Quarterly Journal of Economics (forthcoming)

POWER TO THE PEOPLE: EVIDENCE FROM A RANDOMIZED FIELD EXPERIMENT ON COMMUNITY-BASED MONITORING IN UGANDA*

Martina Björkman and Jakob Svensson

This paper presents a randomized field experiment on community-based monitoring of public primary health care providers in Uganda. Through two rounds of village meetings, localized NGOs encouraged communities to be more involved with the state of health service provision and strengthened their capacity to hold their local health providers to account for performance. A year after the intervention, treatment communities are more involved in monitoring the provider and the health workers appear to exert higher effort to serve the community. We document large increases in utilization and improved health outcomes – reduced child mortality and increased child weight – that compare favorably to some of the more successful community-based intervention trials reported in the medical literature.

*This project is a collaborative exercise involving many people. Foremost, we are deeply indebted to Frances Nsonzi and Ritva Reinikka for their contributions at all stages of the project. We would also like to acknowledge the important contributions of Gibwa Kajubi, Abel Ojoo, Anthony Wasswa, James Kanyesigye, Carolyn Winter, Ivo Njosa, Omiat Omongin, Mary Bitekerezo, and the field and data staff with whom we have worked over the years. We thank the Uganda Ministry of Health, Planning Division, the World Bank's Country Office in Uganda, and the Social Development Department, the World Bank, for their cooperation. We are grateful for comments and suggestions by Paul Gertler, Esther Duflo, Abhijit Banerjee, and seminar and conference participants at Stanford, Berkeley, LSE, Oxford, IGIER, MIT, World Bank, NTNU, Namur, UPF, CEPR/EUDN conference in Paris, and BREAD & CESifo conference in Venice. We also thank three anonymous referees and the editor, Lawrence Katz, for very constructive suggestions. Financial support from the Bank-Netherlands Partnership Program (BNPP), the World Bank Research Committee, the World Bank Africa Region division, and the Swedish International Development Agency, Department for Research Cooperation is gratefully acknowledged. Björkman also thanks Jan Wallander's and Tom Hedelius' Research Foundation for funding.

I. Introduction

Approximately 11 million children under five die each year and almost half of these deaths occur in sub-Saharan Africa. More than half of these children will die of diseases (e.g. diarrhea, pneumonia, malaria, measles, and neonatal disorders) that could easily have been prevented or treated if the children had had access to a small set of proven, inexpensive services (Black et al. 2003; Jones et al. 2003).

Why are these services not provided? Anecdotal, and recently more systematic, evidence points to one possible reason – ineffective systems of monitoring and weak accountability relationships.[1] This paper focuses on one of these accountability relationships, citizen-clients' ability to hold providers accountable, using primary health care provision in rural Uganda as a testing ground.

To examine whether community-based monitoring works, we designed and conducted a randomized field experiment in 50 communities from nine districts in Uganda. In the experiment, local NGOs facilitated village and staff meetings in which members of the communities discussed baseline information on the status of health service delivery relative to other providers and the government standard. Community members were also encouraged to develop a plan identifying key problems and steps the providers should take to improve health service provision. The primary objective of the intervention was to initiate a process of community-based monitoring that was then up to the community to sustain and lead.

The community-based monitoring project increased the quality and quantity of primary health care provision. A year after the first round of meetings we find a significant difference in the weight of infants – 0.14 z-score increase – and a markedly lower number of deaths among children under five – 33 percent reduction in under-five mortality – in the treatment communities. Utilization for general outpatient services was 20 percent higher in the treatment compared to the control facilities and the overall effect across a set of utilization measures is large and significantly positive. Treatment practices, including immunization of children, waiting time, examination procedures, and absenteeism, improved significantly in the treatment communities, thus suggesting that the changes in quality and quantity of health care provision are due to behavioral changes of the staff. We find evidence that the treatment communities became more engaged and began to monitor the health unit more extensively. Using variation in treatment intensity

[1] For anecdotal and case study evidence, see World Bank (2003). Chaudhury et al. (2006) provide evidence on the rates of absenteeism. On misappropriation of public funds and drugs, see Reinikka and Svensson (2004) and McPake et al. (1999).

across districts we show that there is a significant relationship between the degree of community monitoring and health utilization and health outcomes, consistent with the community-based monitoring mechanism.

Community-based, randomized, controlled field trials have been used extensively in medical research to evaluate the effectiveness of various health interventions (see footnote 14). Our paper is related but differs in one important dimension. While the medical field trials address the question of impact of a biological agent or treatment practice when the health workers competently carry out their tasks, we focus on how to ensure that the health workers actually carry out their tasks and the impact that may have on health utilization and health outcomes.

This paper also relates to a small literature on improving governance and public service delivery through community participation. Olken (2007) finds minor effects of an intervention aimed at increasing community participation in the monitoring of corruption in Indonesia. Our work differs in several ways. First, the intervention we evaluate was structure in a way to reduce the risk of elite capture. Second, unlike corruption which is not easily observable, the information discussed in the meetings were basic facts on utilization and quality of services based on the community's own experience. Finally, the intervention sought to address two constraints highlighted in the literature on community monitoring: lack of relevant information and inadequate participation. Banerjee, Deaton, and Duflo (2004) evaluate a project in Rajasthan in India where a member of the community was paid to check whether the nurse-midwife assigned to the health center was present at the center. The intervention had no impact on attendance and the authors speculate that a key reason for this is that the individual community member did not manage to use his or her information on absenteeism to invoke community participation. Here, on the contrary, we explicitly try to address the participation constraint by involving a large number of community members and encourage them to jointly develop a monitoring plan.

Finally, the paper links to a growing empirical literature on the relationship between information dissemination and accountability (Strömberg 2004; Besley and Burgess 2002; Ferraz and Finan 2008). In this paper, however, we focus on mechanisms through which citizens can make providers, rather than politicians, accountable. Thus, we do not study the design or allocation of public resources across communities, but rather how these resources are utilized. Second, we use micro data from households and clinics rather than disaggregated national accounts data. Finally, we identify impact using an experimental design.

The next section describes the institutional environment. The community-based monitoring intervention is described in section 3. Section 4 lays out the evaluation design and the results are presented in section 5. Section 6 concludes. Details about the experiment and additional results are reported in a supplemental appendix.

II. Institutional Setting

Uganda, like many newly independent countries in Africa, had a functioning health care system in the early 1960s. The 1970s and 1980s saw the collapse of Government services as the country underwent political upheaval. Health indicators fell dramatically during this period until peace was restored in the late 1980s. Since then, the Government has been implementing major infrastructure rehabilitation programs in the public health sector.

The health sector in Uganda is composed of four types of facilities: hospitals, health centers, dispensaries, and aid posts or sub-dispensaries. These facilities can be government, private for-profit, or private not-for-profit operated and owned. The impact evaluation focuses on public dispensaries. Dispensaries are in the lowest tier of the health system where a professional interaction between users and providers takes place. Most dispensaries are rural. 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). As of 2001, public health services are free of charge. In our sample, on average, a dispensary was staffed by an in-charge or clinical officer (a trained medical worker), two nurses, and three nursing aids or other assistants.

The health sector in Uganda is decentralized and a number of actors 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. Each dispensary has an HUMC which consists of both health workers and non-political representatives from the community. The HUMC should monitor the day-to-day running of the facility but it has no authority to sanction workers. The next level in the institutional hierarchy is the Health Sub-district. The Health Sub-district monitors funds, drugs and service delivery at the dispensary. Supervision meetings by the Health Sub-district are supposed to appear quarterly but, in practice, monitoring is infrequent. The Health Sub-district has the authority to reprimand, but not dismiss, staff for indiscipline. Thus in severe cases of indiscipline, the errand will be referred to the

Chief Administrative Officer of the District and the District Service Commission, which are the appointing authorities for the district. They have the authority to suspend or dismiss staff.

Various local NGOs, so-called Community-based organizations (CBOs), focusing primarily on health education, are also active in the sector.

III. Experimental Design and Data

III.A. Overview

In response to perceived weak health care delivery at the primary level, a pilot project (Citizen report cards) aimed at enhancing community involvement and monitoring in the delivery of primary health care was initiated in 2004. The project was designed by staff from Stockholm University and the World Bank, and implemented in cooperation with a number of Ugandan practitioners and 18 community-based organizations.

The main objective of the intervention was to strengthen providers' accountability to citizen-clients by initiating a process, using trained local actors (CBOs) as facilitators, which the communities themselves could manage and sustain.

Based on a small but rigorous empirical literature on community participation and oversight, and extensive piloting in the field, our conjecture was that lack of relevant information on the status of service delivery and the community's entitlements, and failure to agree on, or coordinate expectations of, what is reasonable to demand from the provider, were holding back initiatives to pressure and monitor the provider. While individual community members have private information, for example they know whether their own child died or not, and whether the health workers did anything to help them, they typically do not have any information on aggregate outcomes, such as how many children in their community did not survive beyond the age of 5 or where citizens, on average, seek care, or what the community can expect in terms of quality and quantity of service provision (Khemani 2006). Partly as a response to this information problem, and partly because monitoring a public facility is a public good which may be subject to serious free-rider problems, few people actively participate in monitoring their service providers. Relaxing these two constraints was therefore the main objective of the intervention.

