

Information is Power

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

Martina Björkman Nyqvist

Damien de Walque

Jakob Svensson



WORLD BANK GROUP

Development Research Group

Human Development and Public Services Team

August 2014

Abstract

This paper presents the results of two field experiments on local accountability in primary health care in Uganda. Efforts to stimulate beneficiary control, coupled with the provision of report cards on staff performance, resulted in significant improvements in health care delivery and health outcomes in both the short and the longer run. Efforts to stimulate

beneficiary control without providing information on performance had no impact on quality of care or health outcomes. The paper shows that informed users 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.

This paper is a product of the Human Development and Public Services Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at martina.bjorkman.nyqvist@hhs.se; ddewalque@worldbank.org and jakob.svensson@iies.su.se.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Information is Power: Experimental Evidence on the Long-Run Impact of Community Based Monitoring^{*,*}

Martina Björkman Nyqvist[†] Damien de Walque[‡] Jakob Svensson[§]

Keywords: Community Driven Development, Health, Information, Participation

JEL classification: I11, I15, O15

* 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. Bitekerezoo, 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 Adam Wagstaff and seminar and conference participants at Dartmouth, University of Michigan, Monash University, the World Bank, DIAL conference in Paris, and the NEUDC conference at Harvard. Financial support from the BNPP, the World Bank Research Committee, the Gender Action Plan (GAP) and the Strategic Impact Evaluation Fund (SIEF) at the World Bank, the Swedish Research Council (421-2009-2209) and the Program for Development Research, SIDA is gratefully acknowledged.

[†] Stockholm School of Economics and CEPR. Email: martina.bjorkman.nyqvist@hhs.se

[‡] The World Bank, Development Research Group. Email: ddewalque@worldbank.org

[§] IIES, Stockholm University and CEPR. Email: jakob.svensson@iies.su.se

1 Introduction

Poor quality plagues public service provision in many developing countries.¹ In response, policies to enhance beneficiary 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 mixed results about its short-run effectiveness.² Whether it can lead to sustained improvements in service provision is largely unknown.³

First, we provide evidence of the longer run impact of a local accountability intervention in primary health care provision in Uganda. The *participation & information intervention* combined the dissemination of report card information on staff performance with efforts to enhance participation. Four years after the initial intervention, we document significant improvements in health care delivery (increased utilization and improved adherence to clinical guidelines) and health outcomes (reduced child mortality and increased weight-for-age and height-for-age for children), in the treatment as compared to the control group. Thus, properly designed, enhanced beneficiary involvement can result in large and sustained improvements in both health service provision and health outcomes.

Second, we shed light on why this particular community-based intervention resulted in such a large and sustained change in service provision while several other seemingly similar interven-

¹Das et al. (2008) in their overview found quality of care to be very low in many developing countries. For example, doctors in Tanzania completed less than 25% 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) found an average absence rate of 27% among primary school teachers and 37% of primary health center staff in Uganda. Primary students in urban schools in Tanzania spend about one quarter of the daily schedule in a classroom where the teacher is actually present. Roughly half of the primary school teachers in Senegal fail to demonstrate mastery of the curriculum their students are supposed to master. 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 (Bold et al, 2011).

²Olken (2007) finds that grassroot participation in the monitoring of a village road construction program in Indonesia had little average impact. Banerjee et al. (2010) show that a project giving local committees in India the tools to evaluate student performance had no effect. Casey et al. (2012) evaluate an infrastructure project in Sierra Leone involving both relatively large grants and the application of processes to enhance local empowerment and participatory governance in the planning and implementation phases. They do not find any evidence that the intervention led to fundamental changes in collective action at the village level. Duflo et al. (2012), on the other hand, find that a governance program that gave parents specific training on how to monitor and assess teachers’ effort and performance in a contract-teachers program in Kenya resulted in significant improvements in learning. Björkman and Svensson (2009) find that providing users with information on health workers’ performance and facilitating the development of an action plan resulted in significant improvements in both health workers’ performance and health outcomes in Uganda. For a review of the World Bank’s experience with local participatory development programs, see Mansuri and Rao (2013).

³While lack of evidence of the long-run impact is not specific to studies on beneficiary control, it is of particular concern here since public service providers may easily adjust their behavior for a short period of time given the large pre-existing X-inefficiencies in public service provision, but permanent changes in behavior may be harder to institute. Moreover, interventions aimed at increasing community participation are primarily meant to influence norms and collective actions and thus influence local decision making both in the short and the longer run.

tions have not had much of an impact. Our focus here is on the role and impact of information. Specifically, we designed an intervention (the *participation intervention*) replicating the participation but not the information component of the *participation & information intervention*. Most community empowerment projects focus on participation and the *participation intervention* can be viewed, from a process point of view, as being fairly representative of the standard Community Driven Development (CDD) approach promoted by the World Bank and other donors. Holding context constant, we evaluate whether differences in intervention design – the provision (or not) of information on the health clinic’s performance – can help explain the mixed findings in the literature.⁴

The impacts of the interventions with and without information differ markedly. Without information, the process of stimulating participation and engagement (the *participation intervention*) had little impact on the health workers’ behavior, health outcomes or the quality of health care. In contrast, when community members are informed; i.e., have objective and quantitative information about staff behavior, the same type of processes resulted in significant improvements in health care delivery and health outcomes in both the short and the longer run.

Using data from the implementation phases of the two interventions, we investigate why the provision of information appears to have played such a key role. 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 type of issues to be addressed differed significantly. 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 *participation & information 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.

Taken together, our results provide both encouraging, and less encouraging, news for those pro-

⁴As pointed out in Banerjee et al. (2010), from the available evidence, it is difficult to disentangle whether the mixed findings are driven by differences in the details of the intervention or context.

moting greater beneficiary control. We find that an intervention that largely mimics the Community Driven Development (CDD) approach; i.e. the *participation intervention*, had little impact.⁵ The technically more demanding (and more expensive) *participation & information intervention*, however, resulted in a more engaged community and large and long-run improvements in both health service provision and health outcomes.

The next section describes the institutional setting for our study. Section 3 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 *participation & information intervention* are presented in section 4, while the evaluation results of the *participation intervention* are discussed in section 5.⁶ Section 6 presents some suggestive findings in support of the key mechanism, relates our findings to the mixed results in the literatures and discusses the policy recommendations of our findings.

2 Institutional setting

The experiments were implemented in 75 rural communities served by a public primary health facility (or dispensary) in nine districts covering all four 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 6-10 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. However, the baseline data reveal that these institutions, including the HUMC, are not actively involved in the supervision or support of primary health care providers.

The setting for our experimental study – rural Uganda – is characterized by poor public health service provision (large X-inefficiencies). For example, roughly 50% 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

⁵The CDD approach has become one of the most common approaches to poverty reduction (broadly defined) in developing countries. Over the past decade, the World Bank alone has allocated close to \$85 billion to local participatory development programs (Mansuri and Rao, 2013).

⁶The short-run evaluation results of the *participation & information intervention* are presented in Björkman and Svensson (2009).

was used the last time the respondent (or the respondent's child) visited the clinic.⁷

3 Experimental design and data

3.1 Overview

The research project was initiated in 2004 and extended in 2007. Of the 75 rural communities and health facilities, 50 facilities/communities were included in the first phase of the project in 2005 (the *participation & information intervention*) and 25 facilities/communities were added in 2007 (the *participation intervention*). The catchment area or community for each dispensary was defined as the households residing in the 5-km radius around the facility. A community in our sample does, on average, have 2,500 households residing within the five-kilometer radius of the clinic, of which 350 live within a one-kilometer radius.

For the *participation & information* 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).

Trial sizes were set to detect effects on utilization and under-five child 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 households (and collected birth and death statistics from approximately 100 under-five children). Thus, the sample for the *participation & information* 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 are available for 2005 and these data were used to estimate study power (Björkman and Svensson, 2009). The estimated overall under-five mortality rate (the number of child deaths per 1000 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 *participation & information intervention* has 80%

⁷While we have no data on the clinicians' workload, other estimates from similar settings suggest a low workload. For example, combining observational data on time spent per patient with facility records on the number of patients treated per day, Bold et al. (2011) estimate that the total time spent counseling patients per day per clinician in rural primary health clinics in Tanzania and Senegal is only 26 minutes.

[60%] power of detecting significant differences at the 0.05 significance level if the intervention reduced mortality by 25% [20%]. 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% [60%] power of detecting a significant difference at the 0.05 significance level if the intervention reduced mortality by 40% [32%].

The mean utilization in the control group in 2005 was 660 with a standard deviation of 175. The *participation & information intervention* therefore has 80% [60%] power to detect a significant difference at the 0.05 significance level if the intervention increased utilization by 21% [17%]. The *participation intervention* has 80% [60%] power to detect a significant difference at the 0.05 significance level if the intervention increased utilization by 29% [23%].

As a reference point, Björkman and Svensson (2009) show that the *participation & information intervention* reduced under-five mortality by an estimated 35% and increased utilization by 20-29% after one year. Thus, while the power to pick up significant treatment effects in the *participation intervention* is lower than that of the *participation & information intervention*, both experiments are powered to pick up effects of similar size as those reported in Björkman and Svensson (2009).

3.2 Interventions

The aim of the research project was twofold. First, to evaluate the long-run impact of the community monitoring intervention initiated in 2004 (the *participation & information intervention*). Second, to assess what role and impact dissemination of quantitative information on the staff's performance has in these types of beneficiary control programs. To this end, we designed a new intervention in 2007 – the *participation intervention*.

Efforts to stimulate beneficiary involvement, like the Community Driven Development (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 through facilitated meetings. Both the *participation intervention* and the *participation & information intervention* largely followed this approach. The key difference between the two interventions is that the *participation & information intervention* also included the dissemination of a report card with data 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 *participation & information intervention* (reported in Björkman and Svensson, 2009);

- The long run, i.e., four years after the initial intervention, treatment effect of the *participation & information intervention*;
- 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 2005-2008, the short-run treatment effects are estimated over two consecutive time periods (2005 and 2007-2008, respectively). Holding context and timing constant, we can compare the long run (4 years) treatment effects of the *participation & information intervention* with the shorter run (2 years) treatment effects of the *participation intervention* and holding context constant, we can compare the short run effects of two interventions.

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 community members from the catchment area and from all spectra of society and with on average more than 150 participants per day and per community attending; (ii) a health facility meeting – a half-day event, usually held in the afternoon at the health facility, with all staff attending; and (iii) an interface meeting – a half-day event with representatives from the community and the staff attending.⁸

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, including 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.⁹ Second, in the interface meeting, representatives from the community and health facility staff presented and discussed their suggestions and in a facilitated meeting, agreed on a set of issues that were viewed as most critical to address. These issues were put into a joint action plan. The action plan outlines the community's

⁸Altogether, 18 CBOs, each receiving 10 days of training, implemented the two interventions.

