WPS4268

Power to the People:
Evidence from a Randomized Field Experiment
of a Community-Based Monitoring Project in Uganda

Martina Björkman∗ and Jakob Svensson#

Abstract: Strengthening the relationship of accountability between health service providers and citizens is by many people viewed as critical for improving access to and quality of health care. How this is to be achieved, and whether it works, however, remain open questions. The paper presents a randomized field experiment on increasing community-based monitoring. As communities began to more extensively monitor the provider, both the quality and quantity of health service provision improved. One year into the program, there are large increases in utilization, significant weight-for-age z-score gains of infants, and markedly lower deaths among children. The findings on staff behavior suggest that the improvements in quality and quantity of health service delivery resulted from an increased effort by the staff to serve the community. Overall, the results suggest that community monitoring can play an important role in improving service delivery when traditional top-down supervision is ineffective.

1 Introduction

Approximately 11 million children under five die each year. Almost half of these deaths occur in Sub-Saharan Africa where roughly one in five children dies before reaching the age of five. More than half of these children – nearly 6 million – will die of diseases that could easily have been prevented or treated if the children had had access to a small set of proven, inexpensive services.[1]

Why are these services not provided? While there is no simple answer, a wealth of anecdotal, and recently more systematic, evidence shows that the provision of public services to poor people in developing countries is constrained by weak incentives of service providers – schools and health clinics are not open when supposed to; teachers and health workers are frequently absent from schools and clinics and, when present, spend a significant amount of time not serving the intended beneficiaries; equipment, even when fully functioning, is not used; drugs and vaccines are misused; and public funds are expropriated.[2]

The traditional approach to accountability in the public sector relies on external control. This is a top-down approach where someone in the institutional hierarchy is assigned to monitor, control and reward/punish agents further down in the hierarchy. The tacit assumption is that more and better enforcement of rules and regulations will strengthen providers' incentives to increase both the quantity and quality of service provision. But, in many poor countries, the institutions assigned to monitor the providers are typically weak and malfunctioning, and may themselves act under an incentive system providing little incentives to effectively monitor the providers. As a result, the relationship of accountability of provider-to-state is ineffective in many developing countries.[3]

---

[1]See Lancet (2003) and UNICEF (2003). It is estimated that 2 million children under five die from diarrhea, which in most cases can be treated with simple oral rehydration therapy. Another 2 million children die from pneumonia, where once more there is sufficient evidence of effective treatment (antibiotics). Malaria kills one million children under five, most of whom could have been protected by preventive measures and treatment with anti-malarias. Globally, neonatal disorders account for the highest proportion of deaths of children – many of them could have been saved if mothers had had access to basic antenatal and delivery care. Approximately half a million children under five die from measles, for which these is a cheap and effective vaccine (Black et al., 2003; Jones et al., 2003).

[2]For anecdotal and case study evidence, see World Bank (2003). Chaudhury et al. (2006) provide systematic evidence on the rates of absenteeism based on surveys where enumerators made unannounced visits to primary schools and health clinics in seven developing countries. Averaging across countries, 35 percent of the health workers were absent. Banerjee et al. (2004) and Duflo and Hanna (2005) confirm these findings. On misappropriation of public funds and drugs, see Reinikka and Svensson (2004) and McPake et al. (1999).

[3]In addition, while well-functioning legal and financial systems can curtail obvious cases of mismanagement, they only partially constrain the discretionary powers of public sector managers and employees. The complexity of the tasks performed by a typical public sector unit and its informational advantage relative to the monitor (which typically relies on accounting data) make it nearly impossible to design legal and accounting measures to address all types of misuse and thus, to curtail less obvious cases of mismanagement (such as shirking, budget prioritization in favor of staff and political considerations). Finally, audit reports and legal procedures are often difficult for nonspecialists to interpret and will therefore go unnoticed unless the commissioning agency acts on them.

As a complementary strategy, it has therefore been argued that more effort must be placed on strengthening beneficiary control, i.e. strengthening providers' accountability to citizen-clients (see e.g., World Bank, 2003). However, despite the enthusiasm for such an approach, there is little credible evidence on the impact of policy interventions aimed at achieving it (Banerjee and He, 2003; Banerjee and Duflo, 2005). This paper attempts to provide some.

To examine whether beneficiary control works, we designed and conducted a randomized field experiment in 50 "communities" from nine districts in Uganda.[4] In the experiment, or intervention, communities were provided with baseline information on the status of service delivery, both in absolute terms and relative to other providers and the government standard for health service delivery. Community members were also encouraged to develop a plan that identified areas of concerns in health provision, and how to address them within the existing resource envelope.

The intervention sought to relax two constraints typically faced by communities in monitoring providers: lack of access to reliable and structured information on the community's entitlements and the status of service delivery, and inadequate local organizational capacity. Access to reliable and structured information about current status of service delivery and entitlements is critical for citizens' ability to monitor service providers. Although people 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. Provision of information on outcomes and performance improves citizens' ability to challenge abuses of the system, since reliable quantitative information is more difficult for service providers to brush aside as anecdotal, partial, or simply irrelevant. But information provision may not have any considerable impact unless there are members of the community who are willing to make use of the new information. Exerting accountability (monitoring providers) is subject to potentially large free-rider problems. Elite capture further complicates the process of holding providers accountable. By enhancing local organization capacity and encouraging the community to develop its own monitoring strategy, these constraints are sought to be relaxed.

The community-based monitoring project increased the quality and quantity of primary health care provision. One year into the program, we find a significant difference in the weight of infants (0.17 z-scores increase) and a markedly lower number of deaths among children under five (a 33 percent reduction in child deaths) in the treatment communities. Utilization (for general outpatient services) was 16 percent higher in the treatment compared to the control facilities. We also find significant differences in the number of deliveries at treatment facilities and the use of antenatal care and family planning. Treatment practices, as expressed both in perception-based responses by households and in more quantitative indicators (immunization of children, waiting

---

[4]A "community" is operationalized as the households (and villages) residing in the five-kilometer radius around the facility (see section 5 for details). Approximately 110,000 households (600,000 individuals) reside in these communities, of which half reside in the treatment communities.

time, examination procedures, 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. No effect is found on investments, or the level of financial or in-kind support (from the government). Furthermore, supervision of providers by upper-level government authorities remained low in both the treatment and the control group. This reinforces our confidence that the findings on the quality and quantity of health care provision resulted from increased efforts by the health unit staff to serve the community in the light of better community monitoring.

The paper is organized as follows. The next section reviews the literature. Section 3 discusses the concept of community monitoring. Section 4 briefly describes the institutional environment in Uganda and in the project areas. The community-based monitoring intervention is described in section 5. Section 6 lays out the evaluation design and the results are presented in section 7. Section 8 concludes.

2   Literature Review

Improving governance and public service delivery through community participation is an approach that has gained prominence in recent years. For example, the World Development Report 2004 is entirely devoted to the concept of increasing poor citizens' voice and participation in service delivery in order to help them monitor and discipline providers. But despite the enthusiasm for such approaches, there is little credible evidence on the impact of policy interventions aimed at achieving them. On the one hand, most (all) comprehensive community based monitoring initiatives have not been rigorously evaluated. On the other hand, the few studies relying on rigorous impact evaluation strategies have not evaluated more comprehensive attempts to inform and involve the community in monitoring public officials.

On the latter issue, Olken (2005) 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. However, although the invitations increased the number of people participating in the meetings, the meetings were still dominated by members of the village elite. Moreover, corruption is not easily observable and project officials may very well be able to hide it when reporting on how funds were used. The data also reveal that corruption problems were seldom discussed in these meetings.[5]   Thus, it is unclear to what extent non-elite community members were

---

[5]The information problem is illustrated in the novel but burdensome way in which Olken (2005) estimates the extent of corruption. Specifically, Olken (2005) assembled a team of engineers and surveyors who dug samples in roads to estimate the quantity of materials used and then, using price information from local supplies, estimated the extent of "missing" expenditures. The corruption estimates were not reported in the village meetings.

really more informed about corruption in the project, or if they had any means of influencing outcomes, in response to the intervention. Given these constraints, it is not surprising that Olken (2005) only finds minor effects of the intervention.