The key behavioral change induced by more extensive community-based monitoring was expected to be increased effort by the health unit staff to serve the community. In Uganda, as in many other developing countries, health workers have little pecuniary incentives to exert high effort. Public money does not follow patients and hiring, salaries, and promotions are largely determined by seniority and educational qualifications – not by how well the staff performs. An individual worker may of course still put in high effort if shirking deviates from her ideal choice (Akerlof and Kranton 2005). The effort choice may also be influenced by social rewards from community members or social sanctions against shirking workers. Social rewards and sanctions are key instruments available to the community to boost health worker's effort.

III.B. Experimental Design

The experiment involved 50 public dispensaries, and health care users in the corresponding catchment areas, in nine districts covering all four regions in Uganda. All project facilities were located in rural areas. We define a facility's catchment area, or the community, as the five-kilometer radius around the facility.[2] A community in our sample has, on average, 2,500 household residing within the 5-kilometer radius of the clinic, of which 350 live within a one-kilometer radius. For the experimental design, the facilities were first stratified by location (districts) and then by population size. From each group, half of the units were randomly assigned to the treatment group and the remaining 25 units were assigned to the control group.

III.C. Data

Data collection was governed by two objectives. First, data were required to assess how the community at large views the quality and efficacy of service delivery. We also wanted to contrast the citizens' view with that of the health workers. Second, data were required to evaluate impact. To meet these objectives, two surveys were implemented: a survey of the 50 providers and a survey of users. Both surveys were implemented prior to the intervention (data from these surveys formed the basis for the intervention) and one year after the project had been initiated.

A quantitative service delivery survey was used to collect data from the providers. Since agents in the service delivery system 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.

[2] Dispensaries are designed to serve households in a catchment area roughly corresponding to the five-kilometer radius around the facility (Republic of Uganda 2000).

The household survey collected data on both households' health outcomes and health facility performance as experienced by the household. A stratified random sample of households within the catchment area of each facility was surveyed. In total roughly 5,000 households were surveyed in each round.[3] To the extent that it was possible, patient records, i.e., patient exercise books and immunization cards, supported the household's response. The post-intervention household survey also included a shorter module on health outcomes. Specifically, data on under-five mortality were collected and we measured the weight of all infants in the surveyed households.

III.D. Intervention

A smaller subset of the findings from the pre-intervention surveys, including utilization, quality of services, and comparisons vis-à-vis other health facilities, were assembled in report cards. Each treatment facility and its community had a unique report card, translated into the main language spoken in the community, summarizing the key findings from the surveys conducted in their area.

The process of disseminating the report card information, and encouraging participation, was initiated through a series of meetings: a community meeting; a staff meeting; and an interface meeting. Staff from various local NGOs (CBOs) acted as facilitators in these meetings.[4] A time-line of the intervention is depicted in Figure I.

The community meeting was a two-afternoons event with approximately 100 invited participants from the community. To avoid elite capture, the invited participants consisted of a selection of representatives from different spectra of society (i.e. young, old, disabled, women, mothers, leaders). The facilitators mobilized the village members by cooperating with village council representatives in the

[3] The sample strategy for the baseline household survey was designed to generate representative information on the core users' variables in each community (such as the proportion of patients being examined with equipment). In total, 88 percent of the households surveyed in the baseline survey were resurveyed in the ex-post survey. The households that could not be surveyed were replaced.

[4] The participating CBOs, 18 in total, had been active in 64% or the treatment communities and half of the control communities prior to the intervention. A handful of them covered more than one treatment community. The CBOs were primarily focused on health, including issues of health education and HIV/Aids prevention, although other objectives such as agricultural development, women empowerment, support of orphans and vulnerable children, and peace building initiatives, were also common. The CBO facilitators were trained for seven days in data interpretation and dissemination, utilization of the participatory methodology, and conflict resolution and management. Various other CBOs also operate in the project communities.

catchment area. Invited participants were asked to spread the word about the meeting and, in the end, a large number of uninvited participants also attended the meeting. More than 150 participants per day attended a typical village meeting.

In the community meeting, the facilitators used a variety of participatory methods to disseminate the information in the report cards and encouraged community members to develop a shared view on how to improve service delivery and monitor the provider. Information on patient's rights and entitlements was also discussed. The participants were divided into focus groups so that also more marginalized groups such a women and youth could raise their voices and discuss issues specific to their group. At the end of the meeting, the community's suggestions for improvements, and how to reach them without additional resources, were summarized in an action plan. The action plan contained information on health issues/services that had been identified by the community as the most important to address; how these issues could be addressed and how the community could monitor improvements (or lack thereof). While the issues raised in the action plans differed across communities, a common set of concerns included high rates of absenteeism, long waiting-time, weak attention of health staff, and differential treatment.

The health facility meeting was a one-afternoon event held at the facility with all staff present. In the meeting, the facilitators contrasted the information on service provision as reported by the provider with the findings from the household survey.

An interface meeting with members from the community, chosen in the community meeting, and health workers followed the community and health facility meetings. During the interface meeting, the community representatives and the health workers discussed suggestions for improvements. The participants discussed their rights and responsibilities as patients or medical staff. The outcome was a shared action plan, or a contract, outlining the community's and the service provider's agreement on what needs to be done, how, when and by whom. The "community contract" also identified how the community could monitor the agreements and a time plan. Because the problems that were raised in the community meetings constituted the core issues discussed during the interface meetings, the community contract was in many respects similar to the community's action plan.

The three separate meetings aimed at kick-starting the process of community monitoring. Thus, after the initial meetings the communities were themselves in-charge of establishing ways of monitoring the provider. After a period of six months, the communities and health facilities were revisited. The CBOs facilitated a one-afternoon community meeting and a one-afternoon interface meeting with the aim of tracking the implementation of the community contract. Health facility staff and community members jointly discussed suggestions for sustaining or improving progress, or in the case of no improvements, why so.[5]

IV. Evaluation Design and Expected Outcomes

IV.A. Outcomes

The main outcome of interest is whether the intervention increased the quantity and quality of health care provision and hence resulted in improved health outcomes. We are also interested in evaluating changes in all steps in the accountability chain: Did the treatment communities become more involved in monitoring the health workers? Did the intervention change the health worker's behavior?

As a robustness test we also assess alternative explanations. One concern is spillovers. Another concern is that the intervention did not only (or primarily) increase the extent of community monitoring, but had an impact on other agents in the service delivery chain, such as the Health sub district. The intervention could also have affected the health workers' behavior directly, or affected it through the actions of the CBOs, rather than through more intense community-based monitoring as we hypothesize. While this would not invalidate the causal effect of the intervention it would, of course, affect the interpretation. Therefore, these alternative hypotheses are also subject to a battery of tests.

IV.B. Statistical Framework

To assess the causal effect of the intervention we estimate,

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

where y_{ijd} is the outcome of household i (when applicable), in community/health facility j, in district d, T_{jd} is an indicator variable for assignment to treatment, and ε_{ijd} is an error term. Equation (1) also includes a vector, X, of pre-intervention facility-specific covariates and district fixed effects (θ_d).[6] Due to random assignment, T should be orthogonal to X, and the consistency of β does not depend on

[5] Details on the report cards and the participatory methods used, as well as an example of an action plan, are provided in the supplemental appendix.

[6] The baseline covariates included are number of villages in catchment area, number of days without electricity in the last month, indicator variable for whether the facility has a separate maternity unit, distance to nearest public health provider, number of staff with less than advanced A-level education, indicator variable for whether the staff could safely drink from the water source, and average monthly supply of Quinine.

the inclusion of X in the model. The regression adjustment is used to improve estimation precision and to account for stratification and chance differences between groups in the distribution of pre-random assignment (Kling, Liebman, and Katz 2007).

We report the results of estimating equation (1) with X and θ excluded in a supplemental appendix.

For a subset of variables we can also stack the pre and post data and explore the difference-in-differences in outcomes; i.e., we estimate[7]

$$y_{ijt} = \gamma POST_t + \beta_{DD}(T_j * POST_t) + \mu_j + \varepsilon_{ijt}, \tag{2}$$

where $POST$ is an indicator variable for the post-intervention period, μ_j is a facility/community specific fixed effect, and β_{DD} is the difference-in-differences estimate (program impact).

For some outcomes we have several outcome measures. To form judgment about the impact of the intervention on a family of K related outcomes, we follow Kling et al. (2004) and estimate a seemingly unrelated regression system,

$$Y = [I_K \otimes (T\ X)]\,\theta + \upsilon\ , \tag{3}$$

where I_K is a K by K identity matrix. We then derive average standardized treatment effects, $\tilde{\beta} = \frac{1}{K}\sum_{k=1}^{K}\frac{\hat{\beta}_k}{\hat{\sigma}_k}$, where $\hat{\beta}_k$ and $\hat{\sigma}_k$ are the point estimate and standard error, respectively, for each effect (see Duflo et al. 2007). The point estimate, standard error, and p-value for $\tilde{\beta}$ are based on the parameters, $\hat{\beta}_k$ and $\hat{\sigma}_k$, jointly estimated as elements of θ in (3).