⁹Focus group discussions were held with sub-groups (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 aimed at both identifying and aggregating the 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 on which the provider and the community had differential views.

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. After the 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 participation & information intervention

The *participation & information intervention* mirrored the participation intervention with one important exception. Specifically, at the start of both the community and the health facility meetings, the facilitators provided the participants with 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 non-literate process the provided information.

The *participation & information intervention* was initiated at the beginning of 2005 and was followed by small scale repeat engagements in mid-2005, 2007 and 2008 as illustrated in figure 1. In between this fairly "minimal" intervention (a total of 10 days in 4 years), the communities were left to themselves to monitor the providers as outlined in the agreed upon action plan.

3.3 Data

Data collection was governed by two objectives. The first objective was to create report cards for the *participation & information intervention* on staff performance 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 post-treatment 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 of children aged below 5 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 2004 and onward.

In addition to facility and household surveys, we have information from the action plans for the treatment group in the *participation & information intervention* at the start of the intervention in 2005 and mid-way through when they were given the opportunity to revise their plans. We have similar data for the treatment group in the *participation intervention* at the start of the intervention in mid-2007.

3.4 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. Then, we turn to the quantity of health care. To measure utilization, we use data from 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 from the communities and information exchanges, 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 through improved 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 *participation & information intervention* (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.¹⁰

¹⁰These data were only collected in 2009.

To assess the causal effect of the intervention, we compare mean outcomes, after accounting for stratification. That is, we estimate

$$(1) \quad y_{ijd} = \alpha + \beta T_{jd} + \theta_d + \varepsilon_{ijd}$$

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 a subset of variables, we can also stack the pre- and post-data and estimate the following difference-in-difference specification

$$(2) \quad y_{ijt} = \gamma POST_t + \beta T_j \times POST_t + \mu_j + \varepsilon_{ijt}$$

where $POST$ is an indicator for the follow-up survey and μ_j is a facility or community fixed effect.

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 k^{th} outcome regression and $\hat{\sigma}_k$ is the standard deviation of the control group for outcome k (see Duflo et al. 2007).

4 Long-run impact of the participation & information intervention

4.1 Balance at baseline

Tables 1 and 2, panel A, report mean pre-treatment characteristics for the treatment and control groups in the *participation & information intervention* and test statistics for equality of the means. We report both mean differences for a set of key outcome variables and differences in average standardized pre-treatment effects for each family of outcomes (utilization, utilization pattern, quality, catchment area statistics, health facility characteristics, citizen perceptions, supply of drugs, and user charges). There is no systematic difference between the treatment and the control group at the baseline. Thus, overall the sample is balanced.

4.2 Health outcomes and quantity of care

Health outcomes

The primary outcome measure for the trial was under-five 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 summary measure of health and nutrition since conception, while weight is 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.

Table 3 depicts the findings on child mortality for the period 2006-2009. 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. The intervention reduced the number of deaths in all three age categories. The number of under-five deaths dropped by 32% (column i); the number of infant deaths dropped by 36% (column ii); and the number of neonatal deaths dropped by 51% (column iii).

The crude death numbers provide evidence of a fall in child mortality, but the reduction may not necessarily be due solely to a reduction in the risk of child death as cohort sizes may have been differentially affected by the intervention (for instance due to differential fertility rates). To account for this we estimate a mortality rate for children under-five (and infants) over the period of exposure; i.e., between January 2006 and May 2009. We follow the conventional approach used in epidemiology and define the mortality rate as the number of under-five (under 12 months) children that died during the period per per 1000 child-years of exposure over the same time period. We also, conventionally, define neonatal mortality as the number of neonatal deaths per 1000 live births. The results, with the data collapsed at the cluster level, are reported in columns iv-vi. We report treatment effects and for under-five and infant mortality also the rate ratio of the incidence of child deaths (occurrence of death over child-months) in the treatment versus the control group.¹¹

Column iv reports the under-five mortality effects. The estimated rate ratio; i.e., the ratio of the incidence of child deaths in the treatment relative the control group, implies a 23% reduced risk of under-five deaths in the treatment group. The effect is of the same order of magnitude, but less precisely estimated, using a linear model (a reduction of 4.43 deaths per 1000 child-years from a control group mean of 18.7 deaths per 1000). The reduction in infant mortality (column v) is slightly larger – a 28% reduced risk of infant deaths in the treatment versus the control group – but also less precisely estimated.¹² Column vi reports the treatment effect on neonatal mortality.

¹¹The rate ratios are estimated using a Poisson regression model (Preston, 2005).

¹²Under-five and infant mortality rates can also be estimated using a Poisson model with individual (child level)

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%.¹³ As a reference point, the rate ratios for the short run impact of the *participation & information intervention* on under-five and infant mortality are 0.74 and 0.68, respectively. That is, the reduction in child mortality in the longer run is similar to the short run findings, albeit somewhat smaller in magnitude.¹⁴

Table 4 reports the findings on births and pregnancies. The intervention reduced the incidence (column i) and the number of births since 2006 (column ii) by 14% – an effect roughly twice as large in absolute values as compared to the short-run effect reported in Björkman and Svensson (2009). Columns (iii)-(iv) show that the effect is of the same order of magnitude – about two times larger in absolute values in the long (4 years) versus the short (1 year) run – when using the incidence and the number of pregnancies as dependent variables.

We measured the weight of children under-five using portable scales and the height of children under-five years using stadiometers. Table 5 reports the long-run intervention impacts on the two anthropometric outcomes. We report results separately for infants (0-11 months) and children (12-59 months) because reductions in illnesses tend to lead to gains in weight for infants while reductions in illnesses over time, and thus for older children, will tend to manifest in increased height. Columns (i-iii) depict the weight-for-age z-score results. Consistent with the findings in Cortinovis et al. (1997), and the short-run findings, Ugandan children have values of weight far lower than the NCHS/CDC international reference.¹⁵ The treatment effect is 0.22 z score in weight-for-age for infants and is fairly precisely estimated. Adding controls for age and gender does not change the results (column iii). There is no effect for children aged 12-59 months.

Columns (iv)-(vi) report the results for height of children, using height-for-age z-score as the dependent variable. To the extent that health care provision in the treatment group has experienced a sustained improvement, one could plausibly expect an impact on height. There is a positive treatment effect, significant at the 10 percent level, on height for children aged 12-59 months. Four years into the intervention, children aged 12-59 months in the treatment group are 0.10 z-scores taller than the children of the same age in the control group.

data. The point estimates from a model using individual level data are similar. The rate ratios [p-value] for under-five mortality and infant mortality are 0.77 [0.04] and 0.71 [0.07], respectively.

¹³The estimated neonatal mortality rate in the control group, 27.2, is similar to the neonatal mortality rate reported in official statistics; 25 in 2009. (World Bank Open Data)

¹⁴Björkman and Svensson (2009) did not collect data on the month of death. Thus, to derive these rate ratios we assume that children that died in 2005 died midway through the year. In Björkman and Svensson (2009) the under-five 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.

¹⁵Following Björkman and Svensson (2009), we omit observations with a recorded weight above the 90th 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.

As illustrated in figure 5, and consistent with the finding of a positive treatment effect over the course of the program, the treatment effect is increasing in the exposure to treatment. For example, the treatment effect for children that have lived three years in a treatment area is 0.12 z-score in height-for-age, corresponding to a 10 percent increase compared to the similar aged cohort in the control group.

Utilization

Table 6 depicts the findings on utilization. Cross-section estimates based on equation (1) are given in Panel A. Estimates from a value-added specification, where we use average utilization for outpatient services and deliveries pre-intervention as a lagged variable, are shown in Panel B, and estimates from a difference-in-differences specification, equation (2), are reported in Panel C.¹⁶

Four years after the intervention started, we find large positive effects on all four utilization measures. Although the point estimates are imprecisely estimated in the cross-section specification (see Panel A), the average standardized effect, reported in specification (v), is significantly different from zero. The precision improves and the point estimates are larger in the value-added specification, in which utilization for general outpatient services and deliveries are significantly different from zero at the 10 percent and 5 percent level, respectively. For both the value-added and the difference-in-differences specifications, the average standardized effects are highly significant.

The impact on utilization is substantial. For outpatient services, the increase goes from 16 percent in panel A to 27% in panel B. For deliveries, the point estimate in panel B suggests a 50 increase in utilization. For antenatal care, the increase varies between 21 percent (in panel A) to 25 percent in panel B.

The last three columns in table 7, Panel A and C, 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. As is evident, households in the treatment communities switched from traditional healers and self-treatment, specification (vii), to the project facility, specification (vi), in response to the intervention. The average standardized treatment effects, reported in specifications (viii) and (xix), are significant in both the cross-section and the panel model.

¹⁶We have pre-intervention data for outpatients and deliveries but not antenatal care and family planning. Thus, we 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 $t - 1$. The difference-in-differences models can only be estimated on outpatient services and deliveries for which we have pre-intervention data.

4.3 Processes and health treatment practices

The findings in section 4.2 show that the intervention, which resulted in large improvements in health outcomes and utilization in the short run (Björkman and Svensson, 2009), also resulted in long-run improvements in health. 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, would influence health workers' incentives to provide both more and better health care. We next turn to assessing the evidence for these two links.

Processes

In table 7, we use data on a set of process measures that proxy for the community's ability and extent of monitoring and engagement. The data in panel A were collected through visual checks by the enumerators during the post-intervention health facility survey. Four years after the initial intervention, the average standardized effect of 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. For two of the variables, having a suggestion box and a staff duty roster, the individual effects are also significantly positive and the point estimates, relative to the control group, are fairly large.

Panel B uses information collected at the household level. There are positive long-run effect on all individual indicators; i.e. whether the household has received information about the HUMC's role and responsibilities, whether the performance of health workers was discussed in Local Council meetings, whether a household member told the staff about the household's (dis)satisfaction with the health services received, whether they think that the staff at the facility works closely with the community, and whether they know somebody who monitored the health facility and staff performance.

Health treatment practices

Tables 8-11 focus on health workers' behavior; i.e. treatment practices and management.

Table 8 looks at a set of generic measures of behavior. The intervention had no significant long-run impact on the likelihood that any equipment was used during the patient's last examination, column (i), or waiting time before being examined, column (ii). Note, however, that as utilization increased in the treatment group (see table 6), the total number of patients examined with equipment increased. Column (iii) reports the impact on absenteeism based on data from three unannounced visits. The point estimate is close to zero and insignificant. There is suggestive evidence, however, that the absence rate for staff residing in the community fell in the treatment clinics (the results are available upon request).

