD specification can be written in regression notation as

$$y_{ijde} = \pi_1 + \pi_2 X_e + \pi_3 T_{jd} + \pi_4 X_e \times T_{jd} + \theta_{de} + \varepsilon_{ijde}, \tag{3}$$

where y_{ijde} is the outcome of household i (when applicable), in community/health facility j, in district d, for experiment e; θ_{de} are district fixed effects specific for each experiment; and ε_{ijde} is an error term. The key identifying assumption in the DD specification is that any time-variant factor, except for the treatment under study, affect the treatment and control group in the same way for a given experiment; i.e., $E(\varepsilon | E \times T) = 0$. Under that assumption, we can test whether the treatment effects differ across trials by testing the null hypothesis that $\pi_4 = 0$. To improve precision of the DD model, we add a vector V of pre-intervention facility-specific covariates (number of households in the catchment area, average number of patients visiting the facility per month for outpatient care, and the average user-fee for general outpatient service) to equation (3).

III. Results

A. *Balance at Baseline*

Tables 1 and 2 report mean pretreatment characteristics for the treatment and control groups in the *information & participation* intervention (using data from 2004), and the *participation* intervention (using data from 2006), and test statistics for equality of means in the two trials, respectively. We report both mean differences for a set of key outcome variables and differences in average standardized pretreatment effects for each family of outcomes (utilization, utilization pattern, quality, catchment area statistics, health facility characteristics, citizen perceptions, supply of drugs, and user charges). Overall the sample is balanced in both trials.

As reported in Table A1 in the Appendix, the two control groups (compared in the same year) also appear similar on most observable characteristics.

B. *Health Outcomes and Quantity of Care*

Health Outcomes.—The primary outcome measure for the trial was child mortality. We also consider four other health outcomes: number of births and pregnancies, and height and weight of children. Child mortality links to the quality and the quantity of a wide spectrum of services that should be provided by dispensaries. Many of these services, including improved prenatal care, health education, child preventive care, and child curative care, also have the potential to affect height and weight. Height can be viewed as a cumulative measure of health and nutrition since conception, while weight can be viewed as an indicator of current nutrition and illness status, especially for younger children. Number of births and pregnancies link directly to the quantity and the quality of family planning and health education.

Panel A in Table 3 depicts the longer run (2006–2009) findings on child mortality for the *information & participation* intervention. We start by reporting the results using the raw data; i.e., the number of under-five, infant (under 12 months), and neonatal (under 1 month) deaths per year. The *participation & information* intervention reduced the number of deaths in all three age categories. The number of under five-year-old deaths dropped by 32 percent (column 1); the number of infant deaths

Table 1—Pre-intervention Characteristics in the Information & Participation and Participation trials

	Panel A. Information & participation (2004)			*Panel B. Participation (2006)*		
Variables	Intervention (1)	Control (2)	Difference (3)	Intervention (4)	Control (5)	Difference (6)
Outpatient care	593	675	−82	781	790	−9.2
	(373)	(286)	(94)	(301)	(330)	(127)
Delivery	10.3	7.5	2.8	11.9	14.7	−2.8
	(11.1)	(6.8)	(2.6)	(6.1)	(11.1)	(3.6)
Number of households in catchment area	2,140	2,224	−84.4	2,850	2,519	331
	(927)	(1,021)	(276)	(1,218)	(1,144)	(472)
Number of households per village	93.9	95.3	−1.42	121.9	118.1	3.75
	(26.4)	(31.6)	(8.23)	(38)	(41.7)	(16.0)
Access to safe water[a]	0.40	0.32	0.08	0.92	1.00	−0.08
	(0.50)	(0.48)	(0.14)	(0.28)	(0.00)	(0.08)
Access to electricity[a]	18.3	20.4	−2.12	0.46	0.42	0.04
	(14.8)	(14.5)	(4.14)	(0.52)	(0.51)	(0.21)

Notes: In each panel and row we report catchment area/health facility averages (columns 1 and 2, 4 and 5) and the difference in averages (columns 3, 6). Standard deviations are reported in parentheses below the averages and robust standard errors are reported in parentheses under the differences. Columns 1–3 report baseline characteristics for the *information & participation* experiment in 2004. Columns 4–6 reports baseline characteristic for the *participation* experiment in 2006. Description of variables: Outpatient care is average number of patients visiting the facility per month for outpatient care. Delivery is average number of deliveries at the facility per month. Number of households in catchment area and number of households per village are based on census data and Uganda Bureau of Statistics maps. Access to safe water is an indicator variable for whether the health facility staff at the time of the pre-intervention survey reported they could safely drink from the water source (in panel A) and an indicator variable for whether the health facility in 2006 reported to have a functioning water source (in panel B). Access to electricity is the number of days without electricity in the month prior to pre-intervention survey (in panel A) and an indicator variable for whether the main source of electricity at the health facility worked as assessed during the baseline survey (panel B).

[a] The definition of the variable differs between the 2004 and 2006 baseline survey due to differences in survey instruments.

dropped by 36 percent (column 2); and the number of neonatal deaths dropped by 51 percent (column 3).

The crude death numbers provide evidence of a fall in child mortality. The reduction may not necessarily be due solely to a reduction in the risk of child death, however, as cohort sizes may have been differentially affected by the intervention (for instance due to differential fertility rates). To account for this, we estimate mortality rates over the period of exposure; i.e., between January 2006 and May 2009. We follow the conventional approach used in epidemiology and define the under-five [infant] mortality rate as the number of under-five [under 12 months] children that died during the period per 1,000 child-years [infant-years] of exposure over the same time period. We also, conventionally, define neonatal mortality as the number of neonatal deaths per 1,000 live births. The results, with the data collapsed at the cluster level, are reported in columns 4–6. For under five-year-old mortality and infant mortality, we report both treatment effects from a linear regression, equation (1), and the rate ratio using a Poisson regression model (Preston 2005) of the incidence of child deaths (occurrence of death over child-months) in the treatment versus the control group.

TABLE 2—AVERAGE STANDARDIZED PRE-INTERVENTION EFFECTS IN 2004 AND IN 2006

Variables	*Panel A.* *Information & participation* (*2004*) Difference between intervention and control groups	*Panel B.* *Participation* (*2006*) Difference between intervention and control groups
Utilization from health facility records	0.07 (0.27)	−0.15 (0.28)
Utilization pattern of the users	−0.03 (0.03)	−0.04 (0.04)
Quality of services according to users	−0.04 (0.09)	−0.02 (0.16)
Catchment area statistics	0.06 (0.18)	0.45 (0.40)
Health facility characteristics[a]	0.07 (0.09)	−0.17 (0.19)
Citizen perceptions of treatment	0.04 (0.05)	0.04 (0.08)
Supply of drugs	0.17 (0.25)	0.47 (0.43)
User charges	−0.46 (0.32)	−0.34 (0.33)

Notes: In each panel and row we report an average standardized pre-intervention (AST) effect (see text for details). Robust standard errors (for health facility variables) and robust standard errors clustered by catchment area (for household level variables) are reported in parentheses. Column 1 reports AST effects using 2004 baseline data for the *information & participation* experiment and column 2 reports AST effects using 2006 baseline data for the *participation* experiment. Description of variables: Utilization from health facility records summarizes outpatients and deliveries. Utilization pattern of users summarizes use of the project facility, an NGO facility, a private-for-profit facility, other government facility, another provider, a traditional healer and self-treatment, reversing sign of traditional healer and self-treatment. Quality of services summarizes the use of any equipment during the household's last visit to the clinic and waiting time, reversing sign of waiting time. Catchment area statistics summarizes the number of households in the catchment area, the number of households per village, and the distance from the villages to the health facility. Health facility characteristics summarizes presence of piped water, access to a radio, access to a newspaper, the existence of a separate maternity unit, the distance to the nearest Local Council I, the distance to the nearest public health provider, number of staff with advanced A-level education, number of staff with less than A-level education, access to safe water and days without electricity, reversing the sign of days without electricity and distance to nearest local council (in panel A); and presence of piped water, access to safe water, access to electricity, yellow star certification of the health facility, number of staff with advanced A-level education, and number of staff with less than A-level education (in panel B). Citizen perceptions of treatment summarizes whether the staff was polite, whether the staff was attentive, and whether the patient could freely express herself. Supply of drugs summarizes the availability of erythromycin, chloroquine, septrine, quinine, and mebendazole at the facility. User charges summarizes whether the patient needs to pay for medicine, general treatment, injections and deliveries, reversing all signs.

[a] The definition of the variable differs between the 2004 and 2006 baseline survey due to difference in survey instruments.

The estimated rate ratio; i.e., the ratio of the incidence of child deaths in the treatment relative the control group, implies a 23 percent reduced risk of under five-year-old deaths in the treatment relative the control group (column 4). The effect is of the same order of magnitude, but less precisely estimated, using a linear model (a reduction of 4.43 deaths per 1,000 child-years from a control group mean of 18.7 deaths per 1,000). The reduction in infant mortality (column 5) is slightly larger—a 28 percent reduced risk of infant deaths in the treatment versus the control group—but also less precisely estimated. The number of children dying before reaching one month (per 1,000 live births) is 15.2 in the intervention group and 27.2 in the control group, which corresponds to a reduction in neonatal mortality of close to 44 percent (column 6).