Using a randomized design, Banerjee, Deaton and Duflo (2004) evaluate a project in Rajasthan in India where a member of the community was paid to check once a week, on unannounced days, whether the auxiliary nurse-midwife assigned to the health center was present at the center. Unlike Olken's study, getting reliable information is not a concern here. In fact, external monitors confirmed the absence rates documented by the community members assigned to the project. The issue is rather how the informed community member could use his or her information on absenteeism to invoke community participation. The intervention had no impact on attendance. Thus, informing one person, even if this is done is a structured and regular way, may not have much impact.

Jiminez and Sawada (1999) examine how decentralizing educational responsibility to communities and schools affects student outcomes. They study El Salvador's Community-Managed Schools Program, EDUCO, and its effect on students' achievement on standardized tests and attendance. The evidence suggests that enhanced community and parental involvement in EDUCO schools improved students' language skills and diminished student absences. A key estimation issue in this paper is endogenous program participation and although the authors instrument for program participation by using the proportion of EDUCO schools in a municipality, it is not obvious that they succeed in obtaining the causal treatment effect.

There is a growing empirical literature on the relationship between information dissemination (through the media) and accountability. With few exceptions, this literature studies the relationships of accountability of politicians to citizens and deal with one (periodic elections), out of several, mechanism through which citizens can make politicians and policymakers accountable.[6] For example, Strömberg (2003, 2004) considers how the press influences redistributive programs in a model of electoral policies, where the role of the media is to raise voter awareness, thereby increasing the sensitivity of turnout to favors granted. Besley and Burgess (2002) focus on the media's role in increasing political accountability also in a model of electoral policies. Ferraz and Finan (2005) study the effects of making information about corruption in the local governments public on the probability of the incumbent winning the election. Our work differs in several important dimensions. First, 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 or programs, but rather on how these resources are utilized. Second, we use micro data from households and health stations rather than disaggregated national accounts data. Finally, we identify impact using an experimental design. The source of identification will thus come directly from a randomized experiment.

Reinikka and Svensson (2005a) also study the relationship between information, accountability, and outcomes at the provider level. They exploit a newspaper cam-

---

[6]For a review, see Khemani (2006).

paign aimed at reducing capture of public funds by providing schools (parents) with information to monitor local officials' handling of a large education grant program. Head teachers in schools closer to a newspaper outlet are found to be more knowledgeable of the rules governing the grant program and the timing of releases of funds by the central government. These schools also managed to claim a significantly larger part of their entitlement after the newspaper campaign had been initiated. Reinikka and Svensson (2005b) and Björkman (2006) take these results as a starting point to explore the effects of increased "client power" on school outcomes. They show that the reduction in capture had a positive effect on both enrollment and student learning. The newspaper campaign in Uganda, however, may not be easy to scale up in other sectors or for more complex government programs. Specifically, the capitation grant is a very simple entitlement program and a small item in a vast government budget. They also identify impact using a non-experimental approach.

3   Community-based Monitoring

Community-based monitoring, or social accountability, is an approach towards building accountability that relies on civic engagement where citizens and civil society organizations directly or indirectly participate in extracting accountability (Malena et al., 2004). It can take a variety of forms, although most interventions have in common that they inform citizens about their rights and status of service delivery and encourage participation.[7]

Citizens/communities typically face several constraints in initiating local collective action to improve service delivery outcomes. First, citizens may not be able to challenge abuses of the system since they lack reliable information on outcomes. Community members' own experience of service provision, or private information, is typically an imprecise signal of overall (or average) quality (Khemani, 2006). Second, like most local collective actions, community-based monitoring is subject to possibly large free-riding problems: The community would like to ensure that the provider performs, but everyone would rather have someone else monitor performance. Third, the community may lack the ability to sanction, either directly or indirectly, the provider in case of poor performance, or reward good performance.

Community-based monitoring, however, is considered to have several potential advantages. For example, it is likely to be cheaper for the beneficiaries to monitor the providers since they (at least as a group) are better informed about the status of service delivery than the external agent assigned to supervise the provider. They may also have means of punishing the provider that are not available to others, such as verbal

---

[7]Examples of this approach include participatory budgeting in Porto Allegre, Brazil; citizen report cards in Bangalore, India; right to information on public works and public hearings or jan sunwais in Rajasthan, India; public information campaign to reduce capture of school funds in Uganda; and community scorecards in Malawi (see Reinikka and Svensson, 2004; World Bank, 2003; Paul, 2002; and Singh and Shah, 2002).

complaints or social opprobrium (Banerjee and Duflo, 2005). Similarly, they may be able to induce higher effort of health workers by providing non-pecuniary rewards (social rewards) for good performance. To the extent that the service is valuable to them, they should also have strong incentives to monitor the provider – incentives which the external agent assigned to supervise the provider may lack. Naturally, there is no guarantee that community monitoring will work even if the community is informed, can coordinate actions and there is demand for the service. In many developing countries, the beneficiaries of health services in rural areas are socially inferior to health care workers. Beneficiary groups may also be captured by the service provider or other authorities through their social or political connections (Banerjee and Duflo, 2005). Thus, in the end, if and to what extent community monitoring works is an empirical question.

4 Institutional Setting

Uganda, like many newly independent countries in Africa, had a functioning health care system in the early 1960s. Accessibility and affordability were relatively extensive. 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. Some health indicators have improved, while others have not. This is despite a GDP growth rate exceeding 64 percent and a 40-percent reduction in consumption poverty in the 1990s (Appleton 2001)

As of 2001, public health services are free of charge. Anecdotal and survey evidence (see below), however, suggests that users still encounter varying costs when visiting public health facilities.

The health sector in Uganda is composed of four types of facilities: hospitals, health centers, dispensaries (health center III), 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 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 (89 percent). 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). In our sample of facilities, on average, a dispensary was staffed by an in-charge or clinical officer (a trained medical worker/doctor), three nurses (including midwives), and three nursing aids or other assistants.

The health sector in Uganda is decentralized and supervision and control of the dispensaries are governed at the district level. A number of actors are responsible for the functioning of the dispensaries. The Health Unit Management Committee (HUMC) is supposed to be the main link between the community and the health

facility. Each dispensary has an HUMC which consists of members from both the health facility staff and non-political representatives from the community (elected by the sub-county local council). The HUMC should monitor drugs and finances disbursed to the health facility, as well as the day-to-day running of the health facility (Republic of Uganda 2000). The HUMC can warn the health facility staff on issues of indiscipline, rudeness to patients and misappropriations of funds by recommending that the staff is transferred from the health facility. However, the HUMC has no authority to dismiss a worker. In cases of problems at the health facility, the working practice is that the chair person of the HUMC raises the issue with the in-charge. If there is no improvement, the issue should be referred to 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 practise, monitoring is infrequent. The Health Sub-district has the authority to reprimand, but not dismiss, health facility staff for indiscipline. In severe cases of indiscipline, therefore, the errand will be referred to the Chief Administrative Officer of the District and the District Service Commission, which is the appointing authority for the district and has the authority to suspend or dismiss staff.

Another actor in the health sector is Community-based organizations (CBOs). Their main focus is on health education in antenatal care, family planning, and HIV/AIDS prevention.

5 The Project: Citizen Report Card

In response to perceived continued 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 50 project facilities (all in rural areas) were drawn from nine districts in Uganda (see the appendix for details).

With the catchment area (or the community) of each dispensary defined as the households and villages residing within a five-kilometer radius from the clinic, about 110,000 households reside in the communities supposedly served.[8] The facilities were first stratified by location (districts) and then by size (the number of households residing in the catchment areas). From each group, half the units, with corresponding catchment areas, were randomly assigned to the treatment group and the remaining 25 units were assigned to the control group. Each district thus had both treatment and control groups.

The main objective of the Citizen report card project was to strengthen providers' accountability to citizen-clients by enhancing communities' ability to monitor providers.

---

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

Specifically, the project aimed at: (i) providing communities with baseline information on the status of service delivery, both in absolute terms and relative to other providers, and the government standard for health service delivery at the dispensary level; and (ii) encouraging people to develop a plan that identified areas of concerns in health provision, and how to address them within the existing resource envelope. These components are discussed next. A time-line and a schematic view of the intervention and expected outcomes are depicted in figures 1 and 2.

5.1 Data Collection and Report Cards

Data collection was governed by two objectives. First, data were required to assemble report cards on 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 unit staff. Second, data were required to rigorously evaluate impact. To meet these objectives, two surveys were implemented: a survey of health care providers and a survey of health care users. Both surveys were implemented prior to the intervention (data from these pre-intervention surveys formed the basis for the report cards) and one year after the project had been initiated.