The general condition of the clinic – a measure of the conditions of the floor, walls and furniture

and the smell of the clinic – remained significantly better in the treatment group also in the long run (column iv).¹⁷ The effect is substantial, with the mean treatment facility having a 0.5 standard deviation higher score compared to the average control facility. There is also a significant decrease in the probability of drug stock-outs in the long run, column (v), although there is no systematic difference in the supply of drugs between treatment and control groups.¹⁸

Table 9 focuses on the extent to which clinical guidelines for the examination of patients were followed. We collected detailed data on antenatal and postnatal care. Four years after the initial intervention, health workers in the treatment communities appear to provide more effort, or better quality, in their interactions with pregnant women and infants. During antenatal care visits, pregnant women were more likely to be examined by a midwife; to have their weight taken; to have a blood sample taken; to have the fetus checked; and to receive information about pregnancy complications. Newborn children were also more likely to be checked at the facility in the first two months after delivery. For clinical guidelines that are seldom followed in the control group – checking the newborn in the first two months after delivery (a 24 percent increase in the treatment compared to the control group) and measuring weight and taking blood samples of pregnant women (a 22 percent increase in the treatment compared to the control group) – the treatment effects are substantial.

Table 10 includes results on health education on three of the most prevalent diseases in Uganda: HIV/AIDS, tuberculosis and malaria. Column (i) shows, not surprisingly given the mean in the control group of 0.99, that there is no impact on whether household members had heard of AIDS. But the intervention decreased the probability that the household members expressed stigmatization about AIDS (as measured by agreeing with the statement that people living with AIDS should be ashamed of themselves), increased the knowledge that tuberculosis is spread through the air, and increased the probability that children were sleeping under a treated mosquito net.

Table 11 reports the impact on immunization by children’s age group.¹⁹ Two results stand out from the table. First, while the point estimates are positive, the treatment effects are small (between 3%-6%) and insignificantly different from zero. Second, coverage in the control group varies between 78%-94%, which is a doubling of the immunization rates since 2005. As almost nine out of ten children received immunizations according to the recommended immunization

¹⁷The enumerators visually checked the condition of the health clinics; i.e. whether floors and walls were clean, the condition of the furniture and the smell of the facility. From these data, we constructed a summary indicator using a principal components analysis. The variable general condition of the clinic is the standardized (using control group mean and standard deviation) first component.

¹⁸See section 6.3.

¹⁹For 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 1 if child i of cohort (age) j had received the required dose(s) of measles, DPT, BCG, and polio vaccines, respectively, and 0 otherwise. We then estimate equation (3), for each age group, and calculate average standardized effects.

plan, there is little room for improvement. Moreover, the sample sizes in table 11 are insufficient to detect small significant treatment effects at standard levels of significance.²⁰

5 Impact of the participation intervention

The short-run evaluation of the *participation & information* intervention showed that efforts to stimulate community monitoring can have large effects on both the quantity and the quality of service provision and in the end result in improved health outcomes (Björkman and Svensson, 2009). The results presented in Section 4 show that the main results were sustained four years into the intervention. Collecting, assembling, and disseminating information on performance in a rural setting, however, is both technically complicated and costly. Moreover, as the intervention involved both information dissemination and various techniques to stimulate engagement and participation, it is unclear if and to what extent both components are crucial in order to initiate a successful process of community involvement and monitoring.

To address these issues, we designed a new intervention in 2007. The *participation intervention* was designed so as to mimic the *participation & information* intervention in all aspects but one – the participants in the community and health facility meetings did not receive quantitative information on the performance of the provider. Hence, the communities’ informational constraints were not addressed in the *participation intervention*.

5.1 Balance at baseline

As is evident from panel B in table 1, columns (iv)-(vi) and panel B in table 2, column (ii), there are no systematic differences between the treatment and the control group as concerns any of the baseline characteristics. Thus, the sample is balanced. Comparing characteristics of the control groups in the two experiments in the same year (2006) also shows that there was no significant difference across the comparison groups in the two experiments (see Table A.1 for details).

5.2 Main outcomes: Health outcomes and quantity of care

Health outcomes

²⁰The 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 treatment and the control group.

Table 12 reports the short-run findings on health outcomes. The control group means are similar to those reported in tables 3-5 for the *participation & information* intervention, but the treatment effects are small and insignificant.²¹ That is, there are no statistical differences in under-5 mortality (column i); infant mortality (column ii); neonatal mortality (column iii); number of births (column iv); number of pregnancies (column v); and in weight-for-age and height-for age z-scores (columns vi-ix) between the treatment and the control group. Thus, based on health outcomes only, the participation intervention had no impact.

Utilization

Table 13 reports the results on utilization. There is no clear evidence of impact. For general outpatient service, the point estimate goes from -100.2 (p-value=0.38) to -20.3 (p-value=0.89) depending on the specification. For delivery, there is a significant negative effect in two out of three specifications, while the treatment effects are positive, albeit insignificant, for antenatal care and family planning. The average standardized treatment effects are small and insignificantly different from zero in all three specifications.

Columns (vi) and (vii) report changes in utilization patterns based on household data. There is some evidence that treatment households are less likely to use traditional healers and self-treatment, but the average standardized treatment effects are once more insignificant.

5.3 Processes and health treatment practices

The *participation intervention* did not result in improvements in child health or the quantity of health care used. In the appendix, tables A.2-A.6, using the same outcome variables as in section 4.3, we show that there is no systematic evidence that the intervention resulted in increased monitoring and information exchange in the treatment communities. Management of the facility and the extent to which clinical guidelines were followed also remained unchanged.

6 Mechanism and Discussion: The role of information

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

²¹The p-values for the null hypothesis of equal means in the two control groups are 0.65 (for under-5 mortality); 0.62 (for infant mortality); 0.74 (for neonatal mortality); 0.22 (for number of births); 0.28 (for number of pregnancies); 0.50 (for weight-for-age); and 0.17 (for height-for age).

are 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 appeared to have played such a crucial role.²² We start by presenting a stylized example of why information may matter in these kinds of programs. Then, we use process data from the two treatment groups to identify differences in decisions and actions taken in the two treatment groups (with informed and uninformed community members). We discuss alternative mechanisms and also review the literature on beneficiary control and argue that a key difference between successful and unsuccessful community empowerment interventions is access to information on the staff's behavior.

6.1 What role does provision of information play?

Why is access to information about the staff's performance a constraint to the community's ability to monitor and hold providers to account? Data from the baseline survey provide some suggestive evidence. For a set of indicators, like waiting time and absenteeism, we have data both from the in-charge of the facility and data collected either through visual inspections (by enumerators) or collected directly from the users. The comparison reveals a systematic pattern: the in-charge underreport performance problems. For example, the mean waiting time as reported by the clinician was 2 minutes. The household survey data, however, revealed that it was close to 2 hours. Data from unannounced staff surveys showed an absence rate of around 50 percent. In the in-charge survey, however, the clinic officers did 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. Moreover, 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 issues such as weak adherence to clinic guidelines, shirking, or mistreatment of patients were never listed.

Motivated by these examples, we provide a simple stylized model of how the provision of information on mean outcomes, by enabling users to better distinguish between health workers' effort and factors outside the health workers' control, can influence users' willingness to take actions (to monitor the provider) and, in the end, health workers' performance.

Consider a setup where the quality of health service at each visit j and time t , q_{jt} , depends on two factors that are unobserved by the patient: the health worker's performance (p_t) and resources provided from the center (r_{jt}), where

$$(3) \quad q_{jt} = p_t r_{jt} .$$

²²Lack of information was highlighted as a key constraint in the earlier work on community empowerment (see Samuel Paul 1987; Jenkins and Goetz 1999; and Goetz and Jenkins 2001).

Assume that the resource flow from the center, r_{jt} , is a random variable taking the value of 1 with probability μ and 0 with probability $1 - \mu$. The health worker's performance is a function of the health worker's effort (e_t). With probability e , the performance is "appropriate" or "high" ($p = 1$) and with probability $1 - e$, the performance is "inappropriate" or "low" ($p = 0$). The health worker chooses an effort which costs him $c(e) = \frac{1}{2}e^2$.

There are two types (T) of health workers: an honest type (H), who always chooses $e = 1$, and an opportunistic type (O), who chooses e so as to maximize his expected utility. Nature chooses type at the start of the period with a worker of type O chosen with probability $\phi = 1/2$. As an honest type always chooses high effort, our focus is on the opportunistic type.

Patients do not observe p_t , e_t , or r_{jt} . However, they know the distribution of r_j and ϕ and observe q_{jt} after each visit. Patients seek service at the health center and care about the (expected) quality of health service q_{jt} and the cost of taking local actions. As our focus is on the role of information, we disregard collective action problems and can then drop subscript j and just assume that there is one representative community member (patient). We treat the representative patient as myopic, in the sense that she does not aggregate information over visits. Thus, the representative patient observes $q_t = 1$ in period t with probability $e\mu$ and $q = 0$ with probability $1 - e\mu$. Based on the outcome, the patient can take local actions at a cost κ , which inflicts a cost $\lambda(1 - e)$ on the provider, where $\lambda \in (0, 1)$. We can think of the costs $\lambda(1 - e)$ as capturing the shame a low performing health worker would suffer if the community expressed its displeasure with his performance. The costs are proportional to $(1 - e)$ as the cost inflicted on the health workers is increasing in the extent of shirking.

An opportunistic health worker's (per-period) payoff π is

$$(4) \quad \pi = w - c(e) - L\lambda(1 - e)$$

where w is the wage rate and L is a binary variable taking the value of 1 if the community decides to take actions against the provider.

The sequence of events is as follows. Nature chooses type at the start of period 1. The health worker then makes an unobserved choice of effort (e_1) and the resource flow from the center (r) is realized. The (representative) patient experiences the quality of health service, q_1 , and updates her beliefs about the type of health worker. Depending on the outcome, the users then make a decision of whether or not to take local action and the health worker makes a new effort choice (e_2). Finally the patient experiences the quality of health service in period 2 (q_2). We denote by $\phi(\phi|q_1, \tilde{e}^T)$ or ϕ_q the posterior probability that the provider is an opportunistic type, given a realized outcome q_1 , a prior probability ϕ , and the user's expectation of the health worker's choice of effort conditional on type T (\tilde{e}^T).