Table 3—Program Impact on Child Mortality

Dependent variable	Number of under-five deaths per year	Number of infant deaths per year	Number of neonatal deaths per year	Under-five mortality rate	Infant mortality	Neonatal mortality rate
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Information & participation						
Program impact	−0.38 (0.20)	−0.31 (0.17)	−0.26 (0.10)	−4.43 (2.60)	−12.3 (7.92)	−11.5 (5.4)
Rate ratio				0.77 (0.10) [0.04]	0.72 (0.13) [0.08]	
Mean control	1.20	0.87	0.51	18.7	45.0	27.3
Observations	50	50	50	50	50	50
Panel B. Participation						
Program impact	0.05 (0.39)	0.09 (0.25)	−0.02 (0.15)	−1.87 (4.22)	−3.16 (10.1)	−10.0 (8.38)
Rate ratio				0.87 (0.21) [0.56]	0.98 (0.21) [0.92]	
Mean control	1.14	0.79	0.36	14.8	38.6	21.4
Observations	25	25	25	25	25	25
Panel C. Comparisons						
Short-run impact (*I&P*)	−1.12	−0.67	—	0.74	0.68	—
p-value, *I&P* (2005) versus *P* (2007–2009)	0.15	0.03	—	0.27	0.03	—
p-value, *I&P* (2005–2009) versus *P* (2007–2009)	0.13	0.04	—	0.36	0.10	—

Notes: Each column in each panel presents the results of a separate OLS regression, equation (1), with district fixed effects. Robust standard errors are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Specifications 4–5 also report rate ratios derived from a Poisson model with district fixed effects, robust standard errors in parentheses, and *p*-values reported in brackets. Panel A: *Information & participation* experiment, data for 2006–2009. Panel B: *Participation* experiment, data for 2007–2009. Specifications: (1) number of under five-year-old deaths per year; (2) number of infant deaths per year; (3) number of neonatal deaths per year; (4) number of under five-year-old deaths per 1,000 child-years of exposure to the risk of deaths; (5) number of infant deaths per 1,000 child-years of infant exposure to the risk of deaths; (6) number of neonatal deaths per 1,000 live births. Panel C: Short-run impact (*I&P*) reports treatment effect for the *information & participation* intervention after one year (2005) with rate ratios for under-five and infant mortality in columns 4 and 5. *p*-values, *I&P* (2005, 2005–2009) versus *P* (2007–2009) are *p*-values on the test of the null hypothesis of equal treatment effects, based on equation (3), between the *information & participation* intervention (*I&P*) (short run 2005 or long run 2005–2009) and the *participation intervention* (*P*) (2007–2009). Data on neonatal mortality was not collected in 2006.

The first row of panel C reports the estimates from the short-run evaluation of the *information & participation* intervention. As evident, the reduction in child mortality in the longer run is similar to the short run findings (rate ratios of 0.74 and 0.68 for under-five and infant mortality in the short run and 0.77 and 0.72 in the longer run).[5]

Panel B summarizes the effect of the *participation* intervention on child mortality. The control group means are similar to those reported in panel A for the *information*

---

[5] Björkman and Svensson (2009) did not collect data on the month of death. Thus, to derive this rate ratio we assume that children that died in 2005 died midway through the year. In Björkman and Svensson (2009) the under five-year-old mortality rate is estimated by summing the death rates for each cohort (0–1 year old, 1–2 year olds etc.) per community. This life-table approach produces a probability of death rather than an incidence of child death as reported in Table 3. The short-run evaluation survey was not designed to measure neonatal mortality.

& participation intervention, but the treatment effects are small and insignificant. The number of under-five (column 1), infant (column 2), and neonatal (column 3) deaths per year are essentially the same in the treatment and control group, and the estimated rate ratios imply low, and insignificant reductions in child mortality in the treatment relative the control group.

Panel C, last two rows, reports p-values on the null hypothesis that the treatment effects in the two experiments are equal, based on equation (3). That is, we test for differential treatment effects by estimating equation (3) and report the p-value on null hypothesis that $\pi_4 = 0$. We compare both the short-run impacts of the *information & participation* intervention (using data for 2005) and the *participation* intervention (using data for 2007–2009) and the longer run impact of the *information & participation* intervention (using data for 2005–2009) versus the *participation intervention* (using data for 2007–2009). Despite the low power of these tests, we can reject the null hypothesis of equal treatment effects both in the short run (p-value $= 0.03$) and longer run (p-value $= 0.098$) for infant mortality. The p-values on the null hypothesis of equal treatment effects for under five-year-old mortality ranges from 0.27–0.36.

Table 4 reports the findings on births and pregnancies. We observe a reduction in fertility in the treatment group relative the control group, independent of outcome measure, in the *information & participation* experiment (panel A). The number of births (per year and household) fell by 14 percent relative to the control group (panel A, column 1)—an effect roughly twice as large in relative terms as compared to the short-run effect reported in Björkman and Svensson (2009). The effect is of the same order of magnitude—12 percent reduction—for the number of pregnancies (panel A, column 2).

Panel B of Table 4, reports the results for the *participation* experiment. The treatment effects are small and insignificant, although we cannot reject the null hypothesis of equal treatment effects in the two trials (panel C, last two rows).

Table 5 reports the results on weight and height of children under five years old. We measured the weight of children under five years old using portable scales and the height of children under five years old using stadiometers. Panel A reports the longer run impacts of the *information & participation* intervention. We report results separately for infants (0–12 months) and children (13–59 months) to capture both the potentially shorter and longer run effects of improved preventive and curative care. Columns (1–2) depict the weight-for-age z-score results. Consistent with the findings in Cortinovis et al. (1997), Ugandan children have values of weight far lower than the NCHS/CDC international reference (see Figure 4).[6]

The treatment effect is 0.22 z-score in weight-for-age for infants and is fairly precisely estimated. For older children (aged 13–59 months), the treatment effect is essentially 0. For height-for-age we observe the opposite. The treatment effect on height-for-age for infants is close to 0 and insignificant, while children aged

---

[6] Following Björkman and Svensson (2009), we exclude z-scores $> |4.5|$ as implausible and omit observations with a recorded weight above the ninetieth percentile in the growth chart reported in Cortinovis et al. (1997). Since weight is measured by trained enumerators, the reporting error is likely due to misreported age of the child.

TABLE 4—PROGRAM IMPACT ON BIRTHS AND PREGNANCIES

Dependent variable Specification	Number of births (1)	Number of pregnancies (2)
Panel A. Information & participation		
Program impact	−0.02	−0.02
	(0.01)	(0.01)
Mean control group	0.14	0.17
Observations	50	50
Panel B. Participation		
Program impact	−0.00	0.00
	(0.02)	(0.03)
Mean control group	0.22	0.26
Observations	25	25
Panel C. Comparisons		
Short-run impact (*I&P*)	−0.02	—
p-value, *I&P* (2005) versus *P* (2007–2009)	0.40	—
p-value, *I&P* (2005–2009) versus *P* (2007–2009)	0.18	—

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention, data for 2006–2009. Panel B: *Participation* intervention, data for 2007–2009. Specifications: (1) average number of births per year and household in the catchment area; (2) average number of pregnancies per year and household in the catchment area. Panel C: Short-run impact (*I&P*) reports treatment effects for the i*nformation & participation* intervention after one year (2005). *p*-values, *I&P* (2005, 2005–2009) versus *P* (2007–2009) are *p*-values on the test of the null hypothesis of equal treatment effects, based on equation (3), between the *information & participation* intervention (*I&P*) (short run 2005 or long run 2005–2009) and the *participation* intervention (*P*) (2007–2009). Data on number of pregnancies was not collected in 2006.

13–59 months in the treatment group in the *information & participation* trial are 0.10 z-scores taller than the children of the same age in the control group.

As higher height-for-age z-scores can be viewed as a cumulative measure of improved health and nutrition since conception for older children, while weight can be viewed as an indicator of current nutrition and illness status (especially for younger children), these results are consistent with the mortality findings (reported in Table 3): health care provision in the treatment group (in the *information & participation* trial) experienced an improvement in the short run that has been sustained throughout the trial period. Figure 5 illustrates the effect by plotting the treatment effect on height-for-age conditional on trial exposure.

Panel B reports the results for the *participation* intervention. The treatment effects on weight are positive and marginally significant (p-value $= 0.081$ and 0.082, respectively) for both infants and older children—findings that are difficult to reconcile with the results reported in Tables 3 and 4—but small and insignificant for height.