V. Results

V.A. Pre-intervention Differences

The treatment and the control group were similar on most characteristics prior to the intervention. Average standardized pre-treatment effects are estimated for each family of outcomes (utilization, utilization pattern, quality, catchment area statistics, health facility characteristics, citizen perceptions, supply of resources, and user charges) using pre-intervention data. As shown in Table I, we cannot reject the null hypotheses of no difference between the treatment and the control group.[8]

[7] It is a subset of variables since the post intervention surveys collected information on more variables and outcomes.

[8] We report the test of difference in means across control and treatment groups for each individual variable in the supplemental appendix.

V.B. Processes

The initial phase of the project; i.e., the three separate meetings, followed a pre-design structure. A parallel system whereby a member of the survey team originating from the district participated as part of the CBO team also confirmed that the initial phase of the intervention was properly implemented. After these initial meetings it was up to the community to sustain and lead the process. In this section we study whether the treatment communities became more involved in monitoring the providers.

To avoid influencing local initiatives we did not have external agents visiting the communities and could therefore not document all actions taken by the communities in response to the intervention. Still we have some information on how processes in the community have changed. Specifically, the CBOs submitted reports on what type of changes they observed in the treatment communities and we also surveyed the local councils in the treatment communities. We use facility and household survey data to corroborate these reports.

According to the CBO-reports, and the local council survey, the community-based monitoring process that followed the first set of meetings was a joint effort mainly managed by the local councils, HUMC, and community members. A typical village in the treatment group had, on average, six local council meetings in 2005. In those meetings, 89 percent of the villages discussed issues concerning the project health facility. The main subject of discussion in the villages concerned the community contract or parts of it, such as behavior of the staff.

The CBOs reported that concerns raised by the village members were carried forward by the local council to the facility or the HUMC. However, although the HUMC is an entity that should play an important role in monitoring the provider, it was in many cases viewed as being ineffective. As a result, mismanaged HUMCs were dissolved and new members elected. These claims are confirmed in the survey data: More than one third of the HUMCs in the treatment communities were dissolved and new elected or received new members following the intervention, while we observe no dissolved HUMC in the control communities. Further, the CBOs report that the community, or individual members, also monitored the health workers during visits to the clinic, when they rewarded and questioned issues in the community contract that had or had not been addressed, suggesting a more systematic use of non-pecuniary rewards. Monitoring tools such as suggestion boxes, numbered waiting cards, and duty roasters, were also reported to be put in place in several treatment facilities.

In Table II we formally look at the program impact on these monitoring tools.

We use data collected through visual checks by enumerators during the post-intervention facility survey. As shown in columns (i)-(ii), one year into the project treatment facilities are significantly more likely to have suggestion boxes (no control facility had these, while 36 % of the treatment facilities did) and numbered waiting cards (only one control facility had one, while 20 % of the treatment facilities did). Columns (iii)-(iv) show that a higher share of the treatment facilities also post information on free-services and patient's rights and obligations. The enumerators could visually confirm that 70 percent of the treatment facilities had at least one of these monitoring tools, while only 4 out of 25 control clinics had at least one of them. The difference is statistically significant (supplemental appendix, Table A.II). The fifth column reports the average standardized effect of the monitoring tools. The estimate is significantly different from zero at the 1-percent level.

The results based on household data mirror the findings reported in columns (i)-(v). The performance of the staff is more often discussed in local council meetings in the treatment communities, shown in column (vi), and community members in the treatment group are on average better informed about the HUMC's roles and responsibilities, as reported in column (vii). Combining the evidence from the CBO reports and the household survey data thus suggests that both the "quantity" of discussions about the project facility and the subject, from general to specific discussions about the community contract, changed in response to the intervention.

V.C. Treatment Practices

The qualitative evidence from the CBOs and, to the extent that we can measure it, the findings reported in Table II, suggest that the treatment communities became more involved in monitoring the provider. Did the intervention also affect the health worker's behavior and performance? We turn to this next.

We start by looking at examination procedures. The estimate based on equation (2) with the dependent variable being an indicator variable for whether any equipment, for instance a thermometer, was used during examination is shown in the first row in Table III. 50 [41] percent of the patients in the treatment [control] community reported that equipment was used the last time the respondent (or the respondent's child) visited the project clinic. The difference-in-differences estimate, a 20% increase, is highly significant. The cross-section estimate in row (ii), based on equation (1), is less precisely estimated.

In row (iii) we report the result with an alternative measure of staff performance – the waiting time – defined as the difference between the time the user left the facility and the time the user arrived at the facility, subtracting the examination time. On average, the waiting time was 131 minutes in the control facilities and 119 in the treatment facilities. The estimate based on equation (1), shown in column (iv), is less precisely estimated.

The results on absenteeism is shown in the third row.[9] The point estimate suggests a substantial treatment effect. On average, the absence rate, defined as the ratio of workers not physically present at the time of the post-intervention survey to the number of workers on the list of employees as reported in the pre-intervention survey, is 13 percentage points lower in the treatment facilities. Thus, in response to the intervention health workers are more likely to be at work.

Enumerators also 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. We combine these variables through principal components analysis into a summary score. Treatment clinics appear to have put more effort into keeping the clinic in decent condition in response to the intervention. The point estimate, reported in row (vi), implies a 0.56 standard deviations improvement in the summary score in the treatment compared to the control facilities.

According to the government health sector strategic plan preventive care is one of the core tasks for health providers at the primary level. A significantly larger share of households in the treatment communities have received information about the dangers of self-treatment, reported in row (vii), and the importance of family planning, reported in row (viii). The difference is 7 and 6 percentage points, respectively.

There is no systematic difference in the supply of drugs between the treatment and control groups (see section V.F). However, as shown in row (ix), stock-outs of drugs are occurring at a higher frequency in the control facilities even though, as reported below, the control facilities treat significantly fewer patients. These findings suggest that more drugs leaked from health facilities in the control group.[10]

[9] The post-intervention survey was not announced in advance. At the start of the survey the enumerators physically verified the provider's presence. A worker was counted as absent if, at the time of the visit, he or she was not in the clinic. Staff reported to be on outreach were omitted from the absence calculation. Four observations were dropped because the total number of workers verified to be present or reported to be on outreach exceeded the total number of workers on the pre-intervention staff list. Assuming instead no absenteeism in these four facilities yields a point estimate (standard error) of -0.20 (0.065).

[10] The dependent variable is the share of months in 2005 in which stock-cards indicated no availability of drugs, averaged over Erythromycin, Mebendazole, and Septrin. We find no significant difference between treatment and control clinics for Chloroquine – the least expensive

The findings on immunization of children under five are reported in Table IV. We have information on how many times (doses) in total each child has been immunized with polio, DPT, BCG, A-Vitamin supplements and measles. Based on the recommended immunization plan we create indicator variables taking the value of one if child i of cohort (age) j had received the required dose(s) of measles, DPT, BCG, and polio, respectively, and zero otherwise.[11] We then estimate (3), for each age group, and calculate average standardized effects.

The average standardized effects are significantly positive for the younger cohorts. Looking at individual effects (supplemental appendix Table A.IV), there are significant positive differences between households in the treatment and control community for all five vaccines, although not for all cohorts. For example, twice as many newborns in the treatment group have received Vitamin A supplement, 46% more newborns have received the first dose of BCG vaccine, and 42% more newborns have received the first dose of polio vaccine as compared to the control group.

V.D. Utilization

To the extent we can measure it, the evidence presented so far suggests that treatment communities began to monitor the health unit more extensively in response to the intervention and that the health workers improved the provision of health services. We now turn to the question of whether the intervention also resulted in improved quantity and quality of care.

Cross-section estimates based on equation (3) are given in Table V, Panel A. For out-patients and deliveries, we have pre-intervention data and can also estimate difference-in-differences models, shown in Panel B, and value-added models, shown in Table A.V in the supplemental appendix.[12]

---

of the drugs we have data on. Not all clinics had accurate stock-cards and these clinics were therefore omitted.

[11] According to the Uganda National Expanded Program on Immunization each child in Uganda is suppose to be immunized against measles (one dose at 9 months and two doses in case of an epidemic); DPT (three doses at 6 weeks, 10 weeks and 14 weeks); BCG (one dose at birth or during the first contact with a health facility); and polio (three doses, or four if delivery takes place at the facility, at 6 weeks, 10 weeks and 14 weeks). Because measles vaccination should not be given at birth, we exclude immunization against measles in the plan for infants under 12 months.

[12] The value added specification is

$$y_{jt} = \alpha_{VA} + \beta_{VA} T_j + \lambda y_{jt-1} + \varepsilon_{jt} \ .$$

One year into the program, utilization (for general outpatient services) is 20 percent higher in the treatment facilities as shown in specification (i). For the difference-in-differences and the value-added models (reported in specification (ix) in Table V and specification (ix) in Table A.V), the coefficients on the treatment indicator are larger both in absolute magnitude and relative to their standard errors. Thus, controlling for baseline outcomes y_{jt-1}, improves the precision of the treatment effect, which is to be expected given the persistent nature of the outcome variable. The difference in the number of deliveries, shown in specification (ii), albeit starting from a low level, is 58 percent and is fairly precisely estimated. There are also positive differences in the number of patients seeking antenatal care (19 percent increase) and family planning (22 percent increase), although these estimates are not individually significantly different from zero. The average standardized effect, reported in specification (v), however, is highly significant.