Assumption 1: $\mu > \kappa/\lambda > \mu/(2 - \mu)$

Scenario 1: Suppose that no information is provided. Then, provided that assumption 1 holds, it is optimal for the community not to take any local actions. The (opportunistic) health worker exerts no effort ($e_1 = e_2 = 0$) and the quality of health service is low ($q_1 = q_2 = 0$).²³

Scenario 2: Suppose that patients receive information on the mean quality of health service; i.e. $E_q[q] = \bar{q}$. As expected quality conditional on the worker being an honest type is $E_q[q|H] = \mu$ and an opportunistic type at most exerts effort $e \leq \lambda < 1$, implying that $E_q[q|O] \leq \lambda\mu$, patients can tell with certainty, given \bar{q} , whether the worker is an honest or an opportunistic type. If assumption 1 holds, the community will therefore take local actions and the health worker will respond by exerting a higher effort ($e_2 = \lambda$). The expected period 2 quality is $\lambda\mu$.

Intuition: Provision of information solves patients' aggregation problem. This enables users to distinguish between health workers' type or effort and factors outside the health workers' control. As a result, the return to local action increases. In equilibrium, community members become more involved; i.e., take local actions, and the health worker responds by exerting higher effort. Note that assumption 1 does not hold for μ close to 0 and 1. In the first case, public service provision is so poor that the health worker's choice of effort has almost no bearing on health quality, so there are low returns to incentivize the provider. In the second case, the community member can tell with certainty, if observing $q = 0$, that the health worker is an opportunistic type so that there is no aggregation problem. Moreover, the difference between the two thresholds, $\mu - \mu/(2 - \mu)$, has an inverse U-shape with a maximum at $\mu = 2 - \sqrt{2}$; that is, provision of information is most likely (i.e. the parameter space for which assumption 1 holds is the largest) to have an effect when public support to frontline providers is neither too effective or too poor.

6.2 Information and local actions

The above stylized example shows how the provision of information on mean outcomes, by enabling users to better distinguish between health workers' effort and factors outside the health workers' control, can influence the community members' willingness to take actions. In table 14 we use data from the implementation phases of the two interventions to corroborate this prediction.

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 14, 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 i). The two treatment groups also identified roughly the same number of actions

²³See the appendix for details.

to be addressed (column ii). 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 a timely delivery of medicines from the center, the *participation & information 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 iii and iv). Even two years into the *participation & information intervention*, 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 v). Figure 6 illustrates the main result.

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 beneficiary 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. But the findings do not rule out other explanations. 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 before and after treatment in treatment clinics as compared to the mean difference in the control clinics as the dependent variable; i.e., $\bar{y}_j = (y_{tj}^T - y_{t-1j}^T) - (\bar{y}_t^C - \bar{y}_{t-1}^C)$, and regress \bar{y}_j on the share of local issues raised in the action plan. The estimated relationship between the difference in the number of outpatients served before and after treatment compared to control clinics, conditional on 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 *participation & information 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.

6.3 Alternative mechanisms

The findings of large treatment effects in the *participation & information* 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 sub-district management, may have changed their 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 any evidence of this being the case (see the appendix, table A.7). The treatment facilities did

not receive more drugs or funding from the sub-district 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.

6.4 Discussion

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 (Mansuri and Rao, 2013) notes 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.

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 card information 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 an intervention involving only the standard package of facilitated meetings to enhance participation. The focus on participation alone is typical for most Community Driven Development (CDD) programs.

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 almost exclusively identifying issues that either health workers or the users could address themselves, and significant improvements in health care delivery and health outcomes in both the short and longer run.

The results of the two field experiments resonate 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. No information was provided about the teachers' performance. While test scores are likely to be 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 know, 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.²⁴ Thus, it is unclear whether the non-elite 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 et al. (2012) 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 sub-committee of parents was formed to evaluate the contract teacher and deliver a performance report at the end of the year. 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.²⁵ SBM training for PTA committees reduced teachers' absenteeism and increased student test scores.

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

²⁴Corruption is not easily observable and measurable, as evidenced by Olken's (2007) novel but burdensome strategy to estimate it.

²⁵In addition, the intervention provided the PTA with hard incentives (the right to fire or retain the contract teacher).

outcomes.

Collecting data on performance using traditional survey methods is a costly and to some extent technically complex endeavor. Thus, although our findings point to large and sustained positive health effects, it is unclear whether such an intervention is possible to scale up. This 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 et al. (2012), 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?

References

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

Björkman, Martina, and Jakob Svensson, "Power to the People: Evidence from a Randomized Experiment on Community-Based Monitoring in Uganda", *Quarterly Journal of Economics*, 124:2 (2009): 735–769.

Bold, T., B. Gauthier, J. Svensson, and W. Wane, 2011, Service Delivery Indicators: Pilot in Education and Health Care in Africa, Report, The World Bank.

Casey, Katherine, Rachel Glennerster, and Edward Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan", 127(4) (2012): 1755-1812.

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

Cortinovis, I., V. Vella, N. Ndiku, and S. Milani, "Weight, Height and Arm Circumference of Children Under 5 in the District of Mbarara, South-west Uganda," *Annals of Human Biology*, 24 (1997), 557 - 568.

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

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools", 2012, NBER Working Paper #17939.

Duflo, Esther, Rachel Glennerster, and Michael Kremer, "Using Randomization in Development Economics Research: A Tool Kit," CEPR Working Paper No. 6059, 2007.

Jenkins, Rob and Anne Marie Goetz, "Accounts And Accountability: Theoretical Implication of Right to Information Movement in India." *Third World Quarterly*, 20(3) (1999), pp. 603-22.

Jones, Gareth, Richard W. Steketee, Robert E. Black, Zulfiqar A. Bhutta, and Saul S. Morris, Bellagio Child Survival Study Group, "How Many Child Deaths Can We Prevent This Year?," *Lancet*, 362 (2003), 65-71.

Kling, Jeffrey R., Jeffrey B. Liebman, Lawrence F. Katz, and Lisa Sanbonmatsu, "Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment," Princeton University Working Paper No. 5, 2004.

Lawn, Joy, Hannah Blencowe, Robert Pattinson, Simon Cousens, Rajesh Kumar, Ibinabo Ibiebele, Jason Gardosi, Louise T Day, Cynthia Stanton. 2011. "Stillbirths: Where? When? Why? How to make the data count?", *Lancet*, 377, 9775: 1448–1463.

Mansuri, Ghazala, and Vijayendra Rao. 2013. *Localizing Development: Does Participation Work?* Washington, DC: World Bank. DOI: 10.1596/978-0-8213-8256-1. License: Creative Commons Attribution CC BY 3.0.

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

Paul, Samuel, "Community Participation in Development Projects: The World Bank Experience", World Bank Discussion Papers 6, 1987, Washington, D.C. : World Bank.

Preston, D.L., 2005, Poisson regression in epidemiology, *Encyclopedia of Biostatistics*, Wiley Online Library.

Wong, Susan, "World Bank Community-Driven Development Programs? CDD Impact Evaluation Review and Operational and Research Implications" (Washington, DC: The World Bank, Social Development Department, Sustainable Development Network, 2012).

World Bank, *World Development Report 2004: Making Service Work for the Poor People* (Washington, DC: World Bank and Oxford University Press, 2003).