Utilization.—Table 6 and Appendix Tables A2 and A3 summarize the findings on utilization using data collected at the health facility. Cross-section estimates based

Table 5—Program Impact on Weight and Height of Children

Dependent variable	Weight-for-age z-scores		Height-for-age z-scores	
	(0–12m)	(13–59m)	(0–12m)	(13–59m)
Specification	(1)	(2)	(3)	(4)
Panel A. Information & participation				
Program impact	0.22	−0.01	0.04	0.10
	(0.10)	(0.07)	(0.10)	(0.05)
Mean control	−0.70	−0.92	−1.27	−1.50
Observations	669	2,839	565	2,451
Panel B. Participation				
Program impact	0.20	0.11	0.04	0.03
	(0.11)	(0.06)	(0.21)	(0.08)
Mean control group	−0.66	−0.90	−1.13	−1.59
Observations	423	1,643	368	1,451
Panel C. Comparisons				
Short-run impact (*I&P*)	0.14	−0.11	—	—
p-value, *I&P* (2005) versus *P* (2007–2009)	0.18	0.30	—	—
p-value, *I&P* (2005–2009) versus *P* (2007–2009)	0.90	0.14	0.99	0.59

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors clustered by catchment area in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. All specifications exclude observations with z-scores $> |4.5|$ as implausible and observations with recorded weight (columns 1 and 2) or height (columns 3 and 4) above the ninetieth percentile in the growth chart reported in Cortinovis et al. (1997). Panel A: *Information & participation* intervention, measured at the end of the trial (2009). Panel B: *Participation* intervention, measured at the end of the trial (2009). Specifications: (1) WAZ for children 0–12 months; (2) Weight-for-age z-scores (WAZ) for children 13–59 months with recorded weight above the ninetieth percentile in the growth chart reported in Cortinovis et al. (1997); (3) Height-for-age z-scores (HAZ) for children 0–12 months; (4) HAZ for children 13–59 months. Panel C: Short-run impact (*I&P*) reports treatment effect for the *information & participation* intervention after one year (2005). p-values, *I&P* (2005, 2005–2009) versus *P* (2007–2009) are p-values on the test of the null hypothesis of equal treatment effects, based on equation (3), between the *information & participation* intervention (*I&P*) (measured at the beginning of 2006 for the short-run evaluation and 2009 for the long-run evaluation) and the *participation* intervention (*P*) (measured in 2009). Data on height of children was not collected in 2006.

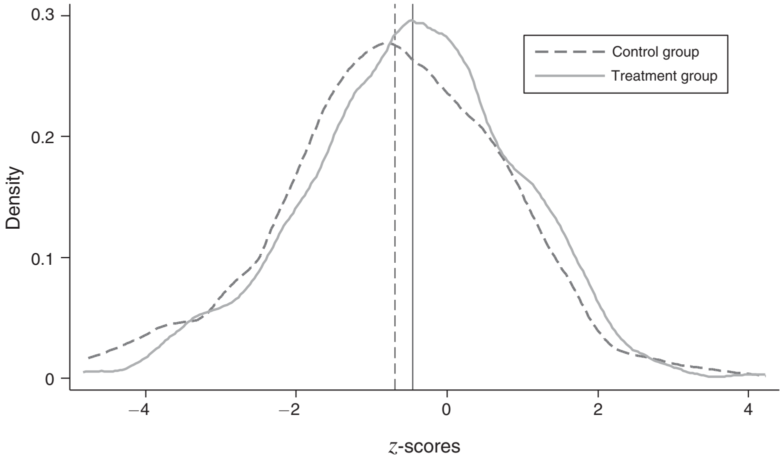


Figure 4. Distribution of Weight-for-Age z-Scores for the Treatment and Control Groups: Information & Participation Model (2006–2008)

Notes: Weight-for-age z-scores for children 0–12 months excluding observations with recorded weight above the ninetieth percentile in the growth chart reposted in Cortinovis et al. (1997). Sample size is 669 children. Solid line depicts the distribution for the treatment group and dashed line the distribution for the control group. Vertical solid line denotes mean in treatment group; vertical dashed line denotes mean in control group.

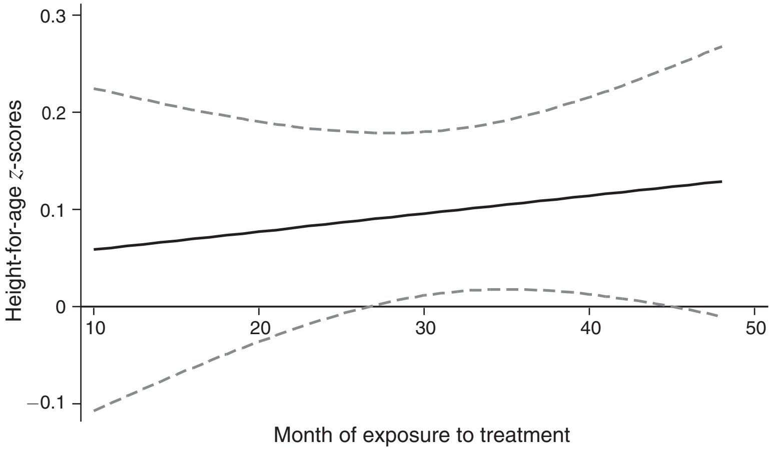


Figure 5. Treatment Effect on Height-for-Age z-Scores Conditional on Time of Exposure: Information & Participation Model

Notes: Figure plots the treatment effect conditional on months of exposure (solid line) and the 95 confidence intervals (dashed lines); i.e. the derivative $\partial z / \partial T$ of equation (1) augmented with *time* (in month) of exposure to treatment (0–48 months) and *time* × T, where z is the outcome variable (height-for-age z-scores for children) and T is the treatment indicator.

on equation (1) are reported in Table 3 and estimates from a value-added specification are reported in Appendix Table A2.[7]

Four years after the intervention was initiated, utilization—derived from daily patient registers during the final full year of the experiments (2008)—across all categories (out-patients, delivery, antenatal care, and family planning) is higher in the treatment relative the control group in the *information & participation* experiment. Although the point estimates are imprecisely estimated, the average standardized effect, reported in specification (5), is significantly different from zero. The precision improves and the point estimates are larger in the value-added specification (see panel A in Appendix Table A3). The impact on utilization is substantial. For outpatient services the increase is 16 percent (27 percent in the value-added specification). For deliveries, the point estimate implies a 46 percent increase in number of deliveries at the health facility. For antenatal care and family planning, we observe an increase of approximately 20 percent in the treatment as compared to the control group. The longer run estimates are similar to the effects observed after one year (panel C, Table 6).

---

[7] We have pre-intervention data for outpatients and deliveries but not antenatal care and family planning. We therefore use the average utilization for outpatient services and deliveries pre-intervention as a lagged variable in the value-added model and estimate

$$y_{jst} = \alpha_{VA} + \beta_{VA} T_j + \lambda \bar{y}_{jt-1} + \varepsilon_j,$$

where y_{sjt} is utilization for service s in facility j at time t, and $\bar{y}_{jt-1}$ is the average utilization for outpatient services and deliveries in the pre-intervention period $t - 1$.

Table 6—Program Impact on Utilization

Dependent variable Specification	Out-patients (1)	Delivery (2)	Antenatal (3)	Family planning (4)	Average std. effect (5)
Panel A. Information & participation					
Program impact	97.5	6.0	18.0	4.9	0.37
	(63.8)	(3.7)	(14.5)	(5.5)	(0.21)
Mean control	598.6	13.0	84.8	26.4	
Observations	50	50	50	50	50
Panel B. Participation					
Program impact	−100.5	−12.3	5.4	3.0	−0.14
	(56.5)	(5.2)	(15.1)	(3.8)	(0.19)
Mean control	649.6	23.9	106.9	18.6	
Observations	25	25	25	25	25
Panel C. Comparisons					
Short-run impact (*I&P*)	102.1	6.0	16.4	5.1	0.47
p-value, *I&P* (2005) versus *P* (2007–2009)	0.00	0.00	0.46	0.78	0.02
p-value, *I&P* (2005–2009) versus *P* (2007–2009)	0.00	0.00	0.97	0.91	0.09

Notes: Estimates from equation (1) with district fixed effects and robust standard errors are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Utilization data is the monthly average for the last full year (2008). Panel A: *Information & participation* intervention. Panel B: *Participation* intervention. Specifications: (1) average number of patients visiting the facility per month for outpatient care; (2) average number of deliveries at the facility per month; (3) average number of antenatal visits at the facility per month; (4) average number of family planning visits at the facility per month; (5) the average standardized effect of the estimates in specifications (1)–(4). Panel C: Short-run impact (*I&P*) reports treatment effect for the *information & participation* intervention after one year (2005). *p*-values, *I&P* (2005, 2005–2009) versus *P* (2007–2009) are *p*-values on the test of the null hypothesis of equal treatment effects, based on equation (3), between the *information & participation* intervention (*I&P*) (2005 data for the short-run evaluation and 2008 data for the long-run evaluation) and the *participation* intervention (*P*) (2008 data).

In the *participation* experiment, on the contrary, we find no evidence of impact. For two of the four utilization measures, the point estimates are in fact even negative. The average standardized treatment effect is small and not significantly different from zero.

The test of the null hypothesis of equal treatment effects, reported in panel C, can be rejected for the joint measure of utilization (column 5), as well as for outpatient services (column 1) and deliveries (column 2).

In Appendix Table A3, we report changes in utilization patterns based on household data. We collected data on where each household member sought care in case of illness that required treatment. Consistent with the findings reported above, households in the treatment communities in the *information & participation* trial switched from traditional healers and self-treatment (column 2), to the project facility (column 1), in response to the intervention.