A quantitative service delivery survey (QSDS) was used to collect data from the health service providers. Since agents in the service delivery system may have a strong incentive to misreport (or not report) 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 submitted to the district-level government. The former, often available in a highly disaggregate format, were considered to suffer the least from any incentive problems in record-keeping.

The household survey collected data on both households' health outcomes and health facility performance, including performance parameters such as usage, availability, access, reliability, quality and satisfaction. To the extent that it was possible, household responses were supported by patient records, i.e., patient exercise books and immunization cards. These records helped the household recall details about its visits to the health facility and also minimized problems of misreporting. 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.

A stratified random sample of households within the catchment area of the facility were surveyed. In total, roughly 5,000 households have been surveyed in each round. The design and implementation of the surveys are explained in more detail in the appendix.

The data from the two pre-intervention surveys were analyzed and a smaller subset of the findings were assembled in report cards for the treatment localities.[9] The data

---

[9] Thus, the design and size of the surveys were largely driven by the second objective – to evaluate impact.

included in the report cards were identified as key areas subject to improvement and include utilization, quality of services, informal user charges and comparisons vis-à-vis other health facilities in the district and the country at large. Each treatment facility and its community had a unique report card summarizing, in a format easily accessible to the communities, the findings from the surveys conducted in their area.

The report cards were translated into the main language spoken in the community.[10] To support the non-literate community members, posters were designed by a local artist so that otherwise complex information and concepts were easily understood. Because the information in the report cards was largely statistical, the posters visually conveyed the main messages, such as where people go to seek medical care and why they do so.[11]

5.2   Dissemination and Participation

Getting people to retain and use information to achieve a specific objective is a complex problem.[12] Extensive piloting concluded that simply reporting the facts would be likely to have little impact. Thus, to maximize the likelihood that the information in the report cards would be used when people decide what actions to take, a participatory approach was chosen where community members themselves actively interpreted and analyzed the information.[13] To this end, the process of providing information and encouraging participation and monitoring was initiated through a series of meetings: a community meeting; a staff meeting; and an interface meeting. Staff from various Community-based organizations (CBO) acted as facilitators in these meetings.[14]

The community meeting was a two-day (afternoons) event with approximately 100 invited participants drawn from the surveyed villages in the catchment area of the health facility. To avoid elite capture, the invited participants consisted of a selection of representatives from different spectra of society (i.e. young, old, disabled, women,

---

[10] In the end, the report cards were translated into six different languages: Ateso (Soroti) (Iganga), Lango (Apac), Luganda (Masaka, Wakiso, Mukono and Mpigi), Runyankore (Mbarara) and Lugbara (Arua).

[11] See the appendix for a prototype poster.

[12] See, for example, Lupia (2004) who systematizes and draws conclusions from clinical, psychological, and economic research on information transmission and processing.

[13] The approach we used is closed linked to the so-called "participatory rural appraisal approach". Participatory rural appraisal (PRA) is a label given to a growing family of participatory approaches and methods with the common aim of enabling people to make their own appraisal, analyses, and plans. PRA evolved from a set of informal techniques used by development practitioners in rural areas to collect and analyze data (World Bank, 1996).

[14] Since the CBOs were in regular interaction with the communities and had a mandate drawn from a long-term presence on the ground working with the community, these facilitators were perceived to be a good conduit through which the project could be delivered. The CBO facilitators were trained for seven days in data interpretation and dissemination, utilisation of the participatory methodology, and conflict resolution and management. It should be noted that various CBOs (including some participating in the project) also operate in the control districts. Thus, the presence (and numbers) of CBOs in the project communities is similar across treatment and control groups.

mothers, leaders). The facilitators mobilized the village members by cooperating with village council representatives in the catchment area. Invited participants were asked to spread the word about the meeting and, in the end, a large number of uninvited participants from other villages who had found out about the event also attended the meeting. A typical village meeting was attended by more than 150 participants per day.

In the community meeting, the facilitators used a variety of methods, including maps, venn diagrams, role-play, and focus group discussions, to disseminate the information in the report cards in a participatory, or interactive, way.[15] Information on patients' rights and entitlements was also discussed.[16] As the objective was not only to inform but to encourage people to participate in developing a shared view on how to improve service delivery and monitor the provider, the facilitators structured the discussions through a series of questions on the various elements of accountability in the primary health sector (*who* is accountable to *whom*?; *what* is a particular actor accountable *for*?; *how* can these actors account for their *actions*?; and how are these elements reflected in the report card findings?). 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 as the most important to address; how these issues could be addressed and how the community could monitor improvements (or lack thereof). An abbreviated version of one such action plan is depicted in the appendix. 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. After the meeting, participants were given posters and copies of the report card to bring back to their villages and share with their village members.

The health facility staff meeting was a one-day (afternoon) meeting held at the health facility with all staff present. In this meeting, the facilitators contrasted the information on service provision as reported by the provider with the findings from the household survey. The meeting enabled the providers to review and analyze their performance, and compare their performance with other health clinics in the district and across the country.

An interface meeting with participants (chosen by people that attended the community meeting) from villages in the catchment area and the health facility staff followed the community and health facility meetings. Based on the action plan developed in the community meeting and the discussions from the health facility meeting, the interface

---

[15]See the appendix for a more detailed description of the various methods used during the meetings.

[16]Information on patients' rights and entitlements was based on the Yellow Star program. In 2000, the MoH developed a quality of care strategy called the Yellow Star Program with the aim of improving and maintaining basic standards of care at government and NGO health facilities. The rationale behind this strategy was the general consensus that the quality of health services had been a major deterrent to service utilization. The Yellow Star Program lists a set of basic standards of quality. The standards fall into six categories: Infrastructure and Equipment; Management systems; Infection prevention: Information: Education and Communication; Clinical skills; and Client services.

meeting devised a strategy for improved health care provision. During the interface meeting, the community representatives and the health facility staff presented and discussed their suggestions for improvements. A role-playing exercise was used to disseminate the results from the survey, with community participants and staff reversing roles. 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 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. Copies of the community contract were kept with the community and the health facility to support the following monitoring process.

5.3   Ongoing process of monitoring

The three separate meetings aimed at kick-starting the process of community monitoring. Thus, after the initial meetings, and based on the agreements in the community contract, the communities were themselves in-charge of establishing ways of monitoring the provider. The facilitators supported the communities in this process with follow-up meetings. This was done as an integrated part of the CBO's ordinary work in the villages. Each community had approximately two follow-up meetings in the six-month period that followed. In these meetings, facilitators raised the issues identified in the community contract with citizens and community leaders.

After a period of six months, the communities and health facilities were revisited to conduct a mid-term review – a repeat engagement on a smaller scale. Including a one-day community meeting and a one-day interface meeting, the review tracked the implementation of the community contract. The earlier community contracts were printed on posters to spark discussions. Health facility staff and community members jointly discussed suggestions on actions for sustaining or improving progress, or in the case of no improvements, why so. Where improvements had been made, suggestions for sustainability were recorded. The community and the health facility kept the updated community contract to assist in further monitoring.

6   Evaluation Design and Expected Outcomes

6.1   A stylized framework

The following sub-section presents a stylized framework that illustrates the channel(s) through which community-based monitoring could affect health outcomes.

The key behavioral change induced by more extensive community-based monitoring is expected to be increased effort by health unit staff to serve the community. Health workers have little pecuniary incentives to exert high effort. Typically, 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. While formal sanctions, such as suspensions and dismissal, are possible, they are in practice uncommon and only applied in cases of severe neglect and mismanagement. An individual worker may still put in high effort if shirking deviates from her ideal choice, given the behavior of other staff and the situation (Akerlof and Kranton, 2005). The effort choice may also be influenced by the social rewards from community members or social sanctions against shirking health workers. Social rewards and sanctions are key instruments available to the community to boost health worker's effort.

The effort of a health worker could theoretically complement or substitute the effort of other health workers. Complementarities could directly arise in health service production. There could also be externalities working through the informal instruments at the community's disposal (social rewards and sanctions). For example, to the extent that social rewards and sanctions are facility specific, a high effort worker may benefit little from social rewards and suffer from social sanctions if other workers are shirking. This, in turn, could generate multiplier effects in average individual effort and thus open up the possibility of multiple health facility equilibria: some with high effort by all staff and others with poor overall service quality.