FIGURE 1: TIMELINE FOR THE PARTICIPATION & INFORMATION INTERVENTION

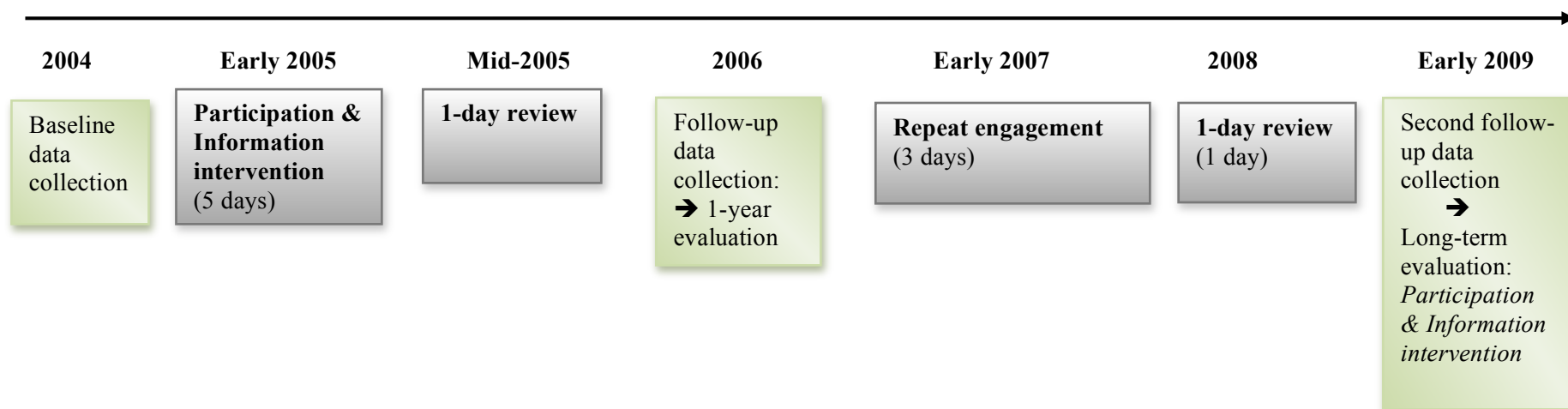


FIGURE 2: TIMELINE FOR THE PARTICIPATION INTERVENTION

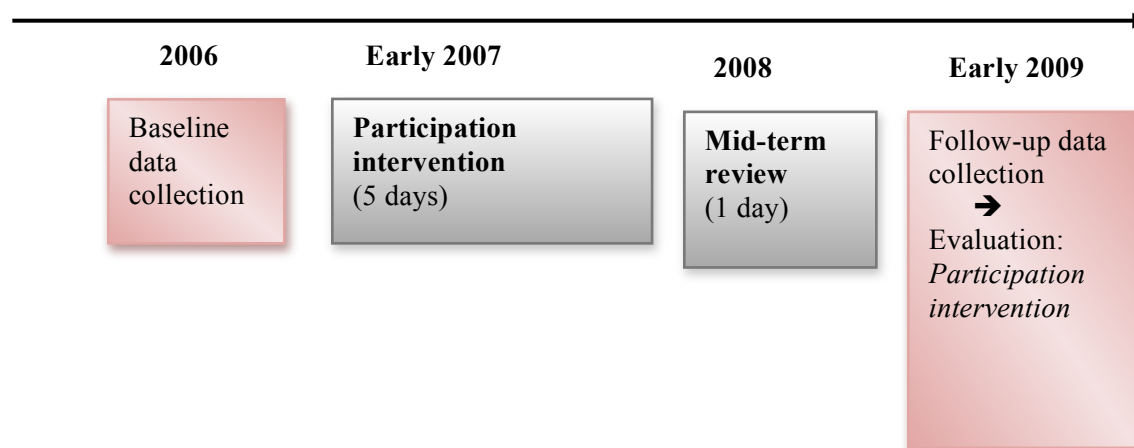


FIGURE 3: DATA COLLECTION

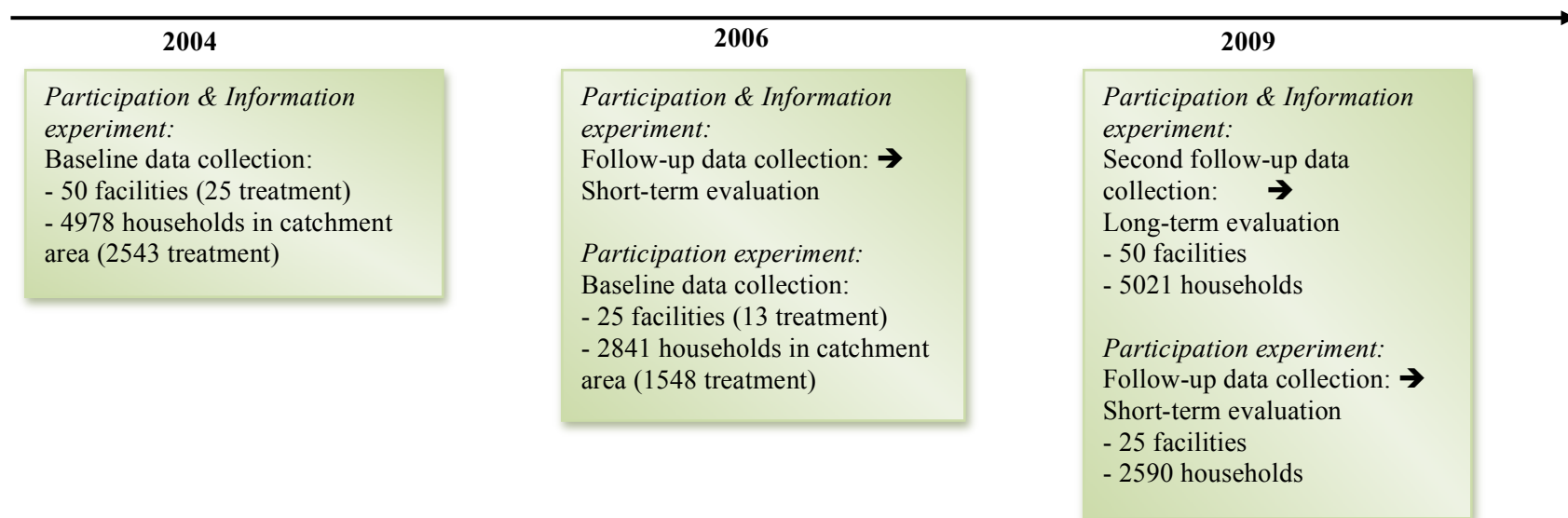
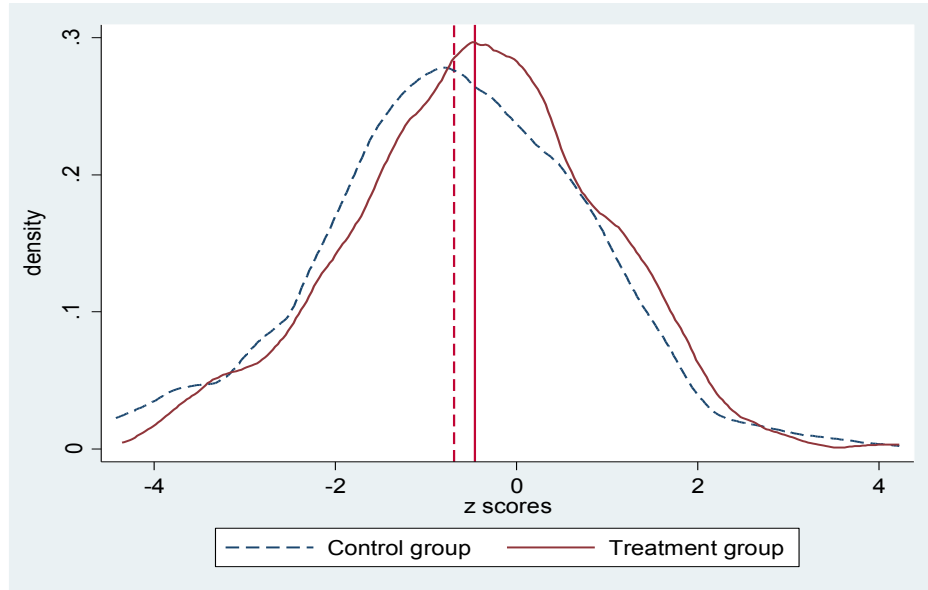
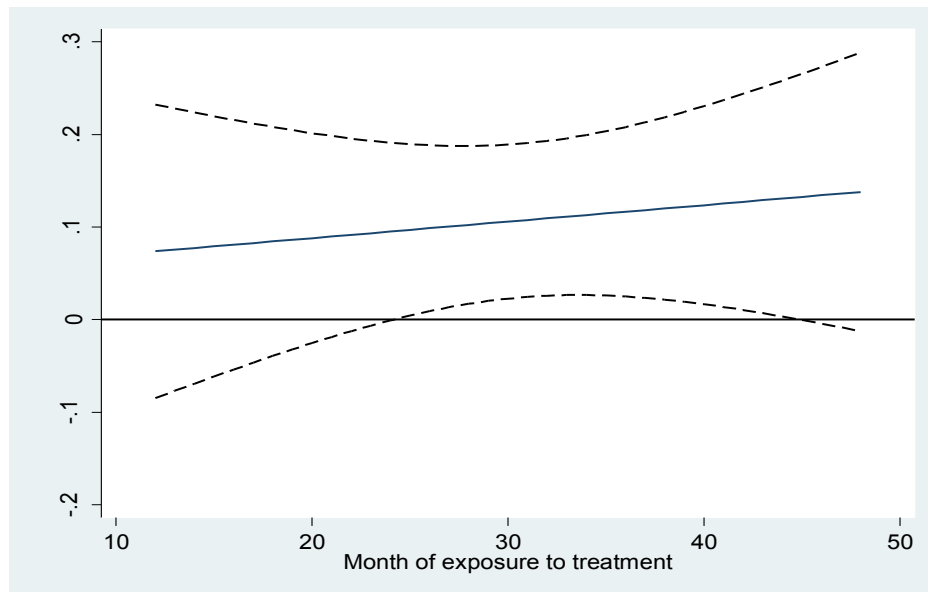


FIGURE 4: DISTRIBUTION OF WEIGHT-FOR-AGE Z-SCORES FOR THE TREATMENT AND CONTROL GROUPS – PARTICIPATION & INFORMATION MODEL (2005-2008)



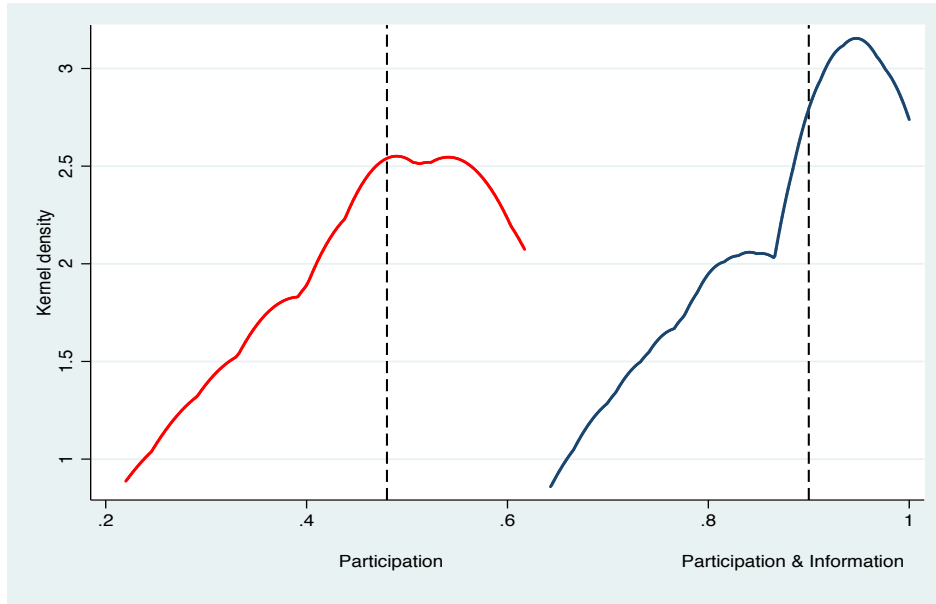
Notes: Weight-for-age z-scores for children under 12 months excluding observations with recorded weight above the 90th percentile in the growth chart reported in Cortinovis et al. (1997). Sample size is 669 children. Solid line depicts the distribution for the treatment group and dashed line the distribution for the control group. Vertical solid line denotes mean in treatment group; dashed line denotes mean in control group.

FIGURE 5: TREATMENT EFFECT ON HEIGHT-FOR-AGE Z SCORES CONDITIONAL ON TIME OF EXPOSURE - PARTICIPATION & INFORMATION MODEL



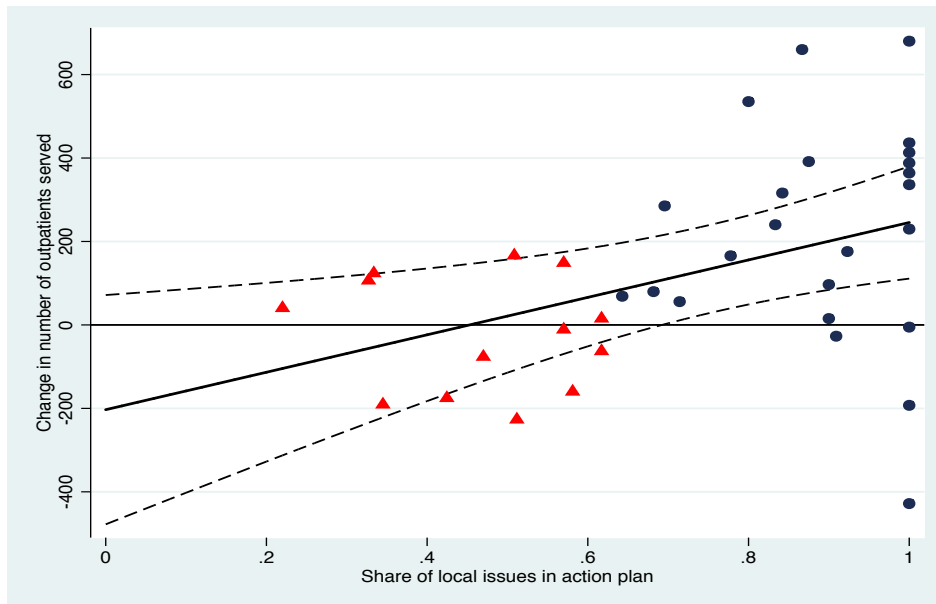
Notes: Figure plots the treatment effect conditional on months of exposure (solid line) and the 95 confidence intervals (dashed lines); i.e. the derivative $\partial z / \partial T$ of equation (1) augmented with *time* (in month) of exposure to treatment (0-48 months) and *time* \times *T*, where *z*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

FIGURE 6: LOCAL ACTIONS FOR THE TWO TREATMENT GROUPS



Notes: 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.