For the *participation* trial (panel B, Table A3), treatment households are less likely to use traditional healers and self-treatment, compared to households in the control group, but not more likely to use the project health facility.

C. *Processes and Health Treatment Practices*

The *information & participation* intervention, which resulted in large improvements in health outcomes and utilization in the short run (Björkman and Svensson 2009), also resulted in longer run improvements in child health and increased use of the services provided by the primary health facility. Below we try to identify possible channels through which these changes came about. The intervention was intended to improve health outcomes by strengthening the community's ability to monitor and hold providers to account which, in turn, was intended to improve the management of the clinics and result in health workers exerting higher effort to provide more and better health care. We next turn to assessing the evidence for these two links.

Monitoring and Management.—We start by assessing whether the interventions resulted in improved monitoring capabilities of community members and better management of the health facility, using both facility and household level data collected at the end of the trial period.

Monitoring processes are measured by two sets of variables. We collected information on the extent to which monitoring tools, such as a visible staff duty roster, a suggestion box for complaints and recommendations, numbered waiting cards, posters informing about free health services, and posters on patients' rights and obligations, were available at the health facility, using visual checks by the enumerators. We also collected data at the household level on whether the household had received information about the Health Unit Management Committee's roles and responsibilities, whether the performance of the staff and the general management of the local health facility had been discussed in local council meetings during the past year, whether the household openly expressed their dis/satisfaction with the quality of service at the facility, and whether the household was aware of who monitored facility performance.

We assess how well the facility is managed using health facility level data, by measuring absence rate (using unannounced visits), the conditions of the floors, walls, and furniture (through visual inspection by the enumerators), and the share of months in 2009 with stock-outs of drugs (using health facility records).

For each group of process outcomes—accountability processes using facility level data, accountability processes using household survey data, and management of the clinic—we derive average standardized treatment effects. The estimates are reported in Table 7. The individual regressions underlying these average standardized treatment effects are reported in Appendix, Table A4a-c.

Panel A in Table 7, columns 1–3, shows that the *information & participation* intervention improved the community's monitoring capabilities. The average standardized effect on having a suggestion box, numbered waiting cards, a staff duty roster, and posters informing patients about their rights and that services are free is significantly positive (column 1). The result using household rather than facility data (column 2) confirms this result. We do not find a significant effect on the aggregate measure of the management of the clinic (column 3), although the intervention group in the *information & participation* intervention performs better (see

Table 7—Processes and Health Treatment Practices

Dependent variable Specification	Monitoring & information (facility) (1)	Monitoring & information (household) (2)	Management of clinic (3)	Health practices (general) (4)	Antenatal and postnatal care (5)	Health education (6)
Panel A. Information & participation						
Program impact	0.34	0.15	0.26	−0.01	0.20	0.07
	(0.16)	(0.03)	(0.17)	(0.04)	(0.05)	(0.02)
Observations	50	5,005	50	3,834	992	2,827
Panel B. Participation						
Program impact	−0.12	0.05	−0.04	0.02	−0.05	−0.04
	(0.17)	(0.03)	(0.14)	(0.05)	(0.07)	(0.03)
Observations	25	2,843	25	2,210	664	1,701
Panel C. Comparison						
Short-run impact (*I&P*)	1.07	—	0.39	0.05	—	—
p-value, *I&P* (2005) versus *P* (2007–2009)	0.00	—	0.07	0.78	—	—
p-value, *I&P* (2005–2009) versus *P* (2007–2009)	0.03	0.10	0.27	0.15	0.04	0.00

Notes: Each column of each panel represents the results of a separate system of linear regressions with district fixed effects. Average standardized treatment effect is reported, with robust standard errors (columns 1 and 3) clustered by catchment area (columns 2, 4–6) in parentheses. Panel A: *Information & participation* intervention. Panel B: *Participation* intervention. Column 1: summarizes whether there is a visible staff duty roster, a suggestion box for complaints/recommendations, numbered waiting cards, posters informing about free health services, and posters on patients' rights and obligations at the facility (see Table A4a for details). Column 2 summarizes whether households have received information about the Health Unit Management Committee's roles/responsibilities; whether the performance of the staff/management of the facility had been discussed in local council meetings during the past year; whether the household openly expressed dis/satisfaction with the quality of service at the facility; and whether the household was aware who monitored facility performance (see Table A4b for details). Column 3 summarizes absence rate (reversing the sign); conditions of the floors, walls, furniture, and share of months in 2008 with stock-outs of drugs (reversing sign) (see Table A4c for details). Column 4 summarizes average waiting time (reversing sign) and whether any equipment was used during examination (see Table A4c for details). Column 5 summarizes whether, during the last antenatal care visit in 2008–2009, the woman was examined by a midwife, informed about pregnancy complications, had her weight checked; blood sample taken, the fetus checked, and whether her newborn was checked at the health facility in the first two months after delivery (see Table A4d for details). Column 6 summarizes whether the respondent had heard about AIDS, did not agree with the statement that people with AIDS should be ashamed of themselves, knew that TB is spread through the air, and whether children sleep under an insecticide treated bed net (see Table A4e for details). Panel C: Short-run impact (*I&P*) reports treatment effect for the *information & participation* intervention after one year (2005). *p*-values, *I&P* (2005, 2005–2009) versus *P* (2007–2009) are *p*-values on the test of the null hypothesis of equal treatment effects, based on equation (3), between the *information & participation* intervention (*I&P*) (measured at the beginning of 2006 for the short-run evaluation and 2009 for the long-run evaluation) and the *participation intervention* (*P*) (measured in 2009). The variables reported in columns 2, 5, and 6 were not collected in 2006.

Table A4c) on two out of five underlying outcomes (conditions of the floors, walls, and furniture, and the share of months without stock-outs).

Panel B in Table 7, columns 1–3, shows the effects for the same monitoring and management outcomes for the *participation* intervention. While there is evidence that community members in the treatment arm of the *participation* experiment were more likely to have discussed the health facility's performance in local council meetings during the last year (column 2 and Table A4b), there is no evidence of increased monitoring activities at the health facility (column 1) nor any evidence that the management of the clinic improved.

In panel C, we report the *p*-values associated with the tests that the treatment effects of the two interventions are equal. We can reject the null hypothesis of equal treatment effects between the *information & participation* and *participation* interventions for two out of three family of outcomes (columns 1–3).

Health Treatment Practices.—We measure treatment practices using household survey data. We use four sets of metrics. First, to measure general health treatment practices we collected data on average waiting time and the extent to which health facility staff used any equipment during examination. Second, we collected detailed data on the extent to which the health facility followed the clinical guidelines during antenatal care visits (examined by a midwife, informed about pregnancy complications, had their weight checked, blood sample taken, and the fetus checked) and postnatal care visits (whether newborns were checked at the health facility in the first two months after delivery). Third, we collected data on the extent and quality of health education (including information on whether the household had heard about AIDS, whether they agreed with the statement that people with AIDS should be ashamed of themselves, whether they know that TB is spread through the air; and whether their children slept under an insecticide treated bed net). Finally, we collected detailed information on immunization of under five-year-old children.

We follow the same procedure as for the monitoring and management outcomes discussed above and derive average standardized treatment effects of each family of outcomes—general health treatment practices, antenatal and postnatal care, health education, and immunization of children—and report the estimates in Tables 7–8. The individual regressions underlying the average standardized treatment effects in Table 7 are reported in Appendix, Tables A4a-e.

Column 4, panel A of Table 7 shows that there is no significant difference in general health treatment practices between assignment arms in the *information & participation* intervention. However, for antenatal and postnatal care and health education (columns 5–6), there are large and precisely estimated differences between the intervention and control group. That is, women that have experienced a pregnancy over the trial period are significantly more likely to have received better quality antenatal and postnatal care. Four three of these outcomes (weight measurement, blood testing, and postnatal care), the effect sizes are over 20 percent (Table A4d). As shown in column 6, panel A, households in the intervention group also have better health knowledge than households in the control group.

In the *participation* intervention (panel B, columns 4–6 of Table 7), there is no systematic difference between the intervention and the control group for any of the aggregate measures of health treatment practices.

The p-values associated with the tests that the treatment effects of the two interventions are equal (Table 7, panel C) show that we can reject that the coefficients are equal for both antenatal and postnatal care and health education.