As the community receives more accurate information about service quality and can coordinate on expected reforms, i.e. the intervention, the community is likely to be in a better position to monitor effort but may also choose to more regularly exploit the instruments at their disposal, i.e., praise workers when service provision improves and complain when it does not. Workers may then find coming to work, or more generally exerting effort, more attractive. Complementarities in workers' effort combined with a more engaged and supportive community can therefore result in a virtuous circle, where higher effort by some staff makes it more attractive for others to also come to work as the social prestige of working in a well functioning health clinic rises. As service quality improves, community members in turn shift from self-treatment to the facility in question, which further boosts the return to effort as more households will commend the staff for its effort. The switch from self-treatment to professional care and the increase in quality could both have a positive effect on health outcomes.

6.2 Outcomes

The main outcome of interest is whether the intervention increased the quantity and quality of health care and, thus, improved health outcomes in the treatment communities. However, we are also interested in evaluating changes (if any) in all steps in the accountability chain depicted in figure 2: Did the intervention increase treatment communities' ability to exercise accountability? Did it result in behavioral changes of the staff (i.e., did staff exert higher effort to serve the community)?

As a robustness test, we also assess alternative explanations. Some of these alternative mechanisms are illustrated in figure 3. One concern is spillovers. Spillovers could affect the estimates in two ways. If information about the intervention spread to control areas and, as a result, control communities became more involved in monitoring the providers, the estimated treatment effect would be biased downward. If, on the other hand, households in control communities shifted from seeking care at the control facility to the nearest treatment clinic, it is possible that the estimated treatment effect would be biased upward. This is a potentially serious concern but also a mechanism which we can test. It is also possible 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. For example, the various upper-level authorities in the health sector (e.g. the Health Sub-district) may have become more involved in monitoring the providers, or the district government may have increased its administrative or financial support, following the intervention. While this would not invalidate the causal effect of the intervention it would, of course, affect interpretation. Therefore, this alternative hypothesis is also subject to a battery of tests.

Given the wealth of information, we report the main results and tables in the text and refer the reader to appendix for additional findings.

6.3 Statistical framework

Given the randomized assignment of the Citizen Report Card project, we expect the 2004 pre-data in the treatment areas to be similar those in the control areas. We have both facility-specific data (on utilization, for example) and household-specific data (on waiting time, for example). Denoting y_{ijdt} the outcome variable of household i (when applicable), health facility j in district d and period t, we start by checking that there is no difference between treatment and control facilities/communities prior to the intervention:

$$y_{ijdPRE} = \alpha_{PRE} + \beta_{PRE}T_{jd} + \varepsilon_{ijdPRE} \; , \tag{1}$$

where $t = PRE$ denotes the pre-intervention period, T_{jd} is a dummy indicating whether health facility j is in the treatment group and ε_{ijdPRE} is the error term. In regressions using household data, the disturbance term is adjusted to allow for correlations within catchment areas (communities).

To estimate the causal effect of the program, we then run the same regression in the post-period ($t = POST$):

$$y_{ijdPOST} = \alpha_{POST} + \beta_{POST}T_{jd} + \varepsilon_{ijdPOST} \; . \tag{2}$$

We also estimate an extended version of equation (2):

$$y_{ijdPOST} = \alpha'_{POST} + \beta'_{POST}T_{jd} + X_{ijdPOST}\pi + \theta_d + \varepsilon_{ijdPOST} \; . \tag{3}$$

Specification (3) includes district fixed effects (θ_d) and facility and household variables (X) controlling for pre-treatment differences across health facilities and communities that were present despite randomization. This increases the precision of the coefficient estimates.

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:[17]

$$y_{ijt} = \gamma POST_t + \beta_{DD}(T_j * POST_t) + \mu_j + \varepsilon_{ijt}, \tag{4}$$

where $POST$ is a post period dummy, μ_j is a facility specific fixed effect, and β_{DD} is the difference-in-differences estimate (program impact).[18]

7   Results

7.1   Pre-intervention differences

Prior to the intervention, the treatment and the control group were similar on most characteristics. We report the test of difference in means across control and treatment groups in table 1. At the baseline, we do not find any statistically significant differences in utilization (number of outpatient treated and deliveries per month), households' use of different service providers (including drug shops) in case of illness, waiting time, equipment usage, government funding of clinics, citizens' perceptions of staff behavior, catchment area characteristics (such as the number of villages and households in catchment area), distances from the health facility to the nearest local council and government facility, or health facility characteristics (such as type of water source, availability of drinking water at the facility, whether a separate maternity unit is available, electricity shortages). In one out of five measures of monthly supply of drugs (i.e., Quinine), the treatment group, on average, has a marginally higher supply in the year prior to treatment. In one out of four user-charge measures, there is some evidence (the estimate is significant at the 10 percent level) that patients served by the treatment facilities are more likely to pay for service delivery. Overall, though, the randomization appears to have been successful.

---

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

[18]A slightly more restricted difference-in-difference (DD) specification substitutes the facility fixed effects for T_{jd}. In that case, time invariant factors will be captured by T_{jd}. Both DD specifications yield identical point estimates of β_{DD}.

7.2 Processes

The initial phase of the project, i.e., the three separate meetings, followed a pre-design structure. A parallel system (visit by a member of the survey team) also confirmed that this initial phase of the intervention was properly implemented. After these initial meetings, it was up to the community to sustain and lead the process that the intervention intended to initiate. In this section, we present some evidence on this first component in the accountability chain depicted in figure 2; namely if the treatment communities become more involved in monitoring the providers.

To avoid influencing local initiatives, the parallel system was only in place during the first round of meetings. Therefore, we are not able to document all actions taken by the communities in response to the intervention. Still, we have two sources of information on how processes in the community have changed. First, the CBOs submitted reports on what type of changes they observed. This evidence is complemented by facility and household survey data as well as data assembled through a local council survey.

According to the CBO reports, the community-based monitoring process that followed the first set of meetings was a joint effort mainly managed by the local councils, HUMC (Health Unit Management Committee) and community members. In the communities, the performance of the health facility was discussed during village meetings. The local council survey also confirms this. 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 report that concerns raised by the village members were carried forward by the local council to the health 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, while others felt the pressure from the community to act and follow up on the issues covered in the community contract. These claims are also 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. In the control communities, we observe no dissolved HUMC. Further, the CBOs report that the community also monitored the health facility staff during health visits to the clinic, when they rewarded and questioned issues in the community contract which had or had not been addressed, suggesting a more systematic use of non-pecuniary rewards. Tools such as suggestion boxes (where community members could anonymously leave suggestions for change or comment on the lack of change that was supposed to have taken place), numbered waiting cards (to ensure a first-come-first serve basis), and duty roasters, were also reported to be put in place in several treatment facilities.

In table 2, we formally look at the program impact on these processes. We use data collected through visual checks by enumerators during the post-intervention survey.

As reported in table 2 (regressions 1-2), 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 these, while 25 % of the treatment facilities did). A higher share of treatment facilities also post information on free-services and patient's rights and obligations (regressions 3-4). The enumerators could visually confirm that 70 percent (17 out of 25) of the treatment facilities had at least one of these "monitoring tools" (suggestion boxes, numbered waiting cards, and/or posters on free-services), while only 4 out of 25 control units had at least one of them. The difference is statistically highly significant (column 5).

The results based on household data mirror the findings reported in columns 1-5. For example, the performance of the staff is more often discussed in local council meetings in the treatment communities (regression 6), suggesting that the treatment communities became more engaged. Three out of four households surveyed have attended at least one village meeting in 2005. Of those attending, 40 percent (13 percentage points) more households in the treatment community report that the functioning of the health facility was discussed. 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.[19]

7.3 Treatment practices

The qualitative evidence from the CBOs and, to the extent that we can measure it, the findings reported in table 2, confirm that the treatment communities became more involved in monitoring the provider. Did community monitoring affect the health worker's behavior and performance? We turn to this next. We report the results on treatment practices and staff behavior, both as expressed in perception responses by households (in appendix A.3) and in quantitative indicators such as the immunization of children, waiting time, staff absenteeism, examination procedures, management of the clinic, and extent of preventive care.