FIGURE 7: OUTPATIENT CARE CONDITIONAL ON THE SHARE OF LOCAL ACTION



Notes: Figure plots the difference in outpatient served before and after treatment in treatment clinics (*participation intervention* in red triangles and *participation & information intervention* in blue dots) compared to control clinics, conditional on the share of local actions agreed upon in the action (solid line) and the 95 confidence intervals (dashed lines).

TABLE 1: PRE-TREATMENT CHARACTERISTICS IN THE PARTICIPATION & INFORMATION EXPERIMENT (2004) AND PARTICIPATION EXPERIMENT (2006)

Variables	Panel A			Panel B		
	2004			2006		
	<i>Participation & Information sample</i>			<i>Participation sample</i>		
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
	Treatment group	Control group	Difference	Treatment group	Control group	Difference
Outpatient care	593 (373)	675 (286)	-82 (94)	781 (300)	790 (330)	-9 (127)
Delivery	10.3 (11.1)	7.5 (6.8)	2.8 (2.6)	12.4 (6.3)	15.9 (10.7)	-3.5 (3.5)
No. of households in catchment area	2140 (927)	2224 (1021)	-84.4 (276)	2850 (1218)	2519 (1144)	331 (473)
No. of households per village	93.9 (26.4)	95.3 (31.6)	-1.42 (8.23)	121.9 (38)	118.1 (41.7)	3.75 (16.0)
Drink safely today	0.40 (0.50)	0.32 (0.48)	0.08 (0.14)	-	-	-
Days without electricity	18.3 (14.8)	20.4 (14.5)	-2.12 (4.14)	-	-	-
Working water source	-	-	-	0.92 (0.28)	1.00 (0.00)	-0.08 (0.08)
Functioning electricity (facility)	-	-	-	0.46 (0.52)	0.42 (0.51)	0.04 (0.21)

Notes: In each panel and row we report catchment area/health facility averages for treatment and control group and difference in averages. Standard deviations are reported in parenthesis below the averages and robust standard errors are reported in parentheses under the differences (columns (iii) and (iv)). Columns (i)-(iii) corresponds to the 2004 baseline characteristics for the Participation and Information Experiment and columns (iv)-(vi) uses 2006 baseline data for the Participation Experiment. 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. Drink safely today is an indicator variable for whether the health facility staff at the time of the pre-intervention survey could safely drink from the water source. Number of days without electricity in the month prior to pre-intervention survey is measured for the last month out of 31 days. Working water source is an indicator variable for whether the health facility in 2006 had a functioning water source. Electricity at the health facility is an indicator variable for whether the main source of electricity health facility functioned during the survey.

TABLE 2: AVERAGE STANDARDIZED PRE-TREATMENT EFFECTS IN 2004 AND IN 2006

	Panel A	Panel B
	2004 <i>Participation & Information sample</i>	2006 <i>Participation sample</i>
Variables	Difference between treatment and control groups	Difference between treatment and control groups
Utilization from health facility records	0.11 (0.77)	0.18 (0.29)
Utilization pattern of the users	-0.48 (0.33)	-0.02 (0.03)
Quality of services according to users	-0.35 (0.84)	-0.02 (0.08)
Catchment area statistics	0.11 (0.66)	0.43 (0.39)
Health facility characteristics	0.14 (0.31)	-0.17 (0.19)
Citizen perceptions of treatment	0.37 (0.67)	0.03 (0.07)
Supply of drugs	0.73 (0.83)	0.47 (0.43)
User charges	-0.65 (0.63)	0.40 (0.30)

Notes: In each panel and row we report an average standardized pretreatment effect (see text for details). Robust standard errors are reported in parentheses. Column (i) corresponds to the 2004 baseline data for the Participation and Information Experiment and column (ii) uses 2006 baseline data for the Participation Experiment. Description of variables: Utilization from health facility records summarizes outpatients and deliveries. Utilization pattern of the users summarizes seven measures including 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 according to users summarizes the use of any equipment during the visit 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 uses different data in columns (i) and (ii) because of differences in data collected at the time of each baseline survey. For column (i) it summarize ten measures about the presence of piped water, access to a radio, a newspaper, the existence of a separate maternity unit, the distance to the nearest Local Council I and to the nearest public health provider, number of staff with advanced A-level education and with less than A-level education, drank safely today and days without electricity, reversing sign of days without electricity and distance to nearest local council. For column (ii) it summarizes six measures about the presence of piped water, working water source, functioning electricity, yellow star certification of the health facility, number of staff with advanced A-level education and with less than A-level education. Citizen perceptions of treatment summarize four measures about politeness, attention, freedom to express themselves and information about drug deliveries. Supply of drugs summarizes five measures about the availability of erythromycin, chloroquine, septrine, quinine and mebendazole. User charges summarize four measures about the existence of user charges for drugs, general treatment, injections and deliveries, reversing all signs.

TABLE 3: PROGRAM IMPACT ON CHILD MORTALITY – PARTICIPATION & INFORMATION MODEL (2006-2008)

Dependent variable	Number of under-5 deaths	Number of infant deaths	Number of neonatal deaths	Under-5 mortality (rate ratio)	Infant mortality (rate ratio)	Neonatal mortality rate
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Program impact	-1.38* (.71)	-1.12* (.61)	-0.93** (.36)	-4.43* (2.60)	-12.3 (7.92)	-11.4** (5.3)
Rate ratio				0.77** (.10)	0.72* (.13)	
Mean control	4.28	3.12	1.84	18.7	45.0	27.2
Observations	50	50	50	50	50	50

Notes: Specifications (i)-(vi): Estimates from equation (1) with district fixed effects and robust standard errors in parentheses. Specifications (iv)-(v): Rate ratios derived from a Poisson model with district fixed effects and robust standard errors in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Specifications: (i) number of under-5 deaths, 2006-2008; (ii) number of infant deaths, 2006-2008; (iii) number of neonatal deaths, 2006-2008; (iv) number of under-five deaths per 1000 child-years of exposure to the risk of deaths, 2006-2008; (v) number of infant deaths per 1000 child-years of infant exposure to the risk of deaths, 2006-2008; (vi) Number of neonatal deaths per 1,000 live births, 2006-2008. *Significant at 10% level ; **Significant at 5% level ; ***Significant at 1% level.

TABLE 4: PROGRAM IMPACT ON HEALTH OUTCOMES (BIRTHS AND PREGNANCIES)
– PARTICIPATION & INFORMATION MODEL (2005-2008)

Dependent variable	Any birth	Number of births	Any pregnancy	Number of pregnancies
Specification	(i)	(ii)	(iii)	(iv)
Program impact	-0.048** (0.018)	-0.061** (0.025)	-0.051*** (0.017)	-0.071*** (0.026)
Mean control group	0.35	0.43	0.39	0.51
Observations	5886	5886	5886	5886

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. Specifications: (i) indicator variable for whether any woman in the household has given birth to a child in 2006-2008; (ii) number of births in the household 2006-2008; (iii) indicator variable for whether any woman in the household is or was pregnant 2006-2008; (iv) number of pregnancies in 2006-2008.

*Significant at 10% level; **Significant at 5% level; ***Significant at 1%

TABLE 5: PROGRAM IMPACT ON HEALTH OUTCOMES (WEIGHT AND LENGTH OF CHILDREN) – PARTICIPATION & INFORMATION MODEL (2005-2008)

Dependent variable	Weight-for-age z-scores (0-11m)	Weight-for-age z-scores (0-11m)	Weight-for-age z-scores (12-59m)	Height-for-age z-scores (0-11m)	Height-for-age z-scores (12-59m)	Height-for-age z-scores (12-59m)
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Program impact	0.22** (0.10)	0.20* (0.10)	0.01 (0.07)	0.04 (0.10)	0.10* (0.05)	0.10* (0.05)
Child age (log)	-	-0.51*** (0.08)	-	-	-	0.22*** (0.07)
Female	-	0.17 (0.12)	-	-	-	0.18*** (0.05)
Mean control	-0.70	-0.70	-0.92	-1.25	-1.59	-1.59
Observations	669	669	2839	565	2451	2451

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. Specifications: (i)-(ii) WAZ for children 0-11 months excluding observations with recorded weight above the 90th percentile in the growth chart reported in Cortinovis et al. (1997); (iii) Weight-for-age z-scores (WAZ) for children 12-59 months excluding observations with recorded weight above the 90th percentile in the growth chart reported in Cortinovis et al. (1997); (iv) HAZ for children 0-11 months excluding observations with recorded height above the 90th percentile in the growth chart reported in Cortinovis et al. (1997); (iv)-(v) Height-for-age z-scores (HAZ) for children 12-59 months excluding observations with recorded height above the 90th percentile in the growth chart reported in Cortinovis et al. (1997). *Significant at 10% level ; **Significant at 5% level ; ***Significant at 1% level.