The last three columns in Table V, Panel A and B, report changes in utilization patterns based on household data. We collected data on where each household member sought care during 2005 in case of illness that required treatment and collapse this information by community. There is a 11-13 percent increase, specifications (vi) and (xii), in the use of the project facility in the treatment as compared to the control group – a result consistent with that reported in specification (i) using facility records.

Households in the treatment community also reduced the number of visits to traditional healers and the extent of self-treatment, specifications (vii) and (xiii), while there are no statistically significant differences across the two groups in the use of other providers (not reported). Thus, as summarized in the average standardized treatment effects, specification (viii) and (xiv), households in the treatment communities switched from traditional healers and self-treatment to the project facility in response to the intervention.

V.E. Health Outcomes

We collected data on births, pregnancies, and deaths of children under five years in 2005. We also measured the weight of all infants (i.e., under 18 months of age) and children (between 18 and 36 months of age) in the surveyed households.

Health outcomes could have improved for several reasons. As noted in the Introduction, access to a small set of proven, inexpensive services could, worldwide, have prevented more than half of all under-five deaths. For a country with an epidemiological profile as in Uganda, the estimate of preventable deaths is 73% (Jones et al. 2003).[13] In the community monitoring project specifically, increased utilization and having patients switching from self-treatment and traditional healers to seeking care at the treatment facility could have an effect. Holding utilization constant, better service quality, increased immunization, and more extensive use of preventive care could also have resulted in improved health status.

As a reference point we review the set of health intervention feasible for delivery at high coverage in low-income settings with sufficient evidence of effect on reducing mortality from the major causes of under-five deaths (Jones et al. 2003). We focus on community-based, randomized, controlled field trials that bear some resemblance (because they are community-based) to our project. Several of these field trials document reductions in under-five mortality rates of 30-50% one to two years into the project.[14] There is, however, a fundamental difference between the

[13] This is likely to be a conservative number since only medical interventions for which cause-specific evidence of effect was available were included in the estimation. For example, increased birth spacing, which has been estimated to reduce under-5 mortality by 19% in India, was not considered. Several perinatal and neonatal health interventions that could be implemented in low-income countries were not included either (Darmstadt et al. 2005).

[14] For example, a project in Tigray, Ethiopia, in which coordinators, supported by a team of supervisors, were trained to teach mothers to recognize symptoms of malaria in their children and provide antimalarials, reduced under-5 mortality by 40% (Kidane and Morrow 2000). Bang et al. (1990) document a 30% reduction in under-five mortality from an intervention that included mass education about childhood pneumonia and case management of pneumonia by trained village health workers — a result similar to the meta-analysis estimate of Sazawal and Black (2003). Bang et al. (1999) evaluate a project in which trained village health workers, assisted by birth attendants and supervisory visits, provided home-based neonatal care, including treatment of sepsis. Two years into the project they document a reduction in infant mortality by nearly 50%. Rahmathullah et al. (2003) assess the impact of a community-based project in two rural districts of Tamil Nadu, India, where newborn infants in the treatment group were allocated oral vitamin A after delivery. The intervention resulted in a 22% reduction in total mortality at age 6 months. Manandhar et al. (2004) evaluate a project in which a facilitator convened nine women's group meeting every month in the Makwanpur district in Nepal in which perinatal problems were identified and strategies to address them formulated. Two year into the project they document a 30% reduction in neonatal mortality. Rahman et al. (1982) evaluate the impact of immunization of women with tetanus injections during pregnancy in rural Bangladesh. The intervention reduced neonatal mortality by 45%. Mtango and Neuvians (1986) evaluate a project in rural Tanzania in which trained village health workers visited families at their homes every six to eight weeks, giving health education on recognition and prevention of acute respiratory infections, treating children with pneumonia with antibiotics or referring them to the next higher level of care. Within a two-year period, they document a 27% reduction in under-five mortality – a reduction slightly lower than that found in a similar study in rural Bangladesh (Fauveau et al. 1992).

interventions discussed in footnote (14) and our work. The medical field trials study the impact of a biological agent or treatment practice in a community setting when the community health workers and/or medical personnel competently carry out their tasks. In the experiment we consider, on the contrary, no new health interventions were introduced and the supply of health inputs were unchanged. Instead we focused on incentivizing health workers to carry out their tasks through strengthened local accountability.

Estimates for births and pregnancies are given in Table VI, columns (i)-(ii). To the extent that the intervention had an effect on fertility, for example through increased use of family planning services, it would primarily affect the incidence of pregnancies in 2005, given the 40-week period between conception to birth. The incidence of births is not significantly different across treatment and control groups. However, the treatment groups have 10 percent fewer incidences of pregnancies in 2005.

The third column shows the treatment effect on under-five mortality.[15] The point estimate suggests a substantial treatment effect. The average under-five mortality rate in the control group is 144, close to the official figure of 133 for 2005 (UNICEF 2006). In the treatment group, the under-five mortality rate is 97, that is a 33 percent reduction in under-five mortality. The difference is significant (and somewhat larger in absolute magnitude) when controlling for district fixed effects as reported in column (iii). While the effect is large, it is worth emphasizing that the 90 percent confidence interval of our estimate also includes much lower effects (90% CI: 8%-64% reduction in under-five mortality rate). With a total of approximately 55,000 households residing in the treatment communities, the treatment effect corresponds to approximately 550 averted under-five deaths in the treatment group in 2005.

Column (iv) shows the age range of the mortality effects. We have information of the birth year of all children (under-five) alive at in the beginning of 2005 and the birth year of all deceased children in 2005. Using this data we estimate (1), replacing the treatment indicator with a full set of year-of-birth indicators and year-of-birth-by-treatment interactions. We can then address the question: Conditional on having a child of age x in the end of 2004, or a child born in 2005, what is the probability that the child died in 2005? As evident, children less than two years old drive the reduction in under-five mortality. The point estimate for the youngest cohort, for example, implies a 35% reduction in the likelihood of

[15] The under-five mortality rate is the sum of the death rates for each cohort (0-1s, 1-2s, 2-3s, 3-4s, and 4-5s) per community in 2005, expressed per 1,000 live births.

death of a child born in 2005 in the treatment compared to the control group.

The program impact on the weight of infants is reported in columns (v)-(vi). Based on weight-for-age z-scores, Ugandan infants have values of weight far lower than the NCHS/CDC international reference and the gap increases for older infants, consistent with the findings in Cortinovis et al. (1997).[16] The difference in means of z scores of infants between the treatment and the control group is reported in column (v). The estimated effect (difference) is 0.14 z score in weight-for-age. Figure II plots the distribution of z scores for the treatment and the control groups. The difference in measured weight is most apparent for underweight children. This is consistent with a positive treatment effect arising from improved access and quality of health care, rather than a general increase in nutritional status, since underweight status causes a decrease in immune and non-immune host defenses and as a consequence underweight children are at a higher risk of suffering from infectious diseases, or severe complications of infectious diseases, and therefore in higher demand of health care. In column (vi), we add controls for age and gender. The results remain qualitatively unchanged.

The treatment effect is quantitatively important. For this purpose, the baseline proportion of infants in each risk category (severe, < -3 z scores; moderately, $-3 \leq z$ scores < -2; mild, $-2 \leq z$ scores < -1) in the control group was calculated. Applying the shift in the weight-for-age distribution (adding 0.14 z score) with the odds ratio for each category – children who are mildly [moderately] {severely} underweight have about a two-fold [five-fold] {eight-fold} higher risk of deaths from infectious disease (Jones et al. 2003) – the reduction in average risk of mortality is estimated to be approximately 7 percent.[17]

V.F. Getting Inside the Box and Robustness Tests

The findings of large treatment effects on our proxies of community-based

[16] The z-score is a normally distributed measure of growth defined as the difference between the weight of an individual and the median value of weight for the reference population (2000 CDC Growth Reference in the U.S.) for the same age, divided by the standard deviation of the reference population. We exclude z scores $> |4.5|$ as implausible and 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. The coefficient estimate (standard error) on the treatment indicator is 0.16 (0.09) when including these outliers.

[17] To put this into perspective, a review of controlled trials designed to improve the intake of complementary food for children aged six months to five years showed a mean increase of 0.35 z score (Jones et al. 2003). Jones et al. (2003) argue that this is one of the most effective preventive interventions feasible for delivery at high coverage in a low-income setting.

monitoring and outcomes are consistent with the community-based monitoring mechanism. But the findings do not rule out other explanations. In this section we assess a number of these alternative hypotheses.

To examine the plausibility of community-based monitoring as a key mechanism for the health utilization and health outcomes treatment effects, we follow the methodology used by Kling, Liebman and Katz (2007). Specifically, we test whether the differences in treatment-control in outcomes across districts are larger in districts with large treatment-control differences in monitoring and information outcomes. This relationship is summarized by the parameter δ, the coefficient on the summary index of monitoring and information, in the outcome equation

$$y_j = \delta M_j + X_j \pi + \varepsilon_j \ . \qquad (4)$$

The summary index of monitoring M in (4) is the first component from a principal components analysis of the six monitoring and information variables in Table II. We examine two outcome measures (y_j), under-five mortality and number of out-patients.