We start by looking at examination procedures.[20] Regression 1, table 3, presents the result of estimating (4) with the dependent variable being an indicator of whether any equipment (for instance thermometer or blood pressure equipment) was used during the examination. 49 percent of the patients in the treatment community reported that equipment was used the last time the respondent (or the respondent's child)

---

[19]Additional evidence on community engagement and monitoring is reported in appendix A.3.

[20]Naturally, the relevant treatment is conditional on illness and the condition of the patient. However, since the project was randomly allocated across communities, there is no reason to believe that the type of illness and the condition of the patients should differ systematically across groups. In fact, we have information on reported symptoms for which the patient seeks care (from the household survey). There are, on average, no systematic differences in reported symptoms across treatment and control communities.

visited the project clinic, as opposed to only 41 percent in the control group. The difference-in-differences estimate, 8 percentage points or a 19% increase, is highly significant.

In regression 2, table 3, we look at 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 133 minutes in the control facilities and 119 in the treatment facilities. The difference is highly significant.[21]

Table 4, column 1, reports the results on absenteeism.[22] The point estimates suggest 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 employed, is 19 percent (10 percentage points) lower in the treatment facilities. Column 2 presents the result when only using the nominator as the dependent variable. In the treatment facilities, 3.1 workers were present on average as compared to 2.3 in the control clinic. Thus, in response to more extensive community monitoring, health workers are more likely to be at work.

Enumerators also visually checked the condition of the health center, i.e. whether floors and walls were clean, the condition of the furniture and the smell of the facility. Each condition was ranked on a score from 1 (dirty) to 3 (clean). Through principal components analysis, we transform these four variables into a summary score (the first component): "condition of the clinic". There is a large and significant improvement in the treatment clinic. The point estimate implies that treatment clinics, on average, score 0.56 standard deviations (in the sample of control facilities) higher than the control facilities. Thus, treatment clinics appear to have put more effort into keeping the clinic in decent condition in response to the intervention.[23]

The findings on immunization of children under five are reported in tables 5a-5d.[24] We have information on how many times (doses) in total each child has been

---

[21]The point estimates for the treatment effect in table 3 are similar, but somewhat less precisely estimated, when only using data from the post-intervention survey, i.e. when estimating (2) instead of (4).

[22]The post-intervention survey was not announced in advance. At the start of the survey, enumerators physically verified the provider's presence. A worker was counted as absent if, at the time of the visit (during facility hours), he or she was not in the clinic. Staff reported to be on outreach were omitted from the absence calculation. In the full sample, 47 percent of the health workers were absent. Chaudhury et al. (2006), based on a larger sample of both rural and urban health centers in Uganda, report that 37 percent of the workers, on average, are absent.

[23]Improvements in treatment practices are also substantiated by the qualitative data assembled. As reported in the appendix, table A.8., for all three subjective measures (overall change in the quality of services provided over the last year, change in staff politeness, change in availability of medical staff), there are positive and significant differences between the treatment and control communities' responses.

[24]We report results of estimating (2) rather than the difference-in-differences equation (4), since the pre-treatment vaccination outcomes were strongly influenced by a mass immunization campaign implemented prior to the survey period. Due to reported irregularities in the top management of the unit in charge of the immunization campaigns, we have not been able to assemble accurate information on the actual timing of the campaign prior to the intervention.

immunized with polio, DPT, BCG, and measles. To the extent that this is possible, these data were collected from households' immunization cards.

According to the Uganda National Expanded Program on Immunization (UNEPI), 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). To account for these immunization requirements, we create dummy 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. We then estimate (2), using these binary indicators (for measles, DPT, BCG, and polio) as dependent variables for each age group (0-12 months, 13-24 months, 25-36 months, 37-48 months, and 49-60 months). The results are reported in tables 5a-5d.

There are significant positive differences between households in the treatment and control community for all four vaccines, although not for all cohorts. The program impact on measles vaccination is presented in table 5a. Approximately 34 percent of the children under one year have received at least one dose against measles. There is no significant difference between treatment and control groups (regression 1). For one-year old children (13-24 months), however, we find a significant difference (regression 2). In the control group, 79 percent of the children have been immunized, while the corresponding number in the treatment group is 5.6 percentage points higher. A smaller, but significant, difference also shows up in the cohort of three-year old children (37-48 months) (regression 4). Table 5b reports the results on immunization against polio. There are positive and significant differences in all but the oldest age group (regressions 6-9). The difference is largest for the youngest cohort (9.5 percentage points). For DPT, in table 5c, we find a significant positive difference in two out of five cohorts and for BCG, in table 5d, we find a positive and significant difference (7.8 percentage points) for the youngest cohort (regression 1).

According to the government health sector strategic plan, preventive care is one of the core tasks for health providers at the primary level. Although we did not collect data on households' knowledge about health and various preventive measures, we have data on to what extent households have been informed about the potential dangers of self-treatment and if they have received information about family planning. Table 6 shows that a significantly larger share of households in the treatment communities have received information about the dangers of self-treatment (regression 1), and the importance of family planning (regression 2). The difference is 9 and 7 percentage points, respectively.[25]

---

[25]As a reference point, the share of households that have received information about the dangers of self-treatment and the importance of family planning are 32 percent and 30 percent, respectively, in the control communities, implying a 28% and 23% increase in health knowledge.

7.4 Utilization

The evidence presented so far shows that treatment communities began to monitor the health unit more extensively in response to the intervention and that in light of better community monitoring, the health unit staff responded by improving the provision of health services. We now turn to the question of whether increased community monitoring also resulted in improved quantity and quality of care.

Tables 7 and 8 report estimates of the treatment effect on quantity. We collected detailed data from the health facilities on the number of out-patients, the number of deliveries, the number of antenatal care patients, and the number of people seeking family planning services.[26]

Table 7 presents the results, for the four different utilization variables, from the estimations of equations (2) and (4). There are positive and significant differences between treatment and control facilities across all four services. One year into the program, utilization (for general outpatient services) is 16 percent higher in the treatment facilities. When controlling for district fixed effects, the point estimate is slightly larger and more precisely estimated (significant at the 1 percent level). The difference in the number of deliveries at the facility (albeit starting from a low level) is even larger (68 percent, regression 4) and fairly precisely estimated. There are also positive and significant differences in the number of patients seeking antenatal care (22 percent, regression 8) and family planning (60 percent, regression 10).

As a complement to the difference approach, columns 3 and 6 present the results from the estimation of a value added specification.[27] Difference-in-differences estimates, i.e., equation (4), are reported in table A.10 in the appendix.[28] The point estimates from both specifications are positive and highly significant. The difference-in-differences estimates also suggest a larger treatment effect (28%) for outpatient services.

Table 8 reports changes in utilization patterns based on household data. We collected data on where each household member sought care during the last year in case of illness that required treatment. Apart from recording visits to the project facility (treatment or control facility), we recorded visits to private providers (for-profit and NGOs), traditional healers, self-treatment (i.e., purchases of medicine in drug shops), or other government facilities (i.e., not a project facility). Consistent with the findings reported in table 7, we find a positive and significant difference in the use of the project facility between treatment and control facilities (regression 1). The increase,

---

[26]As discussed in section 5, these data were assembled by counting the number of patients from daily patient records, maternity unit records, the antenatal care register, and the family planning register.

[27]Data on the number of antenatal care patients and the number of people seeking family planning services were not collected from medical records in the pre-treatment survey.

[28]The value added specification is

$$y_{jdPOST} = \alpha_{VA} + \beta_{VA}T_{jd} + \lambda y_{jdPRE} + \varepsilon_{jdPOST} \ .$$

14 percent higher in the treatment group as compared to the control group, is similar to that reported in table 7 (using facility records).

Table 8 also shows that households in the treatment community reduced the number of visits to traditional healers and the extent of self-treatment (regressions 4 and 5), while there are no statistically significant differences (regressions 2, 3, 6 and 7) across the two groups in the use of other providers (NGO, for-profit, or other government facilities). Thus, households in the treatment communities switched from traditional healers and self-treatment to the project facility in response to the intervention.

7.5 Health outcomes

The main objective of the community-based monitoring project was to improve health outcomes in rural areas of Uganda where health indicators have been stagnating. To achieve this objective, the project intended to enhance communities' abilities to monitor the public health care provider, thereby strengthening providers' incentives to increase both the quality and quantity of primary health care provision. As reported above, the project was successful in raising both utilization and, to the extent that this can be measured, service quality. Next, we turn to health outcomes.