TABLE 6: PROGRAM IMPACT ON UTILIZATION/COVERAGE USING FACILITY AND HOUSEHOLD DATA– PARTICIPATION & INFORMATION MODEL (2005-2008)

Dep. variable	Out-patients	Delivery	Antenatal	Family planning	Average std effect	Use of Project facility	Self-treatment/ traditional healer	Average std effect
<u>PANEL A: Cross section</u>	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Program impact	97.5 (63.8)	6.03 (3.70)	18.0 (14.5)	4.89 (5.45)	0.37* (0.21)	0.018* (0.011)	-0.013* (0.007)	0.049** (0.021)
Observations	50	50	50	50	50	4981	4981	4981
Mean control	598.7	13.0	84.8	26.4		0.30	0.11	-
<u>PANEL B: Value added</u>	(ix)	(x)	(xi)	(xii)	(xiii)			
Program impact	110.1* (62.4)	7.30** (3.67)	21.4 (15.3)	4.86 (5.30)	0.43** (0.21)			
Observations	50	50	50	50	50			
<u>PANEL C: Panel data</u>	(xiv)	(xv)			(xvi)	(xvii)	(xviii)	(xix)
Program impact	162.4** (70.0)	3.70 (2.40)			0.50*** (0.17)	0.037 (0.022)	-0.06* (0.03)	0.17** (0.08)
Observations	100	100			100		9464	9464

Notes: Panel A reports program impact estimates from a cross-sectional model using data from the end line year (2009) with district fixed effects and robust standard errors in parentheses. Panel B reports program impact estimates from a value-added specification, using data from the end line year (2009) and average utilization for outpatient services and deliveries pre-intervention (2004) as lagged variable, with robust standard errors in parentheses. Panel C reports program impact estimates from a difference-in-differences specification with robust standard errors in parentheses. Program impact measures the coefficient on the assignment to treatment indicator in panels A and B, and the assignment to treatment indicator interacted with an indicator variable for 2009 in panel C. Specifications: First column is average number of patients visiting the facility per month for outpatient care; second column is average number of deliveries at the facility per month; third column is average number of antenatal visits at the facility per month; fourth column is average number of family planning visits at the facility per month; fifth column is the average standardized effect of the estimates in specifications (i)-(iv), (ix)-(xiii), and (xiv)-(xv), respectively; sixth column is the households' share of visits to the project facility out of all health visits; seventh column is the households' share of visits to traditional healers and self-treatment out of all health visits; eighth column is the average standardized effect of the estimates in specifications (vi)-(vii) and (xvii)-(xviii), respectively, reversing the sign of use of self-treatment or traditional healers. *Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

TABLE 7: PROGRAM IMPACT ON MONITORING AND INFORMATION – PARTICIPATION & INFORMATION MODEL (2005-2008)

<u>PANEL A:</u>						
Dependent variable (health facility level)	Suggestion Box	Numbered waiting cards	Staff duty roster	Poster informing free services	Poster on patients' rights	Average std effect
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Program impact	0.21** (0.10)	0.00 (0.07)	0.17* (0.09)	-0.02 (0.08)	0.08 (0.07)	0.33** (0.16)
Mean control group	0.08	0.08	0.08	0.12	0.04	-
Observations	50	50	50	50	50	50
<u>PANEL B:</u>						
Dependent variable (household level)	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	(vii)	(viii)	(ix)	(x)	(xi)	
Program impact	0.07*** (0.02)	0.24*** (0.05)	0.04* (0.02)	0.03** (0.015)	0.11*** (0.03)	
Mean control group	0.14	0.72	0.25	0.15	0.77	
Observations	5002	2522	3823	4294	4991	

Notes: Robust standard errors in parentheses. Error terms are clustered by catchment areas in columns (vii)-(xi). Point estimates, standard errors, and average standardized effect, columns (i)-(vi), are derived from equation (1). Program impact measures the coefficient on the assignment to treatment indicator. Outcome measures in columns (i)-(vi) are based on data collected through visual checks by the enumerators during the post-intervention facility survey. Outcome measures in columns (vii)-(xi) are from the post-intervention household survey. Specifications: Column (i) indicator variable for whether the health facility has a suggestion box for complaints and recommendations; (ii) indicator variable for whether the facility has numbered waiting cards for its patients; (iii) indicator variable for whether the health facility has a visible staff duty roster; (iv) indicator variable for whether the facility has a poster informing about free health services; (v) indicator variable for whether the facility has a poster on patients' rights and obligations; (vi) average standardized effect of the estimates in columns (i)-(v); (vii) indicator variable for whether the household has received information about the Health Unit Management Committee's (HUMC's) roles and responsibilities; (viii) the sum of indicator variables for whether the household discussed performance, staff and the functioning of the local health facility at the local council meetings during the past year; (ix) indicator variable for whether the household told health facility staff about dis/satisfaction with the quality of service; (x) indicator variable for whether the household thinks health facility staff work closely with the community; (xi) the sum of indicator variables for whether the household knows community members who monitored facility performance (staff attendance, quality of services, cleaning of the health facility, constructions at the health facility). *Significant at 10%. **Significant at 5%. ***Significant at 1%. #denotes average standardized effect.

TABLE 8: PROGRAM IMPACT ON TREATMENT PRACTICES AND MANAGEMENT – PARTICIPATION & INFORMATION MODEL (2005-2008)

Variable	Equipment used	Waiting time	Absence rate	Condition of clinic	Drug stock- outs
Specification	(i)	(ii)	(iii)	(iv)	(v)
	OLS	OLS	OLS	OLS	OLS
Program impact	0.02 (0.03)	7.06 (6.74)	0.02 (0.03)	0.8* (0.45)	-0.06* (0.04)
Mean control group	0.33	113.08	0.46	-0.36	0.57
Observations	3612	3829	50	50	41

Notes: Point estimates and standard errors are derived from equation (1). Robust standard errors clustered by catchment areas in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Specifications: (i) indicator variable for whether the staff used any equipment during examination when the patient visited the health facility; (ii) difference between the time the patient left the facility and the time the patient arrived at the facility, minus the examination time; (iii) ratio of workers not present at the time of three unannounced visits during three months in 2009 to the number of workers employed (see text for details); (iv) first component from a principal components analysis of four variables: condition of the floors of the health clinic, condition of the walls, condition of furniture, and smell of the facility, where each condition is ranked from 1 (dirty) to 3 (clean) by the enumerators; (v) share of months in 2009 in which stock cards indicated no availability of drugs (see text for details). *Significant at 10%. **Significant at 5%. ***Significant at 1%.

TABLE 9: PROGRAM IMPACT ON ANTENATAL AND POSTNATAL CARE – PARTICIPATION & INFORMATION MODEL (2005-2008)

Dependent variable	Examined by midwife during antenatal visit at the health facility	Weight taken during antenatal visit	Blood sample during antenatal visit	Fetus checked during antenatal visit	Told about pregnancy complications during antenatal visit	Newborn checked at the health facility in the first two months after delivery	Average Standardized Effect
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Program impact	0.09** (0.03)	0.13*** (0.04)	0.13** (0.05)	0.03*** (0.01)	0.06** (0.03)	0.09** (0.03)	0.20*** (0.05)
Mean control	0.87	0.59	0.52	0.95	0.52	0.37	-
Observations	1152	1151	1146	1148	1148	1711	992

Notes: Point estimates and standard errors are derived from equation (1) with district fixed effects. Robust standard errors clustered by catchment areas in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. The data is from the household survey of women who experienced a pregnancy in the last 2 years. Specifications: (i) indicator variable for whether the woman was examined by midwife during the antenatal visit; (ii) indicator variable for whether the woman's weight was taken during the antenatal visit at the health facility; (iii) indicator variable for whether a blood sample was taken during the antenatal visit; (iv) indicator variable for whether the was checked during the antenatal visit; (v) indicator variable for whether the woman was told about pregnancy complications; (vi) indicator variable for whether the newborn was checked at the health facility in the first two months after delivery; (vii) average standardized effect of the estimates in columns (i)-(vi). *Significantly different from zero at 90% confidence level; **Significantly different from zero at 95% confidence level; ***Significantly different from zero at 99% confidence level.

TABLE 10: PROGRAM IMPACT ON HEALTH EDUCATION– PARTICIPATION & INFORMATION MODEL (2005-2008)

Dependent Variable	Have heard about AIDS	AIDS stigma	Knowledge that TB is spread through the air	Children sleeping under a treated mosquito net
Specification	(i)	(ii)	(iii)	(iv)
Program impact	0.00 (0.00)	-0.02*** (0.01)	0.04** (0.02)	0.06** (0.02)
Mean control group	0.99	0.09	0.53	0.38
Observations	5002	4907	4880	2925

Notes: Point estimates and standard errors are derived from equation (1) with district fixed effects. Robust standard errors clustered by catchment areas in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Specifications: (i) indicator variable for whether the household member has heard about AIDS; (ii) indicator variable for whether the household member agree with the statement that people with AIDS should be ashamed of themselves; (iii) indicator variable for whether the household member knows that TB is spread through the air; (iv) indicator variable for whether the household's children are sleeping under a treated mosquito net. *Significantly different from zero at 90% confidence level; **Significantly different from zero at 95% confidence level; ***Significantly different from zero at 99% confidence level.

TABLE 11: PROGRAM IMPACT ON IMMUNIZATIONS – PARTICIPATION & INFORMATION MODEL (2005-2008)

Group Specification	Newborn (i)	Less than 1-year (ii)	1-year old (iii)	2-year old (iv)	3-year old (v)	4-year old (vi)
Average standardized effect	0.03 (0.03)	0.02 (0.03)	0.03 (0.04)	0.05 (0.04)	0.05 (0.04)	0.06 (0.05)
Mean control group	0.78	0.83	0.89	0.91	0.92	0.94
Observations	2535	2481	1894	1325	826	372

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 in parentheses. Groups: (i) Children under 3 months; (ii) Children 0-12 months; (iii) Children 13-24 months; (iv) Children 25-36 months; (v) Children 37-48 months; (vi) Children 49-60 months. *Significantly different from zero at 90% confidence level; **Significantly different from zero at 95% confidence level; ***Significantly different from zero at 99% confidence level.

TABLE 12: PROGRAM IMPACT ON HEALTH OUTCOMES– PARTICIPATION MODEL (2007-2008)

Dependent variable	Under-5 mortality	Infant mortality	Number of births	Number of pregnancies	Weight- for-age z-scores (0-11m)	Weight- for-age z-scores (0-11m)	Height-for- age z-scores (12-59m)	Height-for- age z-scores (12-59m)
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Program impact	-1.87 (4.22)	-3.16 (10.1)	0.00 (.04)	-0.01 (.05)	0.20 (.11)	0.14 (.10)	0.03 (.08)	0.03 (.08)
Rate ratio	0.87 (.21)	0.98 (.21)						
Child age (log)						-0.39*** (.08)		0.45** (.09)
Female						0.30** (.14)		0.11** (.05)
Mean control group	14.8	38.6	0.43	0.52	-0.65	-0.65	-1.50	-1.43
Observations	25	25	3328	3328	423	423	1452	1821

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors in parentheses (i-ii), clustered by catchment area (iii-viii). Program impact measures the coefficient on the assignment to treatment indicator. Specifications: see notes to tables 3-5 for details.

*Significant at 10% level; **Significant at 5% level; ***Significant at 1% level.