Following Kling, Liebman, and Katz (2007), we estimate (4) by 2SLS, using a full set of district-by-treatment interactions as the excluded instruments for the monitoring index M, while controlling for district fixed effects. The IV estimation of (4) will be consistent if M is the mediating factor between treatment and outcomes.

The IV approach is depicted graphically in Figure III.[18] There is a consistent pattern across districts and groups that larger differences in monitoring (relative to the district mean) are associated with larger differences in outcomes – a result in line with the community-based monitoring mechanism.

Estimates based on equation (4) are given in Table VII. The first two columns show 2SLS estimates of δ with district-by-treatment interactions as excluded instruments for the the monitoring index M. To increase precision, we control for baseline outcomes y_{jt-1}, when data allow it (i.e. for number of out-patients treated). The estimates are large in absolute terms and precisely estimated.

A stricter test of whether the extent of the program impact varies with the size of the community monitoring impact is to add a treatment dummy (an over-

[18] If X contains only district indicators, the 2SLS estimate of δ using the district-by-treatment interactions instruments is the slope of the line fit through a scatterplot of the outcome and monitoring index means for the treatment and control groups in each of the nine districts, normalized so that each district has mean 0 (Kling, Liebman, and Katz 2007). We plot the average values by group (treatment and control) for each district for y and M expressed in standard deviation units relative to the control group overall standard deviation for each variable.

all treatment effect regardless of the community monitoring impact) to the IV regressions in equation (4). The community monitoring index is then identified by cross-district variation in changes in community monitoring by treatment from the district-by-treatment interactions as the excluded instruments, with the main effect for treatment no longer excluded; the results are reported in columns (iii) and (iv) of Table VII. Comparing the results without and with controls for treatment are quite similar for both out-patients and under-five mortality, while the coefficients on the treatment indicator have the wrong sign and are small relative to their standard errors, providing some evidence that community monitoring had the primary effects on outcomes as opposed to other effects induced by the intervention.

To examine the hypothesis that differences in monitoring are driving the results as opposed to the supply driven hypothesis that health workers, once being informed that their effort deviates from what is expected (in the health facility staff meeting), decided to exert higher effort into serving the community, we augment specification (4) with a measure of the staff's knowledge about patient's rights and obligations.[19] This model thus have two endogenous variables. If large treatment effects on outcomes across districts are associated with differences in staff knowledge about patients' rights rather than more intense community monitoring, this would be evidence against the community-based monitoring hypothesis. As reported in columns (v)-(vi), the coefficients on community monitoring remain largely unaffected, while the coefficients on staff knowledge are insignificant and with the wrong signs, providing additional evidence, albeit not conclusive, that the demand driven mechanism is more important than the supply driven mechanism.

The CBOs played an integral role in the intervention as facilitators of the meetings. However, it is possible that these CBOs had a role (as educators or activist, for example) beyond the described treatment itself. There is no definitive way to sort out the role of community-based monitoring from the possible roles of the CBOs, but since around 60 percent of the CBOs that took part in the intervention had been operating in the communities before the intervention, and several of them also had activities in the control areas, we can investigate whether

[19] The in-charge was asked to list the patient's rights and obligations according to the Ministry of Health's plan for basic health service delivery. Patient's rights were discussed in the interface meeting. Each correct answer (out of five) was given a score of 0.2, so this test score ranges from 0-1. We also examined other measures of staff engagement, including number of staff meetings in 2005 and if the in-charge had initiated training of staff on proper conduct. The results using these alternative proxies mirror those reported in Table VII.

the outcomes are correlated with pre-intervention CBO activity. This would be the case if the CBOs that participated in the experiment, and that had been present in the communities prior to intervention, had a direct impact on health outcomes (through various preventive activities for example) or indirectly by being more involved in monitoring the provider. The number of out-patients treated per month, shown in column (vii), and the under-five mortality rate, shown in column (viii), are not significantly different in communities where the CBOs had been active prior to the intervention. We have also examined whether the treatment effect vary conditional on observable CBO characteristics/actions. For example, CBOs that are located (have an office) in the community might, everything else equal, be in a better position to monitor the health provider. Moreover, in ten of the treatment sites, the CBOs reported that it regularly visited the clinic. If the CBOs, rather than the community, were pushing the service providers into action, presumably, the effect would be more pronounced in sites where the CBO actually visited the clinic regularly. However, the treatment effects are independent of whether the office of the CBO is located within a five kilometer radius of the health facility or if the CBO reported that it regularly visited the clinic.[20]

Given that within each district there are both treatment and control units, one concern with the evaluation design is the possibility of spillovers from one catchment area to another. In practice, there are reasons to believe spillovers will not be a serious concern. The average (and median) distance between the treatment and control facility is 30 kilometers and in a rural setting it is unclear to what extent information about improvements in treatment facilities has spread to control communities. Still, the possibility of spillovers is a concern. Following Miguel and Kremer (2004), and taking advantage of the variation in distance to nearest treatment clinic induced by randomization, we estimate spillovers from treatment to control groups by enriching X in equation (1) to include an indicator variable for if the control clinic is within 10 kilometers of the nearest treatment clinic. The results are presented in the supplemental appendix (for utilization,

[20] Given the small sample size, we test whether the distribution of outcomes in the subsample $\{T = 1 \; \& \; \textit{CBO located in community} = 1\}$ is the same as in the subsample $\{T = 1 \; \& \; \textit{CBO located in community} = 0\}$, and whether the distribution of outcomes in the subsample $\{T = 1 \; \& \; \textit{CBO regularly carries out monitoring visits to the facility} = 1\}$ is the same as in the subsample $\{T = 1 \; \& \; \textit{CBO regularly carries out monitoring visits to the facility} = 0\}$, using the Wilcoxon rank-sum test. The test statistics (with p-values in parentheses) are 0.88 (0.38) and -1.10 (0.27) for outpatients and 0.31 (0.76) and -0.03 (0.98) for under-5 mortality rate. We get similar results if we enrich equation (1) with an interaction term $T \times \textit{CBO characteristic}$. The estimates of the interaction term are not statistically different from zero in any of the specifications.

delivery, and child death). We do not find evidence in favor of the spillover-hypothesis.

Another concern is if the district or sub-district management changed their behavior or support in response to the intervention. For example, the Health Sub-district or local government may have provided additional funding or other support to the treatment facilities. The results in Table A.VIII in the supplemental appendix do not provide any evidence of this being the case. The treatment facilities did not receive more drugs or funding from the sub-district or district as compared to the control facilities during 2005.

Upper-level authorities could also have increased their supervision of treatment facilities in response to the intervention. As shown in Table A.IX, however, supervision of providers by upper-level government authorities remained low in both the treatment and the control group. As a complement we also assessed sanctions. There is only a handful of staff that have been dismissed or transferred in 2005 and there is no systematic pattern that distinguishes treatment from control facilities. There is also no difference between treatment and control facilities in the number of staff that voluntarily left the facility during 2005 (Table A.IX).

VI. Discussion

Based on a small but rigorous empirical literature on community participation and oversight, and extensive piloting in the field, our conjecture was that lack of relevant information and failure to agree on, or coordinate expectations of, what is reasonable to demand from the provider were holding back individual and group action to pressure and monitor the provider. We designed an intervention aimed at relaxing these constraints. Through two rounds of community meetings, local NGOs initiated a process aimed at energizing the community and agreeing on actions to improve service provision.

We document large increases in utilization and improved health outcomes that compare favorably to some of the more successful community based intervention trials reported in the medical literature. However, while the medical field trials address the question of impact of a biological agent or treatment practice when the health workers do what they are suppose to do, we focus on a mechanism to ensure that health workers exert effort to serve the community.

The project was implemented in nine districts in Uganda with an estimated catchment population of approximately 55,000 households. In this dimension, therefore, the project has already shown that it can be brought to scale. However, the literature on how to enhance local accountability and participation is still in its infancy. And while the results in the paper suggest that community monitoring can play an important role in improving service delivery when traditional top-down supervision is ineffective, there are still a number of outstanding questions. For example, we know little about long-term effects and cross-sector externalities. It may also be the case that combining bottom-up monitoring with a reformed top-down approach could yield even better results. Before scaling up, it is also important to subject the project to a cost-benefit analysis. This would require putting a value on the improvements we have documented. To provide a flavor of such a cost-benefit analysis, consider the findings on averting the death of a child under five. A back-of-the-envelope calculation suggests that the intervention, including the cost for collecting data for the report cards (the main cost item), at a cost of $3 per household in the catchment areas or $160,000 in total, only judged on the cost per death averted must be considered to be fairly cost-effective. The estimated cost of averting the death of a child under five is around $300, which should be compared to the estimate that the average cost per child life saved through the combined and integrated delivery of 23 interventions shown to reduce mortality from the major causes of death in children younger than 5 years is $887 (Bryce at al. 2005).

As argued in a recent *Lancet* article, a systematic program of research to answer questions about how best to deliver health (child survival) interventions is urgently needed (Bryce et al. 2003). In this paper we have focused on a mechanism that have been highlighted, but not examined, in the literature – a mechanism of accountability enabling (poor) people to scrutinize whether or not those in authority have fulfilled their health responsibilities. Future research should address long term effects, identify which mechanisms or combination of mechanisms that are important, and study the extent to which the results generalize to other social sectors.