Data on two health outcomes were collected. First, we collected information on whether the household had suffered from the death of a child (under five years) in 2005, i.e., the first year of the community monitoring project. Second, we 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.[29]

Health outcomes (under-five mortality and weight of infants) 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. In the community monitoring project specifically, having patients switching from self-treatment or traditional healers to seeking care at the treatment facility could have an effect. Holding utilization constant, better service quality and increased immunization of children (particular measles) could also result in a reduction in mortality and improved health status. The increased use of preventive care (health education) may also have an effect.

Table 9 presents the results on child mortality. 3.2 percent of the surveyed households in the treatment community had suffered from the death of a child in 2005. The corresponding number in the control community is 4.9 percent. The difference – a 33 percent reduction in child deaths in the treatment communities – is significant and

---

[29]The weighing scale was a regular hanging baby scale with trousers (Salter type). Two trained enumerators assisted in the task. During the weighing process, the enumerators took help from family members, mostly mothers. When the infant/child was hanging calmly on the scale, the enumerators recorded the weight.

fairly precisely estimated when controlling for district fixed effects (regression 2).[30,31] With a total of approximately 55,000 households residing in the treatment communities, the treatment effect (0.017) corresponds to 546 averted under-five deaths in the treatment group in 2005.[32]

The dependent variable in regression 3, table 9, is under-five mortality rate in the community.[33] Consistent with the findings in columns 1-2, the point estimate suggest a substantial treatment effect. The average under-five mortality rate in the control group is 145, close to the official figure of 133 for 2005 (UNICEF, 2006). In the treatment group, the under-five mortality rate is 97 and the difference is significant at the 5-percent level.[34]

The program impact on the weight of infants is reported in table 10. Growth charts for boys and girls are depicted in figure 4. As in Cortinovis et al's (1997) study of over 4,000 children from 31 villages in Mbarara (a district in south-western Uganda), we find that Ugandan infants have values of weight far lower than the NCHS/CDC international reference. The gap increases for older infants. The median weight of six-month old boys in the sample is close to the 25th percentile of the NCHS/CDC reference chart. For the 18 months old, the median weight for boys lies close to the 10the percentile of the NCHS/CDC chart.

Figure 5 plots the distribution of weight-for-age (z score).[35] A population similar to the reference population (NCHS) will have a mean z score of zero, with approximately 2.5 percent of the population below a z score of -2 (the threshold for moderately underweight). In the sample of measured infants, 17.4 percent fall below this threshold.

---

[30]The numbers on child deaths are comparable to other survey based measures on child mortality in Uganda. In a sample of 1178 children under the age of five from north-western Uganda (from both urban and rural villages), Vella et al (1992) find a mortality rate (percent of children who died during the last year) of 3.9 percent. Mortality rates were around 10% during the first year of life, 3.1% in the second year, 4.0% in the third year, and about 0.5% thereafter.

[31]The treatment effect reported in table 9 is quantitatively important, even as compared to medical field trials where infant mortality is a measured outcome. For example, of the 23 measures (i.e. biological agent or action intended to reduce child mortality) for which Jones et al. (2003) conclude that there is sufficient or limited evidence of effect on child mortality, the mean effect was a 37% reduction in infant mortality.

[32]We get an almost identical estimate (540 averted deaths) when we weight with distance to the health facility. Since villages closer to the facility were oversampled, the sample of treatment villages is not fully representative of the total population in the treatment communities.

[33]The under-five mortality rate is estimated as the number deaths of children under five in the community as a fraction of number of live births in 2005 (i.e. number of infants younger than one year at the end of 2005 plus the number of infants under one year that died in 2005) expressed per 1,000 live births.

[34]To put this into perspective, an under-five mortality rate of 97 implies that child mortality in the treatment group is in parity with that of Ghana - a country with a 50 percent higher GDP per capita (PPP US$) in 2001 (UNDP, 2002).

[35]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. Four observations (out of 1142) with z scores < -4.5 were consequently dropped.

8.5 percent of the infants (up to 18 months) are severely underweight (< -3 z scores). Almost a quarter of the infants falls below the mildly underweight threshold (< -1 z score).

The difference in means of z scores of infants between the treatment and the control group is reported in regression 1, table 10. The estimated effect (difference) is 0.164 z score in weight-for-age. Regression 2 applies a more stringent restriction on the data to avoid problems of misreporting.[36] The difference in mean is 0.17 z score and is precisely estimated. Figure 6 plots the distribution of z scores for treatment and control groups. The difference in measured weight is most apparent for underweight children. Underweight status causes a decrease in immune and non-immune host defenses. Thus, since underweight children are at a higher risk of suffering from infectious diseases (and more severe complications of infectious diseases), and therefore in higher demand for/need of health care, the data in figure 6 are consistent with a positive treatment effect arising from improved access and quality of health care, rather than a general increase in nutritional status.

Regression 3 adds district fixed effects and controls for age and gender. The results remain qualitatively unchanged. The incidence of underweight increases with age. We cannot reject the hypothesis that the treatment effect is the same for girls and boys.

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 0.17 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 8 percent (figure 6).[37]

Columns 4-5 in table 10 report the program impact on child weight for children between 18-36 months of age. The treatment effect is small and insignificant.[38]

---

[36]Specifically, we drop 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.

[37]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). If the present coverage level were increased to universal coverage (99%), Jones et al estimate that complementary feeding alone would prevent 6% of the under-five deaths in the 42 countries with the 90% of worldwide child deaths in 2000. According to Jones et al, this is one of the most effective (in the sense of preventing under-five deaths) preventive interventions feasible for delivery at high coverage in a low-income setting.

[38]Measurement errors due to misreported age of the child are likely to be a more serious concern for children above 18 months than for infants.

7.6 Robustness

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. For example, if a treatment facility improved the quality of health provision due to the intervention, households in villages in the catchment area of a control community might choose to seek service in the treatment facility. If this is the case, we would overestimate the effects (on utilization) of the intervention. Naturally, it is also possible that community members in the control facilities copied the monitoring approach of the treatment facilities, in which case the bias would go in the opposite direction.

In practice, there are reasons to believe that this is not a serious concern. First, the average (and median) distance between the treatment and control facility is 30 kilometers. Second, 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. One way of testing for spillover effects is to estimate an augmented version of (2) for the sample of control facilities.[39] That is, we estimate

$$y_{idPOST} = \alpha + \eta DIST_{id} + \varepsilon_{idPOST}, \tag{5}$$

where $DIST_i$ is the distance (in kilometers) between the control facility i and the closest treatment facility. The results of estimating (5) for the various utilizations measures are reported in table 11. In all specifications, the estimate of η differs insignificantly from zero.

Table 12 reports a difference-in-differences version of (5). Again, the point estimates are insignificantly different from zero..

Another concern, which does not influence the casual effect of the project but the interpretation, 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 tables 13-15 do not provide any evidence of this being the case. Difference-in-differences estimates of the monthly supply of drugs indicate that the treatment and control facilities are similar. If anything, drug supplies are smaller in the treatment clinics (table 13). The treatment facilities did not receive more funding from the sub-district or district (table 12, regression 6) as compared to the control facilities. The difference-in-differences estimate is negative, but insignificant. There are no differences in constructions or infrastructure during the first project year (table 14), and there are no differences in the availability of equipment at the health facility (table 15).

A similar interpretational concern arises if the upper-level authorities increased their supervision and control of treatment facilities in response to the intervention.

---

[39]Pooling the sample of control and treatment facilities and adding a dummy for treatment facilities yields identical results.

However, this does not seem to be the case either. The supervision of providers by upper-level government authorities remained low in both the treatment and the control group (table 16, regressions 1-2).

The incidence of supervision and control visits may be an imprecise measure of the effectiveness of monitoring by the upper-level authorities. A complementary measure is implemented sanctions. We have data on the extent to which staff was dismissed or transferred during the first year of the project. As noted in section 4, only the District Service Commission has the authority to dismiss and transfer staff. There is only a handful of staff that has been dismissed or transferred in 2005 and there is no systematic pattern that distinguishes treatment from control facilities (table 16, regressions 3-4). Likewise, there is no difference between treatment and control facilities in the number of staff that voluntarily left the facility during 2005 (regression 5).