TABLE 13: PROGRAM IMPACT ON UTILIZATION/COVERAGE USING FACILITY AND HOUSEHOLD DATA– PARTICIPATION MODEL (2007-2008)

Dep. variable	Out-patients	Delivery	Antenatal	Family planning	Average std effect	Use of Project facility	Self-treatment/ traditional healer	Average std effect
<u>PANEL A: Cross section</u>	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Program impact	-100.2 (114.4)	-11.9** (5.9)	9.6 (19.2)	3.0 (4.6)	-0.27 (0.30)	0.004 (0.014)	-0.022** (0.009)	0.046 (0.028)
Observations	25	25	25	25	25	2825	2825	2825
Mean control	649.6	23.9	106.9	18.6		0.32	0.11	-
<u>PANEL B: Value added</u>	(ix)	(x)	(xi)	(xii)	(xiii)			
Program impact	-34.4 (77.7)	-5.7** (2.6)	2.9 (14.2)	4.0 (4.0)	0.01 (0.14)			
Observations	25	25	25	25	25			
<u>PANEL C: Panel data</u>	(xiv)	(xv)	(xvi)	(xvii)	(xviii)	(xix)	(xx)	(xxi)
Program impact	-20.3 (150.4)	-5.9 (7.5)	0.32 (26.3)	5.4 (6.8)	-0.01 (0.41)	0.02 (0.03)	-0.03 (0.03)	0.09 (0.09)
Observations	50	50	50	50	50	5657	5657	5657

Notes: Panel A reports program impact estimates from a cross-sectional model using data from the end line year (2009) with robust standard errors in parentheses. Panel B reports program impact estimates from a value-added specification, using data from the end line year (2009) and average utilization for outpatient services and deliveries pre-intervention (2006) as lagged variable, with robust standard errors in parentheses. Panel C reports program impact estimates from a difference-in-differences specification with robust standard errors in parentheses. Program impact measures the coefficient on the assignment to treatment indicator in panels A and B, and the assignment to treatment indicator interacted with an indicator variable for 2009 in panel C. Specifications: see notes in Table 6 for details. *Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

TABLE 14: INFORMATION PROCESSES AND LOCAL ACTIONS

Dependent variable	Participants in community meetings	Number of actions in contract	Share of upper-level actions at baseline	Share of local-level actions at baseline	Share of local-level actions in 2007
Specification	(i)	(ii)	(iii)	(iv)	(v)
Participation & information model	131.4 (36.4)	13.0 (5.3)	0.12 (0.12)	0.88 (0.12)	0.80 (0.16)
Participation model	144.5 (36.4)	12.9 (5.1)	0.50 (0.13)	0.50 (0.13)	0.50 (0.13)
Difference	-13.2 (12.4)	0.04 (1.8)	-0.39 ^{***} (0.04)	0.39 ^{***} (0.04)	0.30 ^{***} (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 treatment group, with standard deviations reported in parenthesis, and the difference between the two treatment groups with robust standard errors in parentheses. Specifications: (i) number of participants in the community meeting; (ii) number of actions agreed upon by community members and health facility staff in the joint action plan; (iii) share of upper-level actions agreed upon in the baseline action plan (2005 for the Participation & Information experiment and 2007 for the Participation experiment); (iv) share of local-level actions out of the total number of actions agreed upon in the baseline action plan (2005 for the Participation & Information experiment and 2007 for the Participation experiment); (v) Share of local-level actions in the action plans conducted in 2007 (follow-up action plan for the Participation & Information Model and baseline action plan for the Participation Model).

TABLE A.1: COMPARISON OF CONTROL GROUP CHARACTERISTICS IN 2006

Variables	Participation & information, control group	Participation, control group	Difference between control groups
Specification	(i)	(ii)	(iii)
Out-patient care	651.5 (215)	730.2 (235.1)	-78.7 (79.6)
Delivery	12.3 (12.4)	19.9 (14.1)	-7.7 (4.7)
No. households in catchment area	2224 (1021)	2519 (1144)	295 (385)
No. households per village	95.3 (31.6)	118.1 (41.7)	-22.8* (13.4)
Working water source	0.96 (0.2)	1.0 (0.0)	-0.04 (0.04)
Functioning electricity at health facility	0.44 (0.51)	0.42 (0.51)	0.02 (0.18)
<u>Average standardized effects</u>			
Utilization from health facility records			-0.49 (0.33)
Utilization pattern of the users			-0.04 (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			0.01 (0.06)
Supply of drugs			-0.56 (0.34)
User charges			-0.86*** (0.32)

Notes: Each row reports catchment area/health facility averages for the two control groups and difference in averages. Standard deviations are reported in parenthesis below the averages and robust standard errors are reported in parentheses under the differences (columns (iii)). characteristics are catchment area/health facility averages for the two control groups (Participation & Information model and Participation model) in 2006 and difference in averages, with robust standard errors in parentheses. Description of variables: see Table 1 for details. *Significant at 10%. **Significant at 5%. ***Significant at 1%.

TABLE A.2: PROGRAM IMPACT ON MONITORING AND INFORMATION – PARTICIPATION MODEL (2007-2008)

Panel A:						
Dependent variable (health facility level)	Suggestion box	Numbered waiting cards	Staff duty roster	Poster informing free services	Poster on patients' rights	Average standardized Effect
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Program impact	-0.03 (0.12)	-0.08 (0.11)	0.08 (0.11)	-0.27** (0.11)	-0.00 (0.13)	-0.12 (0.17)
Mean control group	0.17	0.17	0.08	0.25	0.17	-
Observations	25	25	25	25	25	25
Panel B:						
Dependent variable (household level)	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	(vii)	(viii)	(ix)	(x)	(xi)	
Program impact	0.02 (0.02)	0.19** (0.07)	-0.03 (0.02)	0.01 (0.02)	0.04 (0.03)	
Mean control group	0.12	0.77	0.28	0.15	0.79	
Observations	2840	1505	2202	2505	2838	

Notes: Robust standard errors in parentheses. Disturbance terms are clustered by catchment areas in columns (vii)-(xi). Point estimates, standard errors, and average standardized effect, columns (i)-(vi), are derived from equation (1). Program impact measures the coefficient on the assignment to treatment indicator. Specifications: see Table 8 for details. *Significant at 10%. **Significant at 5%. ***Significant at 1%. #denotes average standardized effect.

TABLE A.3: PROGRAM IMPACT ON TREATMENT PRACTICES AND MANAGEMENT– PARTICIPATION MODEL (2007-2008)

Variable	Equipment used	Waiting time	Absence rate	Condition of clinic	Drug stock- outs
Specification	(i)	(iii)	(v)	(vi)	(vii)
	OLS	OLS	OLS	OLS	OLS
Program impact	0.01 (0.04)	-2.83 (11.7)	-0.04 (0.04)	-0.34 (0.46)	0.04 (0.07)
Mean control group	0.38	129.03	0.55	0.23	0.54
Observations	2110	2206	25	25	19

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors clustered by catchment areas in parentheses. Program impact measures the coefficient on the assignment to treatment indicator in the OLS models. Specifications: see Table 9 for details. *Significantly different from zero at 90% confidence level; **Significantly different from zero at 95% confidence level; ***Significantly different from zero at 99% confidence level.

TABLE A.4: PROGRAM IMPACT ON ANTENATAL AND POSTNATAL CARE – PARTICIPATION MODEL (2007-2008)

Dependent variable	Examined by midwife during antenatal visit at the health facility	Weight taken during antenatal visit	Blood sample during antenatal visit	Fetus checked during antenatal visit	Told about pregnancy complications during antenatal visit	Newborn checked at the health facility in the first two months after delivery	Average Standardized Effect
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Program impact	0.01 (0.02)	-0.05 (0.06)	0.10 (0.07)	0.00 (0.01)	-0.044 (0.03)	-0.08** (0.04)	-0.05 (0.07)
Mean control	0.97	0.74	0.44	0.96	0.65	0.46	-
Observations	787	787	788	788	788	1033	664

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors clustered by catchment areas in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. The data come from household interviews with women who experienced a pregnancy in the last 2 years. Specifications: see Table 10 for details. *Significantly different from zero at 90% confidence level; **Significantly different from zero at 95% confidence level; ***Significantly different from zero at 99% confidence level.

TABLE A.5: PROGRAM IMPACT ON HEALTH EDUCATION – PARTICIPATION MODEL (2007-2008)

Dependent Variable	Have heard about AIDS	AIDS stigma	Knowledge that TB is spread through the air	Children sleeping under a treated mosquito net
Specification	(i)	(ii)	(iii)	(iv)
Program impact	-0.002* (0.001)	0.01 (0.00)	-0.01 (0.03)	-0.02 (0.03)
Mean control group	1.00	0.06	0.54	0.28
Observations	2844	2796	2761	1766

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors clustered by catchment areas in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. The data come from household interviews. Specifications: see Table 11 for details. *Significantly different from zero at 90% confidence level; **Significantly different from zero at 95% confidence level; ***Significantly different from zero at 99% confidence level.

TABLE A.6: PROGRAM IMPACT ON IMMUNIZATIONS – PARTICIPATION MODEL (2007-2008)

Group Specification	Newborn (i)	Less than 1-year (ii)	1-year old (iii)	2-year old (iv)	3-year old (v)	4-year old (vi)
Average standardized effect	-0.04 (0.08)	-0.02 (0.06)	-0.02 (0.06)	-0.05 (0.07)	0.02 (0.08)	0.11 (0.09)
Mean control group	0.79	0.83	0.88	0.90	0.90	0.92
Observations	1498	1466	1097	756	467	226

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 in parentheses. For more details see notes to Table 12. *Significantly different from zero at 90% confidence level; **Significantly different from zero at 95% confidence level; ***Significantly different from zero at 99% confidence level.

TABLE A.7: ALTERNATIVE MECHANISMS – PARTICIPATION & INFORMATION MODEL (2005-2008)

Dependent Variable	Funds received in USD	Number of staff left health facility	Supervision visits	Drugs received per patient
Specification	(i)	(ii)	(iii)	(iv)
Program impact	12.7 (11.7)	-0.004 (0.13)	0.07 (0.09)	0.20 (0.15)
Mean control group	79.4	0.20	0.32	-
Observations	50	50	50	48

Notes: Estimates from equation (1) with district fixed effects. Robust standard errors clustered by catchment areas in parentheses. Program impact measures the coefficient on the assignment to treatment indicator. Specifications: (i) average yearly amount of Primacy Health Care (PHC) funds received per health facility over the years 2006-2008 in USD; (ii) number of staff that has left the health facility during 2006-2008; (iii) if the health facility receives regular (once a month) monitoring visits from village health teams; (iv) average standardized treatment effect of the number of tablets of (a) erythromycin, (b) coartem/artemether, (c) lumefantrine, (d) cotrimoxazole/septrin, and (e) quinine, received per patient and health facility. *Significantly different from zero at 90% confidence level; **Significantly different from zero at 95% confidence level; ***Significantly different from zero at 99% confidence level.