IGIER, Bocconi University, and CEPR
Institute for International Economic Studies, Stockholm University, NHH, and CEPR

References

Akerlof, George A., and Rachel E. Kranton, "Identity and the Economics of Organizations," *Journal of Economic Perspectives,* 19 (2005), 9-32.

Banerjee, Abhijit, Angus Deaton, and Esther Duflo, "Wealth, Health, and Health Service Delivery in Rural Rajasthan," *American Economic Review,* 94 (2004), 326-330.

Bang, Abhay T., Rani A. Bang, Sanjay B. Baitule, M. Hanimi Reddy, and Mahesh D. Deshmukh, "Effect of Home-based Neonatal Care and Management of Sepsis on Neonatal Mortality: Field Trial in Rural India," *Lancet*, 354 (1999), 1955-1961.

Bang, Abhay T., Rani A. Bang, O. Tale, P. Sontakke, J. Solanki, R. Wargantiwar, and P. Kelzarkar, "Reduction in Pneumonia Mortality and Total Childhood Mortality by Means of Community-based Intervention Trial in Gadchiroli, India," *Lancet*, 336 (1990), 201-206.

Besley, Timothy, and Robin Burgess, "The Political Economy of Government Responsiveness: Theory and Evidence From India," *Quarterly Journal of Economics,* 177 (2002), 1415-51.

Black, Robert E., Saul S. Morris, and Jennifer Bryce, "Where and Why are 10 million Children Dying Every Year?," *Lancet,* 361 (2003), 2226-34.

Bryce, Jennifer, Shams el Arifeen, George Pariyo, Claudio F. Lanata, Davidson Gwatkin, and Jean-Pierre Habicht, the Multi-Country Evaluation of IMCI Study Group, "Reducing Child Mortality: Can Public Health Deliver?," *Lancet,* 362 (2003), 159-64.

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.

Darmstadt, Gary L., Zulfiqar A. Bhutta, Simon Cousens, Taghreed Adam, Neff Walker, and Luc de Bernis, Lancet Neonatal Survival Steering Team, "Evidence-based, Cost-effective Interventions: How Many Newborn Babies Can We Save?," *Lancet,* 365 (2005), 977-88.

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

Fauveau, Vincent, M. K. Stewart, J. Chakraborty, and S. A. Khan, "Impact on Mortality of a Community-based Programme to Control Acute Lower Respiratory Tract Infections," *Bulletin of the World Health Organization*, 70 (1992), 109-16.

Ferraz, Claudio, and Frederico Finan, "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes," *Quarterly Journal of Economics,* 123 (2008), 703–745.

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.

Khemani, Stuti, "Can Information Campaigns Overcome Political Obstacles to Serving the Poor," World Bank Working Paper, 2006.

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

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.

Kidane, Gebreyesus, and Richard H. Morrow, "Teaching Mothers to Provide Home Treatment of Malaria in Tigray, Ethiopia: A Randomised Trial," *Lancet*, 356 (2000), 550-555.

Manandhar, Dharma S., David Osrin, Bhim Prasad Shrestha, Natasha Mesko, Joanna Morrison, Kirti Man Tumbahangphe, Suresh Tamang, Sushma Thapa, Dej Shrestha, and Bidur Thapa, "Effect of a Participatory Intervention with Women's Groups on Birth Outcomes in Nepal: Cluster-randomised Controlled Trial," *Lancet*, 364 (2004), 970-979.

Mtango F.D., and D. Neuvians, "Acute Respiratory Infections in Children under Five Years. Control Project in Bagamoyo District, Tanzania," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 80 (1986), 851-8.

McPake, Barbara, Delius Asiimwe, Francis Mwesigye, Mathias Ofumbi, Lisbeth Ortenblad, Pieter Streefland, and Asaph Turinde, "The Economic Behavior of Health Workers in Uganda: Implications for Quality and Accessibility of Public Health Services," *Social Science and Medicine,* 49 (1999), 849-865.

Miguel, Edward, and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica,* 72 (2004), 159-217.

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

Rahman, M., L.C. Chen, J. Chakraborty, M. Yunus, A.I. Chowdhury, A.M. Sarder, S. Bhatia, and G.T. Curlin, “Use of Tetanus Toxoid for the Prevention of Neonatal Tetanus. 1. Reduction of Neonatal Mortality by Immunization of Non-pregnant and Pregnant Women in Rural Bangladesh,” *Bulletin of the World Health Organization*, 60 (1982), 261-7.

Rahmathullah, Lakshmi, James M. Tielsch, R D. Thulasiraj, Joanne Katz, Christian Coles, Sheela Devi, Rajeesh John, Karthik Prakash, A V. Sadanand, N. Edwin, and C. Kamaraj, “Impact of Supplementing Newborn Infants with Vitamin A on Early Infant Mortality: Community Based Randomised Trial in Southern India,” *BMJ,* 327 (2003), 254.

Reinikka, Ritva, and Jakob Svensson, “Local Capture: Evidence from a Central Government Transfer Program in Uganda,” *Quarterly Journal of Economics,* 119 (2004), 679-705.

Republic of Uganda, *National Health Policy and Health Sector Strategic Plan 2000/01-2004/05* (Kampala, Uganda: Ministry of Health, 2000).

Strömberg, David, “Radio’s Impact on Public Spending”, *Quarterly Journal of Economics,* 119 (2004), 189-221.

Sazawal, Sunil, and Robert E. Black, Pneumonia Case Management Trials Group, “Effect of Pneumonia Case Management on Mortality in Neonates, Infants, and Preschool Children: A Meta-analysis of Community-based Trials,” *Lancet Infectious Diseases*, 3 (2003), 547-556.

United Nations Children’s Fund, *The State of the World’s Children 2007: Women and Children - The Double Dividend of Gender Equality* (New York, NY: United Nations Children’s Fund, 2006).

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

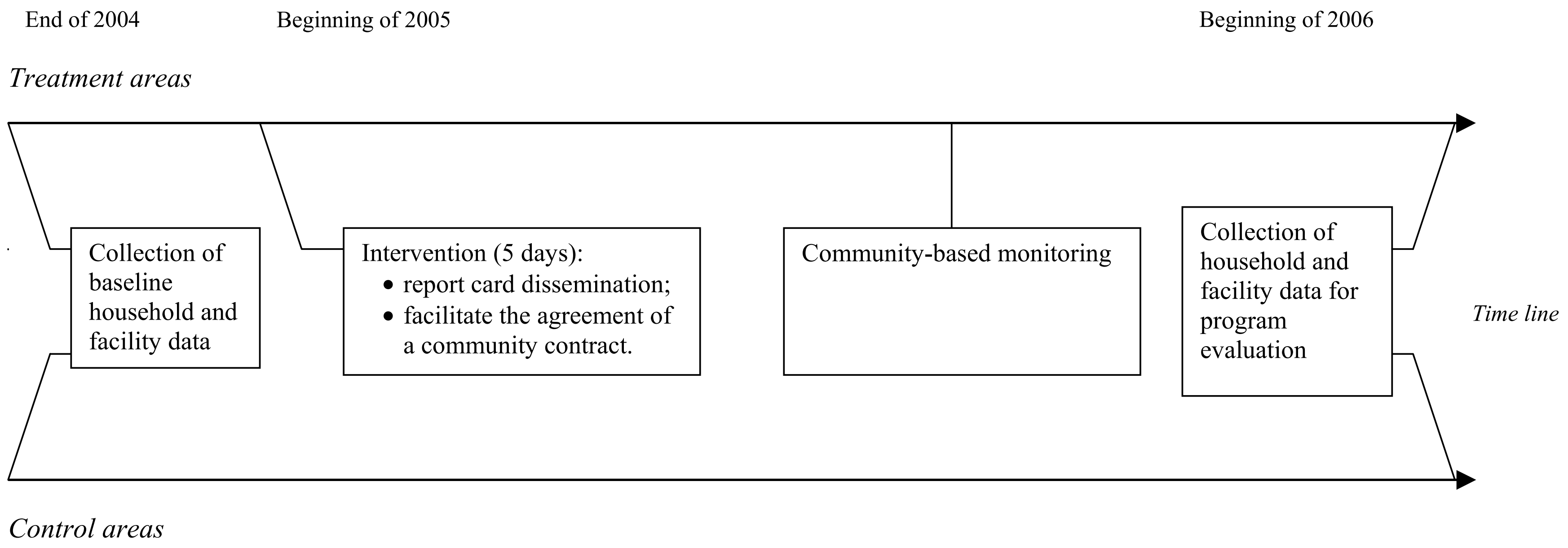

FIGURE I
Timing of the project

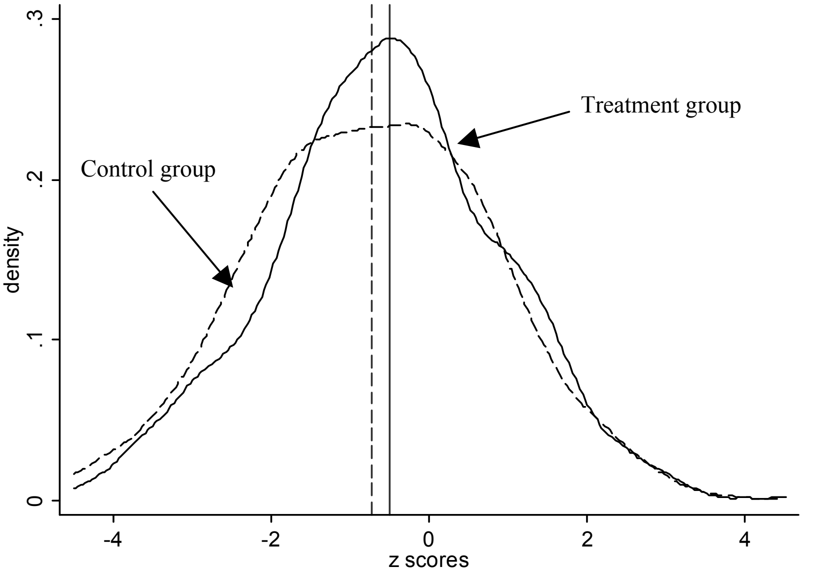

FIGURE II

Distributions of weight-for-age z-scores for treatment and control groups

Notes: Weight-for-age z-scores for children under 18 months excluding observations with recorded weight above the 90th percentile in the growth chart reported in Cortinovis et al (1997). Sample size is 1135 children. Vertical solid line denotes mean in treatment group, dashed line denotes mean in control group.