Table 8 reports the impact on immunization of under five-year-old children for the *information & participation* intervention (panel A) and for the *participation* intervention (panel B).[8] Two results stand out. First, while the point estimates are

---

[8] For each age group, we use information on how many times (doses) in total each child has received of polio, DPT, BCG, and measles vaccines and vitamin A supplements. On the basis of the recommended immunization plan, we create indicator variables taking the value of one if child i of cohort (age) j had received the required dose(s) of

Table 8—Program Impact on Immunizations

Group Specification	Newborn (1)	< 1 year old (2)	1 year old (3)	2 years old (4)	3 years old (5)	4 years old (6)
Panel A. Information & participation						
Average standardized effect	0.03	0.02	0.03	0.05	0.05	0.06
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)
Mean control group	0.78	0.83	0.89	0.91	0.92	0.94
Observations	2,535	2,481	1,894	1,325	826	372
Panel B. Participation						
Average standardized effect	−0.04	−0.02	−0.02	−0.05	0.02	0.11
	(0.08)	(0.06)	(0.06)	(0.07)	(0.08)	(0.09)
Mean control group	0.79	0.83	0.88	0.90	0.90	0.92
Observations	1,498	1,466	1,097	756	467	226
Panel C. Comparison						
Short-run impact (*I&P*)	0.32	0.09	0.10	0.01	0.16	0.07
p-value, *I&P* (2005) versus *P* (2007–2009)	0.00	0.03	0.19	0.70	0.15	0.84
p-value, *I&P* (2005–2009) versus *P* (2007–2009)	0.17	0.36	0.60	0.22	0.75	0.77

Notes: Average standardized effects with the dependent variables being indicator variables for whether the child has received at least one dose of measles, DPT, BCG, and polio, respectively, (see text for details), and with district fixed effects. Robust standard errors clustered by catchment areas are in parentheses. Groups: (1) Children under three months; (2) Children 0–12 months; (3) Children 13–24 months; (4) Children 25–36 months; (5) Children 37–48 months; (6) Children 49–60 months. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention, data for 2006–2008. Panel B: *Participation* intervention, data for 2007–2008. Panel C: Short-run impact (*I&P*) reports treatment effect for the *information & participation* intervention after one year (2005). *p*-values, *I&P* (2005, 2005–2009) versus *P* (2007–2009) are *p*-values on the test of the null hypothesis of equal treatment effects, based on equation (3), between the *information & participation* intervention (*I&P*) (2005 for the short-run evaluation and 2006–2009 for the long-run evaluation) and the *participation* intervention (*P*) (2007–2009).

positive for the *information & participation* intervention, the treatment effects are small (3–6 percent) and insignificantly different from zero. Likewise, there are no significant differences in immunization patterns across the intervention and control group in the *participation* intervention. Second, and unlike the situation in 2005, immunization coverage in the control groups is high (between 78–94 percent). In fact, as almost nine out of ten children received immunizations according to the recommended immunization plan, there is little room for improvement.[9]

IV. Mechanism and Discussion: The Role of Information

The results of the *participation* and the *information & participation* interventions differ markedly. Without information, the process of stimulating participation and engagement had little impact on health workers' performance or the quality

---

measles, DPT, BCG, and polio vaccines, respectively, and zero otherwise. We then estimate equation (3), for each age group, and calculate average standardized effects.

[9] The World Health Organization and other UN agencies have worked intensively with the Ministry of Health in Uganda on large-scale and country-wide immunization campaigns during the period of the study. This may explain the improvement in immunization coverage in both the intervention and the control groups.

of health care. In contrast, when community members were informed about staff performance, the same type of process intervention resulted in significant improvements in health care delivery and health outcomes in both the short and the longer run.

In this section we investigate why the provision of information appears to have played such an important role. We start by using process data from the two intervention groups to identify differences in decisions and actions taken in the two groups. We also discuss alternative mechanisms and review the literature on beneficiary control.

To identify possible mechanisms, it is illustrative to view the community contract (joint action plan) as the (contractual) solution to the community's (the principal's) problem of eliciting effort from the service provider (the agent). The service providers may be intrinsically motivated but hampered and demoralized by the same system of weak accountability relationship that frustrates users. In such a contract, the final outcome depends on the principal's ability to identity tasks that both maps into better health service and can be influenced by the health workers' actions. It also depends on the principal's ability to observe actions and the compensation scheme. When applying the framework to community monitoring one must also take into account the fact that there are multiple principals (the users) that may disagree about both objectives and the mapping between actions and outcomes.

In general, one can view the participatory components of a community monitoring project as dealing with the process of reaching an agreement. Providing information on performance can facilitate that process and also help the principal to identify, and possibly, over time, observe actionable tasks. The compensation scheme, i.e., the ability to reward and punish the agent, is typically not explicitly addressed in these type of program. Instead the process relies on informal forms of social recognition and/or social opprobrium. Within reasonable bounds, it is possible that non-financial rewards (social recognition) may be just as effective at eliciting effort as financial incentives.

A. *Information and Local Actions*

The provision of information on the health facility's performance was intended to assist the community (and health facility staff) to better distinguish between health workers' actions and factors outside the health workers' control, and thereby improve the community's ability to hold the provider to account.

Baseline survey data provide suggestive evidence that providers and users have different views and perceptions about the performance of the health clinic. For example, at baseline, the mean waiting time as reported by the in-charge was two minutes. The waiting time reported by users, on the other hand, was close to two hours on average. In the health facility survey the in-charge did typically not report problems with absenteeism and explained the low presence at the time of the survey as the result of a high, but unsubstantiated, level of staff training and outreach. Data from the unannounced staff surveys, however, showed an absence rate of around 50 percent. Finally, when the in-charge was asked to list key constraints facing the clinic, lack of funding, staff, material, and drugs were high on the agenda, while

TABLE 9—INFORMATION PROCESSES AND LOCAL ACTIONS

Dependent variable Specification	Participants in community meetings (1)	Number of actions in contract (2)	Share of upper level actions at baseline (3)	Share of local-level actions (2005/2007) (4)	Share of local-level actions (2007) (5)
Information & participation group	131.4	13.0	0.12	0.88	0.80
	(36.6)	(5.3)	(0.12)	(0.12)	(0.16)
Participation group	144.5	12.9	0.50	0.50	0.50
	(36.4)	(5.1)	(0.13)	(0.13)	(0.13)
Difference	−13.2	0.04	−0.39	0.39	0.30
	(12.4)	(1.8)	(0.04)	(0.04)	(0.05)
p-value	0.30	0.98	0.00	0.00	0.00
Observations	38	38	38	38	38

Notes: Mean outcomes in the two intervention groups, with standard deviations reported in parentheses. Difference is difference between the intervention groups with robust standard errors in parentheses. Specifications: (1) number of participants in the community meeting; (2) number of actions agreed upon by community members and health facility staff in the joint action plan; (3) share of upper level actions agreed upon in the baseline action plan (2005 for the *information & participation* experiment and 2007 for the *participation* experiment); (4) share of local-level actions out of the total number of actions agreed upon in the initial action plan (2005 for the *information & participation* experiment and 2007 for the *participation* experiment); (5) share of local-level actions in the 2007 action plans (follow-up action plan for the *information & participation* experiment and initial action plan for the *participation* experiment).

issues, such as weak adherence to clinic guidelines, shirking, or mistreatment of patients, were never listed.

Given such differences in perceived (or at least reported) performance, can the provision of information on the staff's behavior in various dimensions influence the community members' willingness to take actions? If so, will it drive the process toward actionable tasks, i.e., issues the health workers or the users could address themselves? We turn to these questions next.

A key component of both experiments was the agreement of an action plan. The action plan outlined the community's and the provider's joint agreement on what needs be done and by whom. As shown in Table 9, the process of reaching an agreement looks similar on some observable characteristics. About the same number of community members participated in the community meetings (column 1). The two treatment groups also identified roughly the same number of actions to be addressed (column 2).

A closer look at the type of actions agreed upon, however, reveals that while the *participation* group mostly identified issues that required third-party actions, e.g., more financial and in-kind support from upper-level authorities and NGOs, and timely deliveries of medicines from the center, the *information & participation* group almost exclusively identified (88 percent on average) local issues, which either the health workers or the users could address themselves, including absenteeism, opening hours, waiting time, and patient-clinician interactions (columns 3 and 4). Even two years into the *information & participation* trial, i.e., in 2007, when a shorter repeat engagement was implemented, more than four-fifths of the actions identified in the joint action plans dealt with local issues (column 5). Figure 6 illustrates the findings.

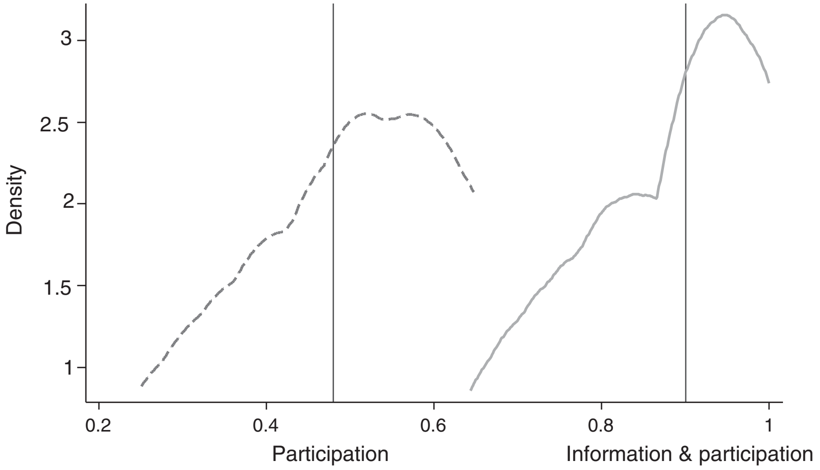


Figure 6. Local Actions for the Two Treatment Groups

Note: Figure plots the kernel density for the share of local actions for both treatment groups out of the total amount of actions listed in the joint action plan.