Taken together, these findings reinforce our confidence that the improved quality and quantity of health care provision resulted from increased efforts by the health unit staff to serve the community in light of better community monitoring.

8   Conclusion

In this paper, we have studied the effects of enhancing rural communities' ability to hold primary health care providers accountable. We find that both the quality and quantity of the health service provision improved in the treatment communities: One year into the program, average utilization was 16 percent higher in the treatment communities; the weight of infants higher, and the number of deaths among children under five markedly lower. Treatment communities became more extensively involved in monitoring the providers following the intervention and the results suggest that the health unit staff responded by exerting a higher effort in serving the community. By strengthening the providers' incentives to serve the poor, health provision and in the end health outcomes can be significantly improved.

The starting point of this work is the mounting evidence showing that the provision of public services to poor people in developing countries is constrained by weak incentives of service providers. As argued in Chaudhury et al. (2006), this evidence is symptomatic of failures in "street-level" institutions and governance. However, although these failures constitute a direct obstacle to economic and social development, they have, until recently, received much less attention in the literature than weaknesses in macro institutions. This paper is an attempt at partly closing this gap.

Although the Citizen report card project appears to be successful, it is too early to use these findings as a basis for continued or increased support and funding for various activities with the aim of strengthening beneficiary control. There are still a number of outstanding issues. One important concern is to what extent the processes initiated by the Citizen report card project are permanent. Since the project is ongoing and scaled up to involve an additional 25 project units, this question can be answered at a

later stage. At the same time, it is possible that the treatment communities' ability to coordinate citizen actions has also been applied to other areas of concern (education, local road construction, etc.), in which case the aggregate return is even larger than what the above results suggest. It is also possible that even better results can be achieved by combining bottom-up monitoring with a top-down approach (supervision and possibly sanctions/rewards from someone in the institutional hierarchy assigned to monitor and control the primary health care providers).

Before scaling up, it is also important to subject the project to a cost-benefit analysis and relate the cost-benefit outcomes to other possible interventions. 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. The intervention resulted in 1.7 percentage points fewer child deaths in the treatment communities during the first project year. To the extent that this number is representative of the total treatment population, this would imply that approximately 550 under-five deaths were averted as a result of the intervention. A back-of-the-envelope calculation then suggests that the intervention, 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 in the Citizen report card project. This can be compared to the numbers reported by Filmer and Pritchett (1999). They contrast the cost of averting the death of a child derived from increasing public expenditures on health (regression estimates range from $47,112 to $100,927), to more conventional health interventions based on cost-effectiveness estimates of the minimum required cost to avert a death (ranges from $1,000 to $10,000 for diarrheal diseases, from $379 to $1,610 for acute respiratory infection, $78 to $990 for malaria, and $836-$3,967 for complications of pregnancy).[40]

The Citizen report card project was implemented in nine different districts of Uganda and reached approximately 55,000 households. Thus, in this dimension, the project has already shown that it can be brought to scale. Still, this project is a controlled experiment in some dimension. Specifically, data collection and data analyses were supervised by the evaluators. To the extent that these tasks were delegated to local actors in the various communities, they could have been subject to capture. This is an issue on which our findings do not shed any light. What our findings strongly suggest, though, is that experimentation and evaluation of new tools to enhance accountability should be an integral part of the research agenda on improving the outcomes of social services. This is an area where at present, research on what works and what does not is lagging behind policy.

---

[40] These numbers should be viewed with caution. Naturally, the 95 percent confidence interval would also include a much smaller estimate of program impact than the 1.7 percentage points used here. Moreover, since the largest cost item was the collection of data and these data were partly used in the intervention and partly to evaluate impact, the cost is a rough estimate. Filmer and Pritchett's (1999) estimates of the cost of averting a child death derived from increasing public expenditures on health are subject to a variety of estimation problems and the health interventions based cost-effectiveness estimates of the minimum required cost to avert a death are, as noted by Filmer and Pritchett, at best suggestive.

References

Appleton, Simon (2001), "The Rich Are Just Like Us, Only Richer': Poverty Functions or Consumption Functions?", *Journal of African Economies* 10(4): 433-469.

Banerjee, Abhijit and Esther Duflo (2005), "Addressing Absence", *Journal of Economic Perspectives* 20 (1): 117-132.

Banerjee, Abhijit, Angus Deaton and Esther Duflo (2004), "Wealth, Health, and Health Service Delivery in Rural Rajasthan", *American Economic Review Papers and Proceedings* 94(2): 326-330.

Banerjee, Abhijit, and Ruimin He (2003), "The World Bank of the Future", *American Economic Review* 93(2): 39-44.

Besley, Timothy, and Robin Burgess (2002), "The Political Economy of Government Responsiveness: Theory and Evidence From India", *The Quarterly Journal of Economics* 177(4):1415-51.

Besley, Timothy, and Andrea Prat (2005), "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability", *American Economic Review*, forthcoming.

Björkman, Martina (2006), "Does Money Matter for Student Performance? Evidence from a Grant Program in Uganda", Working Paper, IIES, Stockholm University.

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

Esther, Duflo and Rema Hanna (2005), "Monitoring Works: Getting Teachers to Come to School", Working Paper, Department of Economics and Poverty Action Lab, MIT.

Filmer, Dean and Lant Pritchett (1999), "The Impact of Public Spending on Health: Does Money Matter?", *Social Science and Medicine* 49(10).

Jimenez, Emmanuel and Yasuyuki Sawada (1999), "Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program", *The World Bank Economic Review* 13(3): 415-441.

Khemani, Stuti (2006), "Can Information Campaigns Overcome Political Obstacles to Serving the Poor," mimeo, The World Bank.

Malena, Carmen, Reiner Forster and Janmejay Singh (2004), "Social Accountability: An Introduction to the Concept and Emerging Practice", Social Development Papers 76, Participation and Civic Engagement Group, The World Bank.

McPake, Barbara, Delius Asiimwe, Francis Mwesigye, Mathias Ofumbi, Lisbeth Ortenblad, Pieter Streefland and Asaph Turinde (1999), "The Economic Behavior of Health

Workers in Uganda: Implications for Quality and Accessibility of Public Health Services," *Social Science and Medicine* 49(7): 849-865.

Moeller, Lars Christian (2002), "Uganda and the Millennium Development Goals", Human Development Network, World Bank, Washington D.C, Processed.

Olken, Ben (2005), " Monitoring Corruption: Evidence from a Field Experiment in Indonesia", NBER Working Paper No.11753.

Reinikka, Ritva and Jakob Svensson (2004), "Local Capture: Evidence from a Central Government Transfer Program in Uganda", *The Quarterly Journal of Economics* 119 (2): 679-705.

Reinikka, Ritva and Jakob Svensson (2005a), "The Power of Information: Evidence from a Newspaper Campaign to Reduce Capture", Working paper, IIES, Stockholm University.

Reinikka, Ritva and Jakob Svensson (2005b), "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda", *Journal of the European Economic Association* 3 (2-3): 259-267.

Reinikka, Ritva and Jakob Svensson (2005c), "Working for God", Working paper, IIES, Stockholm University.

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

Republic of Uganda (2002), "Infant Mortality in Uganda 1995-2000: Why the Non-Improvement?", Discussion Paper No. 6, Planning and Economic Development, Kampala.

Singh, Janmejay and Meera Shah (2002), "Community Score Cards in rural Malawi", World Bank, Washington, D.C. Processed

Strömberg, David (2003), "Mass Media and Public Policy", *European Economic Review* 45(4-6): 652-63.

Strömberg, David (2004), "Radio's Impact on Public Spending", *The Quarterly Journal of Economics* 119(1): 189-221.

Samuel Paul (2002), *Holding the State to Account: Citizen Monitoring in Action*, Books for Change, Bangalore.

UNICEF (2006), *The State of the World's Children 2007: Women and Children - The double dividend of gender equality*, UNICEF.

UNDP (2002), *Millennium Development Goals: A compact among nations to end human poverty*, Human Development Report 2003, UNDP.

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

World Bank (1996), "World Bank Participation Sourcebook", Environmental Department papers No. 19, Washington.

World Health Organization (2006), *World Health Statistics 2006*, WHO Press, Switzerland.