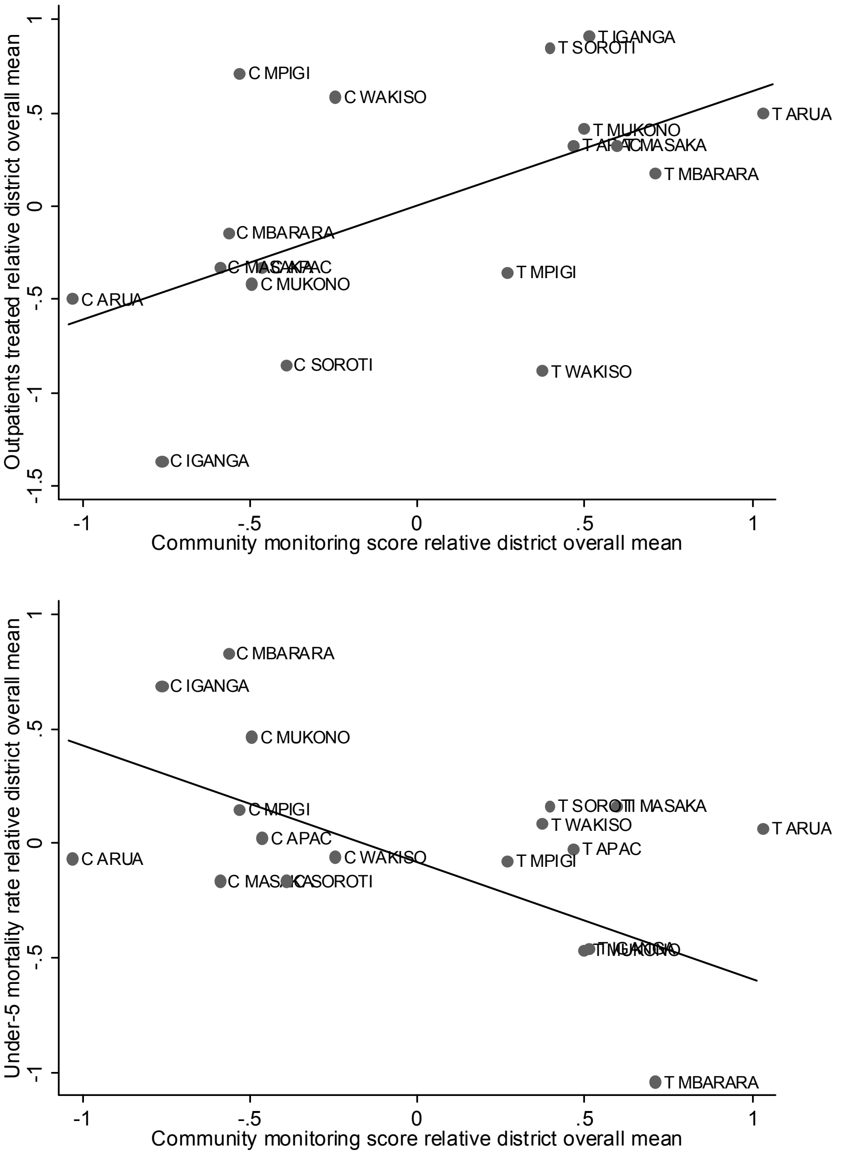

FIGURE III

Differences in treatment-control in outcomes and monitoring across districts

Note**:** Partial regression plots. The community monitoring index, outpatients, and under-five mortality rate in the community (all three variables are described in the main text) are expressed in standard deviation units relative to the control group overall standard deviation for each variable. The points are the average values by group (treatment and control) for each district, normalized so that each district has mean zero. The line passes through the origin with the slope from the 2SLS estimation of equation (4) of the outcome on community monitoring and district indicators, using district-by-treatment interactions as instrumental variables. "T" ["C"] denotes treatment [control] group.

TABLE I
PRE-TREATMENT FACILITY AND CATCHMENT AREA CHARACTERISTICS AND AVERAGE STANDARDIZED EFFECTS

Variables	Treatment group	Control group	Difference
Key characteristics:			
Out-patient care	593	675	-82
	(75)	(57)	(94)
Delivery	10.3	7.5	2.8
	(2.2)	(1.4)	(2.6)
No. of households in catchment area	2140	2224	-84.4
	(185)	(204)	(276)
No. of households per village	93.9	95.3	-1.42
	(5.27)	(6.32)	(8.23)
Drink safely today	0.40	0.32	0.08
	(0.10)	(0.10)	(0.14)
No. of days without electricity in last month	18.3	20.4	-2.12
	(2.95)	(2.90)	(4.14)
Average standardized pre-treatment effects:			
Utilization			0.11
			(0.77)
Utilization pattern			-0.48
			(0.33)
Quality measures			-0.35
			(0.84)
Catchment area statistics			0.11
			(0.66)
Health facility characteristics			0.14
			(0.31)
Citizen perceptions			0.37
			(0.67)
Supply of drugs			0.73
			(0.83)
User charges			-0.65
			(0.63)

Notes: Key characteristics are catchment area/health facility averages for treatment and control group and difference in averages. Robust standard errors in parentheses. Significantly different from zero at 99 (***), 95 (**), and 90 (*) percent confidence level. Description of variables: Out-patient care is average number of patients visiting the facility per month for out-patient 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 out of 31 days. Average standardized pre-treatment effects are derived by estimating (3) on each family of outcomes. Utilization summarizes outpatients and deliveries. Utilization pattern summarizes the seven measures in Table A.I., reversing sign of traditional healer and self treatment. Quality measure summarizes the two measures in Table A.I., reversing sign of waiting time. Catchment area statistics summarize the four measures in Table A.I. Health facility characteristics summarize the eight measures in Table A.I. and drink safely today and days without electricity, reversing sign of days without electricity and distance to nearest local council. Citizen perceptions summarize the four measures in Table A.I. Supply of drugs summarizes the five measures in Table A.I. User charges summarize the four measures in Table A.I. reversing all signs. The Chi(2) test-statistic on the joint hypothesis that all average standardized effects are zero is 4.70 with p-values = 0.79.

TABLE II
PROGRAM IMPACT ON MONITORING AND INFORMATION

Dependent variable	Suggestion box	Numbered waiting cards	Poster informing free services	Poster on patients' rights	Average standardized effect	Discuss facility in LC meetings	Received information about HUMC
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Program impact	0.32***	0.16*	0.27***	0.14	2.55***	0.13***	0.04***
	(0.08)	(0.09)	(0.09)	(0.10)	(0.55)	(0.03)	(0.01)
Mean control group	0	0.04	0.12	0.12	-	0.33	0.08
Observations	50	50	50	50	50	3119	4996

Notes: Robust standard errors in parentheses. Disturbance terms are clustered by catchment areas in columns (vi)-(vii). *** [**] (*) denote significance at the 1 [5] (10) percent level. Point estimates, standard errors, and average standardized effect, columns (i)-(v), are derived from equation (3). Program impact measures the coefficient on the assignment to treatment indicator. Outcome measures in columns (i)-(iv) are based on data collected through visual checks by the enumerators during the post-intervention facility survey. Outcome measures in columns (vi)-(vii) are from the post-intervention household survey. The estimated equations all include district fixed effects and the following baseline covariates: number of villages in catchment area, number of days without electricity in the last month, indicator variable for whether the facility has a separate maternity unit, distance to nearest public health provider, number of staff with less than advanced A-level education, indicator variable for whether the staff could safely drink from the water source, and average monthly supply of Quinine. Specification: (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 facility has a poster informing about free health services; (iv) Indicator variable for whether the facility has a poster on patients' rights and obligations; (v) Average standardized effect of the estimates in columns (i)-(iv); (vi) Indicator variable for whether the household discussed the functioning of the health facility at a Local council meeting during the past year; (vii) Indicator variable for whether the household has received information about the Health Unit Management Committee's (HUMC's) roles and responsibilities.