The difference in the content of the action plans across the two treatment arms and the difference in the impact (health outcomes and staff behavior) of the two interventions are consistent with the hypothesis that when the principal (i.e., community) is informed, efforts to stimulate local control result in an actionable plan that affects staff behavior, and thereby health utilization and health outcomes. That is, the report card component helped build a reform agenda on the true as opposed to the perceived status of service provision.

To further examine the plausibility of the information and local actions channel as a key mechanism for the health utilization and health outcomes treatment effects, we exploit variation within the two treatment arms. Specifically, we use the difference in the number of outpatients served at baseline and at the end of the trial in treatment as compared to the average control clinic as the dependent variable, i.e., $\overline{y}_j = \left(y_{tj}^T - y_{t-1j}^T\right) - \left(\overline{y}_t^C - \overline{y}_{t-1}^C\right)$, and regress $\overline{y}_j$ on the share of local issues raised in the action plan. The estimated relationship between $\overline{y}_j$ and the share of local actions agreed upon in the action plan is illustrated in Figure 7. There is a consistent pattern across clinics and treatment arms. The *information & participation* group identified a significantly larger share of local actions, and a higher share of local actions is associated with a larger, and significant, increase in the number of outpatients served over time and as compared to the control group.

B. *Alternative Mechanisms*

The findings of large treatment effects in the *information & participation* intervention are consistent with the community-based monitoring mechanism discussed above. But the findings do not rule out other explanations. In particular, other agents in the supply chain, e.g., district or subdistrict management, may have changed their

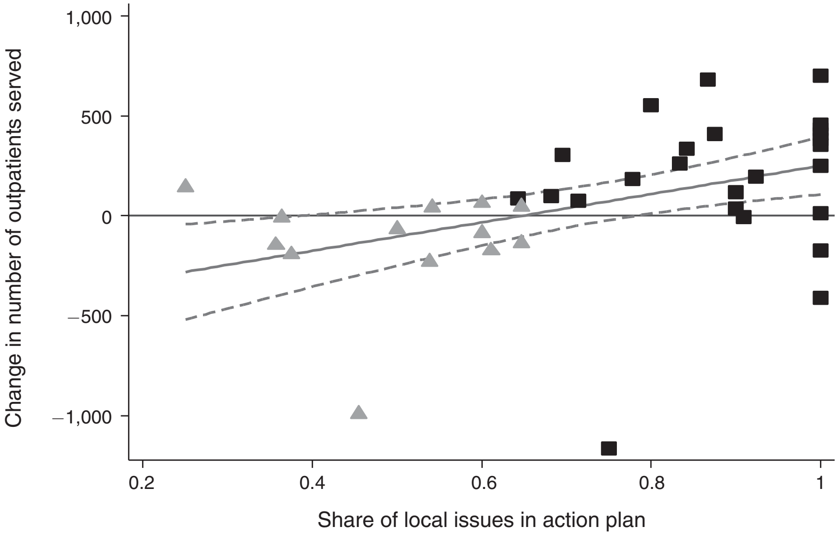


Figure 7. Outpatient Care Conditional on the Share of Local Action

Notes: This figure plots the difference in outpatients served before and after treatment in treatment clinics (*participation intervention* in triangles and *information & participation intervention* in squares) compared to control clinics, conditional on the share of local actions agreed upon in the action (solid line) and the 90 confidence intervals (dashed lines).

behavior or support in response to the intervention, for example, by providing additional funding or other support to the treatment facilities. We do not find evidence of this being the case (see the Appendix, Table A5). The treatment facilities did not receive more drugs or funding from the subdistrict or district as compared to the control facilities. Moreover, the level of supervision of providers by government elected authorities remained low in both the treatment and the control group. There is also no difference between treatment and control facilities in the number of staff that voluntarily left the facility.

C. *Discussion*

In this paper, we present the results of two field experiments. First, we assess the longer run impact of an intervention combining the standard package of facilitated meetings to enhance participation with the dissemination of report cards on the facility's performance. While programs to stimulate community involvement are nowadays common in most developing countries, there is little evidence of whether and when these types of bottom-up initiatives have the intended impact and no evidence based on rigorous evaluations of the longer run impact.

Second, we assess the short-run impact of a cheaper and technically less demanding intervention involving only the standard package of facilitated meetings to enhance participation.

We find that the process of stimulating participation, when the community does not have access to information on performance, resulted in a joint agreement focusing on issues requiring third-party actions. The intervention had little impact on health workers' behavior or the quality of health care. In contrast, with an informed community, the same type of process intervention resulted in a joint agreement that almost exclusively identified issues that either health workers or the users could address themselves, and large improvements in health care delivery and health outcomes in both the short-run and longer run.

The results of the two field experiments resonate relatively well with the mixed findings in the literature. Banerjee et al. (2010) study beneficiary control in Uttar Pradesh, India. In one of their treatment arms, trained facilitators held small-group discussions with parents and provided them with simple tools to enable them to generate their own information about their children's learning outcomes. This information was later shared in a village meeting where the school teachers were also asked to provide general information about the resources available at the school. While test scores, on average, are likely correlated with the teachers' effort, it is plausible that parents would (correctly) infer that the variation in test scores is to a large extent driven by child- and household-specific factors, rather than teacher effort, and they might, possibly incorrectly, infer that the lack of resources that has been highlighted in the meetings is also a quantitatively important factor. Thus, even if parents viewed these test scores as a credible and objective measure of what their children have learned, they might be skeptical about using these test scores as an indicator of the school's performance. As in the *participation* intervention evaluated above, the village meetings were dominated by discussions about third-party actions (e.g., the state government's scholarship program and a new school-meals program), not the teachers' performance or the learning environment. The intervention prompted no increased teacher effort and no improvement in educational outcomes.

Olken (2007) evaluates different ways of monitoring corruption in a road construction project in Indonesia. In one of the experiments, invitations were sent out to village-level meetings where project officials documented how they spent project funds for local road construction. This provided villagers with new information, but at best only indirect information about the key outcome variable—corruption—as project officials and/or elite community members may be able to hide it when reporting on how funds were used. Thus, it is unclear whether the nonelite community members were really more informed about corruption in the project and consistent with the above results, corruption problems were seldom discussed in the village meetings and the intervention had little impact.

Duflo, Dupas, and Kremer (2015) evaluate a contract teachers program in Kenya. One of their treatment arms included School-Based Management (SBM) training of the PTA committee. Committee members were then given specific training on how to monitor and assess teachers' effort and performance, and a set of parents were asked to perform attendance checks on the teachers on a regular basis. A formal subcommittee of parents was formed to evaluate the contract teacher and deliver a performance report at the end of the year. The PTA was also given influence over the decision whether to retain the contract teacher or not. While this intervention did not directly provide quantitative information on performance, it provided detailed training to the community on how to measure it throughout the year and how to combine the information in a performance report that could be shared with others. SBM training for PTA committees reduced teachers' absenteeism and increased student test scores.

Barr et al. (2012) study different ways to provide SMC members with tools to monitor teacher performance. In line with the findings in Bjorkman and Svensson (2009), they show that the provision of tools for assembling performance information, combined with an explicit participatory component, resulted in reduction in teacher absenteeism and increased student test scores.

Finally, Andrabi, Das, and Khwaja (2015) focus on providing information only. They show that providing households with report cards that displayed test results of their own children and test results from all schools, private and public, that catered to the students in the village, raised test scores both in private schools and in public schools with limited market or administrative disciplining mechanisms—a result consistent with the hypothesis that information about relative performance can strengthen parents' ability to hold public providers to account.

Our findings provide both encouraging and less encouraging news for those promoting greater beneficiary control. On the one hand, we show that a standard and CDD inspired beneficiary involvement intervention, where the core of the strategy is the process through which (local) decisions are made, had no impact. On the other hand, we show that the same process based intervention in a scenario where the community is informed about how the public health facility is performing resulted in large and long-run improvements in both health service provision and health outcomes. While the intervention involving the production and dissemination of report cards was significantly more expensive, the findings of large and sustained positive health effects suggest that the returns may still be high. In fact, a back-of-the-envelope calculation shows that the intervention, including the cost for collecting data for the report cards (the main cost item), but not the costs of collecting data for the evaluation or the researchers time, was about \$10,000 per facility/community over a four-year period. Only judged on the cost per death averted this still must be considered to be a fairly cost-effective intervention. The estimated cost of averting the death of a child under five is \$278, which should be compared to the estimate that the average cost per child life saved through the combined and integrated delivery of 23 interventions shown to reduce mortality from the major causes of death in children younger than five years is \$887 (Bryce at al. 2005).[10]

The *information & participation* intervention was implemented at a relatively large scale (in nine districts across all four regions of Uganda), but it was not designed to provide a blueprint for a scalable intervention. This in turn opens up important questions for future research. For example, is it possible to provide beneficiaries with tools to collect performance data and strategies to use them also in the health sector, for instance, along the line of Duflo, Dupas, and Kremer (2015), instead of directly providing them with report cards? Can recent advances in ICT be used to identify ways of assembling and disseminating information at a lower cost?

---

[10] The number of under five-year-old deaths per year and cluster dropped by 0.4 on average (Table 3, panel A, column 1) or by 1.8 over the full trial period (54 months) in the treatment group. We surveyed about 5 percent of the households in the catchment area on average, so we estimate that the intervention over the full trial period reduced the number of under five-year-old deaths by 36 (=1.8/0.05) in the catchment area. The estimated cost of averting the death of a child is therefore \$10,000/36 = \$278.

APPENDIX

TABLE A1—COMPARISON OF CONTROL GROUP CHARACTERISTICS IN 2006

Variables	Information & participation (1)	Participation (2)	Difference between control groups (3)
Outpatient care	660.8	790.1	−129.2
	(175)	(330.2)	(100.2)
Delivery	9.2	14.7	−5.5
	(8.1)	(11.1)	(3.5)
Number of households in catchment area	2,224	2,519	295
	(1,021)	(1,144)	(385)
Number of households per village	95.3	118.1	−22.7
	(31.6)	(41.7)	(13.4)
Access to safe water	0.96	1.0	−0.04
	(0.2)	(0.0)	(0.04)
Access to electricity	0.44	0.42	0.02
	(0.51)	(0.52)	(0.18)
Average standardized effects			
Utilization from health facility records			−0.71
			(0.41)
Utilization pattern of the users			−0.02
			(0.03)
Quality of services according to users			0.05
			(0.14)
Catchment area statistics			−0.22
			(0.25)
Health facility characteristics			−0.23
			(0.19)
Citizen perceptions of treatment			0.00
			(0.07)
Supply of drugs			−0.67
			(0.49)
User charges			−0.47
			(0.29)

Notes: Each row reports catchment area/health facility averages for the two control groups (column 1 and 2) and difference in averages (column 3). Standard deviations are reported in parentheses below the averages and robust standard errors are reported in parentheses under the differences. Column 1 reports the control group characteristics for the *information & participation* experiment in 2006. Column 2 reports the control group characteristics for the *participation* experiment in 2006. Description of variables: see Table 1 and Table 2 for details.

Table A2—Program Impact on Utilization/Coverage Using Facility Data (*value added regression*)

Dependent variable Specification	Out-patients (1)	Delivery (2)	Antenatal (3)	Family planning (4)	Average std. effect (5)
Panel A. Information & participation					
Program impact	110.1	7.3	21.4	4.9	0.43
	(62.4)	(3.7)	(15.3)	(5.3)	(0.2)
Lagged variable	0.2	0.0	0.0	−0.0	
	(0.1)	(0.0)	(0.0)	(0.0)	
Mean control	598.6	13.0	84.8	26.4	
Observations	50	50	50	50	50
Panel B. Participation					
Program impact	−96.9	−11.8	6.1	3.5	−0.12
	(54.2)	(4.7)	(14.7)	(3.4)	(0.2)
Lagged variable	0.2	−0.0	0.0	0.0	
	(0.2)	(0.0)	(0.1)	(0.0)	
Mean control	649.6	23.9	106.9	18.6	
Observations	25	25	25	25	25
Panel C. Comparison					
Short-run impact (*I&P*)	136.4	6.6	19.3	5.4	0.56
p-value, *I&P* (2005) versus *P* (2007–2009)	0.02	0.00	0.46	0.78	0.02
p-value, *I&P* (2005–2009) versus *P* (2007–2009)	0.00	0.00	0.97	0.91	0.09

Notes: Estimates from equation (1) with district fixed effects and robust standard errors are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Lagged variable is the average utilization for outpatient services and deliveries pre-intervention (2004 for *information & participation* intervention and 2006 for *participation intervention*). Utilization data is the monthly average for the last full year (2008). Panel A: *Information & participation* intervention. Panel B: *participation* intervention. Specifications: (1) average number of patients visiting the facility per month for outpatient care; (2) average number of deliveries at the facility per month; (3) average number of antenatal visits at the facility per month; (4) average number of family planning visits at the facility per month; (5) the average standardized effect of the estimates in specifications (1)–(4). Panel C: short-run impact (*I&P*) reports treatment effect for the *information & participation* intervention after one year (2005). *p*-values, *I&P* (2005, 2005–2009) versus *P* (2007–2009) are *p*-values on the test of the null hypothesis of equal treatment effects, based on equation (3), between the *information & participation* intervention (*I&P*) (2005 data for the short-run evaluation and 2008 data for the long-run evaluation) and the *participation* intervention (*P*) (2008 data).

Table A3—Program Impact on Utilization/Coverage Using Household data

Dependent variable Specification	Use of project facility (1)	Self-treatment/ traditional healer (2)	Average std. effect (3)
Panel A. Information & participation			
Program impact	0.02	−0.00	0.04
	(0.01)	(0.00)	(0.02)
Mean control	0.21	0.01	—
Observations	5,000	5,000	5,000
Panel B. Participation			
Program impact	−0.01	−0.00	−0.02
	(0.01)	(0.00)	(0.03)
Mean control	0.23	0.01	—
Observations	2,843	2,843	2,843
Panel C. Comparison			
Short-run impact (*I&P*)	0.01	-0.01	0.04
p-value, *I&P* (2005) versus *P* (2007–2009)	0.03	0.25	0.37
p-value, *I&P* (2005–2009) versus *P* (2007–2009)	0.00	0.98	0.00

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors clustered by catchment areas are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention, data from 2009. Panel B: *Participation* intervention, data from 2009. Specifications: (1) the adult households' share of visits to the project facility out of all health visits; (2) the adult households' share of visits to traditional healers and self-treatment out of all health visits; (3) the average standardized effect of the estimates in specifications (1)–(2), reversing the sign of use of self-treatment or traditional healers. Panel C: short-run impact (*I&P*) reports treatment effect for the *information & participation* intervention after one year (2005). *p*-values, *I&P* (2005, 2005–2009) versus *P* (2007–2009) are *p*-values on the test of the null hypothesis of equal treatment effects, based on equation (3), between the *information & participation* intervention (*I&P*) (2005 data for the short-run evaluation and 2008 data for the long-run evaluation) and the *participation* intervention (*P*) (2008 data).

Table A4a—Intermediate Outcomes: Monitoring and Information at the Health Facility Level

Dependent variable Specification	Suggestion box (1)	Numbered waiting cards (2)	Staff duty roster (3)	Poster informing free services (4)	Poster on patients (5)
Panel A. Information & participation					
Program impact	0.21	0.00	0.17	−0.02	0.08
	(0.11)	(0.08)	(0.10)	(0.08)	(0.08)
Mean control group	0.08	0.08	0.08	0.12	0.04
Observations	50	50	50	50	50
Panel B. Participation					
Program impact	−0.03	−0.08	0.08	−0.27	−0.00
	(0.14)	(0.13)	(0.14)	(0.13)	(0.16)
Mean control group	0.17	0.17	0.08	0.25	0.17
Observations	25	25	25	25	25

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention. Panel B: *Participation* intervention. Outcome measures in columns 1–5 are based on data collected through visual checks by the enumerators during the post-intervention health facility survey. Column 1: indicator variable for whether the health facility has a suggestion box for complaints/recommendations; column 2: indicator variable for whether the facility has numbered waiting cards for its patients; column 3: indicator variable for whether the health facility has a visible staff duty roster; column 4: indicator variable for whether the facility has posters informing about free health services; column 5: indicator variable for whether the facility has posters on patients' rights and obligations.

Table A4b—Intermediate Outcomes: Monitoring and Information at the Household level

Dependent variable	Knowledge about HUMC	Discuss health facility performance and staff in LC meetings	Telling staff if dis/satisfied with quality of service	Health facility staff work closely with community	Knowing someone who monitored facility performance
Specification	(1)	(2)	(3)	(4)	(5)
Panel A. Information & participation					
Program impact	0.07	0.24	0.04	0.03	0.11
	(0.02)	(0.05)	(0.02)	(0.015)	(0.03)
Mean control group	0.14	0.72	0.25	0.15	0.77
Observations	5,002	2,522	3,823	4,294	4,991
Panel B. Participation					
Program impact	0.02	0.19	−0.03	0.01	0.04
	(0.02)	(0.07)	(0.02)	(0.02)	(0.03)
Mean control group	0.12	0.77	0.28	0.15	0.79
Observations	2,840	1,505	2,202	2,505	2,838

Notes: Estimates from equation (1) with district fixed effects and robust standard errors clustered by catchment areas are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention. Panel B: *Participation* intervention. Outcome measures are from the post-intervention household survey. Column 1: indicator variable for whether the household has received information about the Health Unit Management Committee's (HUMC's) roles and responsibilities; column 2: the sum of indicator variables whether the performance of staff/management and the functioning of the local health facility had been discussed at the local council meetings during the past year; column 3: indicator variable for whether the household openly expressed dis/satisfaction with the quality of service at the health facility and performance of the health facility staff; column 4: indicator variable for whether the health facility staff works closely with the community; column 5: the sum of indicator variables for whether the household was aware of community members who monitored facility performance (staff attendance, quality of services, cleaning of the health facility, constructions at the health facility).