TABLE III

PROGRAM IMPACT ON TREATMENT PRACTICES AND MANAGEMENT

Spec.	Dep. variable	Model	Program impact	2005	Mean control group 2005	Obs.
(i)	Equipment used	DD	0.08**	-0.07***	0.41	5280
			(0.03)	(0.02)		
(ii)	Equipment used	OLS	0.01		0.47	2758
			(0.02)			
(iii)	Waiting time	DD	-12.3*	-12.4**	131	6602
			(7.1)	(5.2)		
(iv)	Waiting time	OLS	-5.16		131	3426
			(5.51)			
(v)	Absence rate	OLS	-0.13**		0.47	46
			(0.06)			
(vi)	Management of clinic	OLS	-1.20***		0.49	50
			(0.33)			
(vii)	Health information	OLS	0.07***		0.32	4996
			(0.02)			
(viii)	Importance of family planning	OLS	0.06***		0.31	4996
			(0.02)			
(ix)	Stock-outs	OLS	-0.15**		0.50	42
			(0.07)			

Notes: Each row is based on a separate regression. The DD model is from equation (2). The OLS model is from equation (1) with district fixed effects and baseline covariates as listed in Table II. Robust standard errors, clustered by catchment areas in columns (i)-(iv) and (vii)-(viii), in parentheses. Significantly different from zero at 99 (***), 95 (**), and 90 (*) percent confidence level. Program impact measures the coefficient on the assignment to treatment indicator in the OLS models and the assignment to treatment indicator interacted with an indicator variable for 2005 in the DD models. Specification: (i)-(ii) Indicator variable for whether the staff used any equipment during examination when the patient visited the health facility; (iii)-(iv) Difference between the time the citizen left the facility and the time the citizen arrived at the facility, subtracting the examination time; (v) The ratio of workers not physically present at the time of the post-intervention survey to the number of workers employed pre-intervention (see text for details); (vi) The first component from a principal components analysis of the variables Condition of the floors of the health clinic, Condition of the walls, Condition of furniture, and Smell of the facility. Each condition is ranked from 1 (dirty) to 3 (clean) by the enumerators; (vii) Indicator variable for whether the household has received information about the importance of visiting the health facility and the danger of self-treatment; (viii) Indicator variable for whether the household has received information about family planning; (ix) Share of months in 2005 in which stock-cards indicated no availability of drugs (see text for details).

TABLE IV

PROGRAM IMPACT ON IMMUNIZATION

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	1.30*	1.44**	1.24**	0.72	2.01***	0.86
	(0.70)	(0.72)	(0.63)	(0.58)	(0.67)	(0.80)
Observations	173	929	940	951	1110	526

Notes: Average standardized effects are derived from equation (3) with the dependent variables being indicator variables for whether the child has received at least one dose of measles, DPT, BCG, polio, and Vitamin A supplement, respectively (see text for details) and with district fixed effects and baseline covariates listed in Table II included. Robust standard errors clustered by catchment areas in parentheses. *** [**] (*) denote significance at the 1 [5] (10) percent level. 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.

TABLE V

PROGRAM IMPACT ON UTILIZATION/COVERAGE

Dep. variable	Out-patients	Delivery	Antenatal	Family planning	Average std effect	Use of project facility	Use of self treatment /traditional healers	Average std effect
PANEL A: Cross-section data	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Program impact	130.2**	5.3**	15.0	3.4	1.75***	0.026*	-0.014	1.43*
	(60.8)	(2.1)	(11.2)	(3.2)	(0.63)	(0.016)	(0.011)	(0.87)
Observations	50	50	50	50	50	50	50	50
PANEL B: Panel data	(ix)	(x)			(xi)	(xii)	(xiii)	(xiv)
Program impact	189.1***	3.48*			2.30***	0.031*	-0.046**	1.96**
	(67.2)	(1.96)			(0.69)	(0.017)	(0.021)	(0.89)
Observations	100	100			100	100	100	100
Mean control gr. 2005	661	9.2	78.9	15.2	-	0.24	0.36	-

Notes: Panel A reports program impact estimates from cross-section models with district fixed effects and baseline covariates as listed in Table II, with robust standard errors in parentheses. Panel B reports program impact estimates from difference-in-differences models with robust standard errors clustered by facility in parentheses. *** [**] (*) denote significance at the 1 [5] (10) percent level. Point estimates, standard errors, and average standardized effects in specifications (i)-(v), (vi)-(viii), (ix)-(xi), and (xii)-(xiii) are derived from equation (3). Program impact measures the coefficient on the assignment to treatment indicator in the OLS models and the assignment to treatment indicator interacted with an indicator variable for 2005 in the DD models. Specification: First column is average number of patients visiting the facility per month for out-patient 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 average standardized effect of estimates in specification (i)-(iv) and (ix)-(x), respectively. Sixth column is the share of visits to the project facility of all health visits, averaged over catchment area; Seventh column is the a share of visits to traditional healers and self-treatment of all health visits, averaged over catchment area; Eight column is average standardized effect of estimates in specification (vi)-(vii) and (xii)-(xiv), respectively, reversing the sign of use of self treatment/traditional healers.

TABLE VI

PROGRAM IMPACT ON HEALTH OUTCOMES

Dependent variable	Births	Pregnancies	U5MR	Child death	Weight-for-age z-scores	
Specification	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Program impact	-0.016	-0.03**	-49.9*		0.14**	0.14**
	(0.013)	(0.014)	(26.9)		(0.07)	(0.07)
Child age (log)						-1.27***
						(0.07)
Female						0.27***
						(0.09)
Program impact×Year of birth 2005				-0.026**		
				(0.013)		
Program impact× Year of birth 2004				-0.019**		
				(0.008)		
Program impact× Year of birth 2003				0.003		
				(0.009)		
Program impact× Year of birth 2002				0.000		
				(0.006)		
Program impact× Year of birth 2001				0.002		
				(0.006)		
Mean control gr. 2005	0.21	0.29	144	0.029	-0.71	-0.71
Observations	4996	4996	50	5094	1135	1135

Notes: Estimates from equation (1) with district fixed effects and baseline covariates as listed in Table II included. Specification (iv) also includes a full set of year-of-birth indicators. Robust standard errors in parentheses (iii), clustered by catchment area (i)-(ii), (iv)-(vi). *** [**] (*) denote significance at the 1 [5] (10) percent level. Program impact measures the coefficient on the assignment to treatment indicator. Specification: (i) Number of births in the household in 2005; (ii) Indicator variable for whether any women in the household are or have been pregnant in 2005; (iii) U5MR is under-5 mortality rate in the community expressed per 1,000 live births (see text for details); (iv) Indicator variable for child death in 2005; (v)-(vii) Weight-for-age z-scores for children under 18 months excluding observations with recorded weight above the 90th percentile in the growth chart reported in Cortinovis et al (1997).

TABLE VII

MECHANISMS AND ROBUSTNESS

Dep. variable Specification	Out-patients (i)	U5MR (ii)	Out-patients (iii)	U5MR (iv)	Out-patients (iii)	U5MR (iv)	Out-patients (vii)	U5MR (viii)
Community monitoring index	0.77***	-0.43*	0.86*	-0.43	0.77**	-0.54*		
	(0.22)	(0.25)	(0.53)	(0.82)	(0.21)	(0.30)		
Staff's knowledge about patients' rights					-0.01	0.47		
					(0.28)	(0.29)		
Program impact			-0.12	0.01			190.5**	-41.3
			(0.66)	(0.88)			(92.6)	(45.8)
CBO presence							-8.3	-21.0
							(69.4)	(37.9)
Program impact×CBO presence							-127.9	-4.0
							(126.1)	(58.4)
F-test on program impact							6.17	
							(0.05)	
F-test on CBO presence							0.37	
							(0.83)	
F-test on Program impact×CBO presence							1.03	
							(0.60)	

Notes: Columns (i)-(iv) report 2SLS estimates from equation (4) with district-by-treatment interactions as the excluded instruments and district fixed effects and outpatients$_{t-1}$ (in specifications (i) and (iii)) as controls. The variables in columns (i)-(iv) are expressed in standard deviation units relative to the control group overall standard deviation for each variable. Robust standard errors in parentheses. *** [**] (*) denote significance at the 1 [5] (10) percent level. Program impact measures the coefficient on the assignment to treatment indicator. F-test statistics (with p-values in parentheses) on the excluded instruments Community monitoring and Staff's knowledge about patient's rights are 15.9 (0.00) and 7.23 (0.00), respectively. Point estimates and standard errors in columns (v)-(vi) and columns (vii)-(viii), respectively, are jointly estimated from equation (3). Explanatory variables: Community monitoring is the first component from a principal components analysis of the six monitoring and information proxies presented in Table II. Staff's knowledge about patients' rights is a measure of the in-charge's knowledge about patients' rights and obligations (see text for details). CBO presence is an indicator variable for whether a participating CBO had been operating in the community before the intervention. F-test on program impact [CBO presence] {Program impact×CBO presence} is the test statistic, with p-values in parenthesis, on the test that the coefficients on program impact [CBO presence] {Program impact×CBO presence} are jointly zero in columns (v)-(vi) and (vii)-(viii), respectively.