Table A4c: Intermediate Outcomes: Treatment Practices and Management

Variable	Equipment used	Waiting time	Absence rate	Condition of clinic	Drug stock-outs
Specification	(1)	(2)	(3)	(4)	(5)
Panel A. Information & participation					
Program impact	0.02	7.06	0.03	0.80	−0.06
	(0.03)	(6.74)	(0.04)	(0.45)	(0.04)
Mean control group	0.33	113.08	0.45	−0.36	0.57
Observations	3,610	3,829	50	50	41
Panel B. Participation					
Program impact	0.01	−2.83	−0.06	−0.30	0.04
	(0.04)	(11.7)	(0.04)	(0.40)	(0.07)
Mean control group	0.38	129.03	0.54	0.44	0.51
Observations	2,109	2,206	25	25	19

Notes: Estimates from equation (1) with district fixed effects and robust standard errors clustered by catchment areas are in parentheses in columns 1 and 2. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention. Panel B: *Participation* intervention. Column 1: indicator variable for whether any equipment was used during examination when the patient visited the health facility; column 2: waiting time for the patient calculated as the difference between the time the patient left the facility and the time the patient arrived at the facility, minus the examination time; column 3: absence rate of the health facility staff during an unannounced visit excluding staff on-leave and reported on outreach visits; column 4: first component from a principal components analysis of condition of the floors, walls, furniture, and the smell at the health facility, where each condition is ranked from 1 (dirty) to 3 (clean) by the enumerators; column 5: share of months with stock-outs of drugs over the period 2006–2008 (see text for details).

Table A4d—Intermediate Outcomes: Antenatal and Postnatal Care

Dependent variable Specification	Examined by midwife during antenatal visit at the health facility (1)	Weight taken during antenatal visit (2)	Blood sample during antenatal visit (3)	Fetus checked during antenatal visit (4)	Told about pregnancy complications during antenatal visit (5)	Newborn checked at the health facility in the first two months after delivery (6)
Panel A. Information & participation						
Program impact	0.09	0.13	0.13	0.03	0.06	0.09
	(0.03)	(0.04)	(0.05)	(0.01)	(0.03)	(0.03)
Mean control	0.87	0.59	0.52	0.95	0.52	0.37
Observations	1,152	1,151	1,146	1,148	1,148	1,711
Panel B. Participation						
Program impact	0.01	−0.05	0.10	0.00	−0.04	−0.08
	(0.02)	(0.06)	(0.07)	(0.01)	(0.03)	(0.04)
Mean control	0.97	0.74	0.44	0.96	0.65	0.46
Observations	787	787	788	788	788	1,033

Notes: Estimates from equation (1) with district fixed effects and robust standard errors clustered by catchment areas are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention. Panel B: *Participation* intervention. The sample is women who had experienced a pregnancy in the last two years. Column 1: indicator variable for whether the woman was examined by midwife during the antenatal visit; column 2: indicator variable for whether the woman's weight was taken during the antenatal visit at the health facility; column 3: indicator variable for whether a blood sample was taken during the antenatal visit; column 4: indicator variable for whether the fetus was checked during the antenatal visit; column 5: indicator variable for whether the woman was told about pregnancy complications; column 6: indicator variable for whether the newborn was checked at the health facility in the first two months after delivery.

Table A4e—Intermediate Outcomes: Health Education

Dependent variable Specification	Have heard about AIDS (1)	AIDS stigma (2)	Knowledge that TB is spread through the air (3)	Children sleeping under a treated mosquito net (4)
Panel A. Information & participation				
Program impact	0.00	−0.02	0.04	0.06
	(0.00)	(0.01)	(0.02)	(0.02)
Mean control group	1.00	0.09	0.53	0.38
Observations	5,002	4,907	4,880	2,925
Panel B. Participation				
Program impact	−0.002	0.01	−0.01	−0.02
	(0.001)	(0.00)	(0.03)	(0.03)
Mean control group	1.00	0.06	0.54	0.46
Observations	2,843	2,796	2,762	1,766

Notes: Estimates from equation (1) with district fixed effects and robust standard errors clustered by catchment areas are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention. Panel B: *Participation* intervention. Column 1: indicator variable for whether the respondent has heard about AIDS; column 2: indicator variable for whether the respondent agrees with the statement that people with AIDS should be ashamed of themselves; column 3: indicator variable for whether the respondent knows that TB is spread through the air; column 4: indicator variable for whether the household's children are sleeping under an insecticide treated net.

Table A5—Alternative Mechanisms: Information & Participation Model (2006–2008)

Dependent variable	Funds received in USD	Number of staff left health facility	Supervision visits	Drugs received per patient
Specification	(1)	(2)	(3)	(4)
Panel A. Information & participation				
Program impact	12.7	–0.00	0.07	0.09
	(11.7)	(0.13)	(0.09)	(0.18)
Mean control group	79.4	0.20	0.32	—
Observations	50	50	50	48
Panel B. Participation				
Program impact	27.3	0.08	0.05	–0.16
	(31.2)	(0.32)	(0.13)	(0.24)
Mean control group	83.7	0.33	0.17	—
Observations	25	25	25	23

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors are in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Panel A: *Information & participation* intervention. Panel B: *Participation* intervention. Column 1: average yearly amount of Primacy Health Care (PHC) funds received at the health facility over the years 2006–2008 in USD; column 2: number of staff that has left the health facility during 2006–2008; column 3: if the health facility receives regular (once a month) monitoring visits from village health teams; column 4: average standardized treatment effect of the number of tablets received per patient and health facility of the drugs erythromycin, coartem/artemether, lumefantrine, cotrimoxazole/septrin, and quinine.

REFERENCES

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja. 2015. “Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets.” World Bank Policy Research Working Paper 7226.

Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. 2010. “Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India.” *American Economic Journal: Economic Policy* 2 (1): 1–30.

Barr, Abigail, Frederick Mugisha, Pieter Serneels, and Andrew Zeitlin. 2012. “Information and collective action in community-based monitoring of schools: Field and lab experimental evidence from Uganda.” Unpublished.

Besley, Tim, and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton: Princeton University Press.

Björkman, Martina, and Jakob Svensson. 2009. “Power to the People: Evidence from a Randomized Experiment on Community-Based Monitoring in Uganda.” *Quarterly Journal of Economics* 124 (2): 735–69.

Björkman Nyqvist, Martina, Damien de Walque, and Jakob Svensson. 2017. “Experimental Evidence on the Long-Run Impact of Community-Based Monitoring: Dataset.” *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20150027.

Bold, Tessa, Jakob Svensson, Bernard Gauthier, and Waly Wane. 2011. *Service Delivery Indicators: Pilot in Education and Health Care in Africa*. Chr. Michelsen Institute (CMI). Bergen, Norway, August.

Bryce Jennifer, Cynthia Boschi-Pinto, Kenji Shibuya, and Robert E. Black. 2005. “WHO estimates of the causes of death in children.” *Lancet* 365 (9465): 1147–52.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2006. “Missing in Action: Teacher and Health Worker Absence in Developing Countries.” *Journal of Economic Perspectives* 20 (1): 91–116.

Cortinovis, I., V. Vella, N. Ndiku, and S. Milani. 1997. “Weight, height and arm circumference of children under 5 in the district of Mbarara, south-west Uganda.” *Annals of Human Biology* 24 (6): 557–68.

▶**Das, Jishnu, Jeffrey Hammer, and Kenneth Leonard.** 2008. "The Quality of Medical Advice in Low-Income Countries." *Journal of Economic Perspectives* 22 (2): 93–114.

▶**Duflo, Esther, Pascaline Dupas, and Michael Kremer.** 2015. "School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools." *Journal of Public Economics* 123: 92–110.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using Randomization in Development Economics Research: A Tool Kit." Center for Economic Policy Research (CEPR) Working Paper 6059.

Kling, Jeffrey R., Jeffrey B. Liebman, Lawrence F. Katz, and Lisa Sanbonmatsu. 2004. "Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment." https://www.innovations.harvard.edu/sites/default/files/2609662.pdf.

Mansuri, Ghazala, and Vijayendra Rao. 2013. *Localizing Development: Does Participation Work?* Washington, DC: World Bank.

▶**Olken, Benjamin A.** 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–249.

Olken, Benjamin A., and Rohini Pande. 2012. "Corruption in developing countries." *Annual Reviews* 4 (1): 479–509.

Preston, Dale L. 2005. "Poisson Regression in Epidemiology." In *Encyclopedia of Biostatistics*, edited by Peter Armitage and Theodore Colton. Chichester, UK: John Wiley and Sons.

Republic of Uganda. 2000. *Health Sector Strategic Plan 2000/01-2004/05*. Kampala, Uganda: Ministry of Health.

Strömberg, David. 2015. "Media Coverage and Political Accountability: Theory and Evidence." In *Handbook of Media Economics*, edited by Simon P. Anderson, David Strömberg, and Joel Waldfogel, 595–622. Amsterdam: North-Holland.

World Bank. 2015. "Mortality rate, under-5, overview per country." *World Development Indicators Database*. Washington, DC: World Bank.