A Appendix

A.1 Sampling Strategy

The starting point for the sample frame is the QSDS data set for 2000 and the second round of QSDS data for 2004 (Reinikka and Svensson, 2005c). The QSDS data set consists of a total of 155 health facilities. The sample design for the QSDS was governed by three principles. First, the attention was restricted to dispensaries (i.e., health centre III) to ensure a degree of homogeneity across sampled facilities. Second, subject to security constraints, the sample was meant to capture regional differences. Finally, the sample had to include facilities from the main ownership categories: government, private non-profit, and private for-profit providers. These three considerations led to the choice of a stratified random sample, which was based on the Ministry of Health facility register for 1999. The register includes government, private non-profit, and private for-profit health facilities, but is known to be inaccurate with respect to the latter two. A total of 155 health facilities were sampled. On the basis of existing information on private-for profit and non-profit, it was decided that the sample would include 81 government facilities, 44 private non-for-profit facilities, and 30 private for-profit facilities. As a first step in the sampling process, 8 districts (out of 45) had to be dropped from the sample frame due to security concerns.[41] From the remaining districts, 10 districts, stratified according to geographical location, were randomly sampled in proportion to district population size. Thus, three districts were chosen from the Eastern and Central regions, and two from the Western and North regions.

A.1.1 Part 1: Sampling of Villages

Our initial sample frame for the household survey thus consists of 81 government facilities and their "catchment" areas. The catchment area of a facility is operationalized as the five-kilometer radius around the facility. For different reasons, all these facilities/catchment areas could not be included in the sample. First, three government facilities in Soroti could not be surveyed in the second round of the QSDS due to security concerns. Second, detailed maps (covering at least the five-kilometer radius around the facility) and the corresponding census data could not be collected for three units.[42] Third, for some facilities, a significant part of the catchment area lies outside the facilities' administrative boundaries. These facilities/catchment areas were there-

---

[41] The eight districts were Bundibugyo, Gulu, Kabarole, Kasese, Kibaale, Kitgum, Kotido, and Moroto.

[42] Uganda Administrative Maps from the Cartography Department at the Uganda Bureau of Statistics. These maps are drawn with the sub-county level as the highest administrative unit and the village as the smallest unit. The maps were drawn in September 2001 (some earlier) as a preparation for the 2001/2002 Census.

fore dropped from the sample.[43] Finally, five districts had been split since the initial survey; Kaberamaido previously part of Soroti, Kayunga previously part of Mukono, Mayuge previously part of Iganga, Sironko previously part of Mbale, and Wakiso previously part of Mpigi. As a result, for some districts, we end up with too few facilities. The districts with too few (less than four) facilities were therefore dropped. Altogether, we end up with a sample of 50 government facilities/catchment areas (CA).

Combining information on geographical location (from the detailed maps provided by Uganda Bureau of Statistics (UBOS)) and census data, we could list all villages and enumeration areas and their size (number of households) for each catchment area (CA). Summary data on the number of villages in CA are provided in Tables A.1-A.3. Altogether, there are 804 enumeration areas, covering 1,194 villages and 109,296 households in the 50 CAs. On average, a CA consists of 20 enumeration areas and 29 villages, half of which are outside the 3 km radius. The average (median) village has 92 (84) households.

Three general principles governed our choice of sample. First, we wanted our sample of households to be representative of the potential users of the facility in the CA. This, in turn, is a function of both the size of the population in the CA and the distance to the facility. Second, for the intervention to be feasible (and within our budget constraint), we wanted to restrict the number of villages to be surveyed within a given CA. For the same reason, we wanted to ensure that the villages surveyed are clustered together in a smaller set of clusters within each CA. Finally, we wanted to include the village where the facility was located (typically the village where the staff resides).

To ensure this, we chose a four-stage sampling design. First, we determined how many villages should be selected from each CA. Balancing the need of being representative of the potential users of the facility in the CA and designing a financially and logistically feasible survey strategy, the "village rule" was set to

$$ no. \ villages = 3.3 + 0.1 * (no. \ villages \ in \ CA). \tag{6} $$

Second, we determined the share of these villages that should be sampled from the one, three, and five kilometer radius (strata 1, 3, and 5), i.e., the "strata rule".[44] For each CA, these shares were set so as to replicate the shares of villages in the different strata in the CA, with one exception. Since households in villages closer to the facility, everything else equal, are more likely to visit the facility, we oversampled the villages from the one-kilometer radius by a factor of 2 and undersampled the share of facilities within the five-kilometer radius (excluding the facilities within the three-kilometer radius) by a factor of 0.7.

Third, to ensure that the villages surveyed are clustered together and that the

---

[43]Specifically, we dropped facilities/catchment areas where more than 25, [33] or {50} percent of the catchment area were outside the 1 [3] {5} km radius.

[44]Strata 1 is defined as the area within the one-kilometer area; strata 3 is defined as the area within the three-kilometer area excluding the area within the one-kilometer area; strata 5 is defined as the area within the five-kilometer area excluding the area within the three-kilometer area.

village where the facility is located is included in the sample, we first identified the enumeration areas (EA) of the village where the facility is located and second, we selected an additional 2-4 EAs within each CA, with a probability proportional to population size. The number of EAs selected was determined by (6).[45]

Finally, within the sampled EAs, we randomly selected the stipulated number of villages in the 1, 3, and 5 kilometer strata in the CA.

The total and the average number of villages sampled according to the sampling strategy and the actual number of villages surveyed are depicted in Table A.4.[46]

Summary statistics of the sample of villages surveyed are depicted in Table A.5 and Table A.6. Overall, 293 villages were surveyed (from 242 EAs) with a total population of 29,405. The average village in the sample has 102 households, slightly larger than the average village in the sample frame.

A.1.2 Part 2: Sampling of Households in Selected Villages

Using the most updated census data, we enumerated all 293 villages included in the final sample and coded them. Two codes were created; one unique code for each household in each village (HHSVC), and one unique code for each household in the whole sample of households in the 293 villages (HHSSC). Then, we determined the number of households that should be surveyed in each village (SHHS). The rule was set as follows:

SHHS	Condition
10	if total no. of households in village $\in [20, 50]$
0.2*(no. hhs in village)	if total no. of households in village $\in [50, 100]$
20	if total no. of households in village $\in [100, 200]$
25	if total no. of households in village > 200

This resulted in a total sample of 4,978 households to be surveyed in the final sample. The sample design to select the households to be surveyed from the set of eligible households (i.e., the enumerated households) is as follows. First, a random number between 1-10 (between 1-5 in villages with less than 100 households) was drawn. This number is denoted "START" and is the first household selected. Let the last number in the village list of households (HHSVC) be denoted by "LNO". Then, the remaining (SHHS-1) sampled households are determined by selecting every xth (denoted "EVERY") household, starting from START up to the point where the total number of sampled households is equal to SHHS. The variable EVERY is defined as the maximum integer such that

---

[45]That is, enough EAs were chosen so that the stipulated number of villages in the 1, 3, and 5 kilometer radius could be surveyed.

[46]Four villages were dropped due to too few households residing in the village (less than 20 households). We also had to replace a handful of villages where enumeration was not possible. This accounts for the difference between the sample rule and the actual sample.

$$EVERY = (max\,[integer \leq LNO] - START)/(SHHS - 1) \qquad (7)$$

Intuitively, we determined EVERY such that the sequence of households to be sampled is evenly distributed over the list of households in the village, i.e. evenly distributed over HHSVC.[47]

A replacement strategy was also designed. The replacements are selected as follows. If a selected household with HHSVC code x could not be surveyed, the household with HHSVC code x+1 should be selected. If that is not feasible (because there is no x+1 household or because that household could not be interviewed, or because that household has already been interviewed), the household with HHSVC code x-1 should be selected. If that is not feasible, the household with HHSVC x+2 should be selected, and thereafter x-2, etc.

A.1.3   Ex-post Survey

The same sample of health facilities, villages and households that were sampled and surveyed in 2004 were re-surveyed in the ex-post survey at the beginning of 2006, and there was likely to be cases where the previously surveyed household could not be interviewed for some reason (i.e. the household had moved or died etc.), a replacement strategy was designed. The replacements were selected as follows. If a selected household with HHSVC code x could not be surveyed, pick the household residing to the right of household x. If that is not feasible (because there is no household to the right or because that household could not be interviewed either, or because that household has already been interviewed), pick the household residing to the left of household x. If that is not feasible, pick the household residing two houses to the right of household x, and then two houses to the left of household x, etc.

In total, 4,996 households were surveyed in the ex-post survey, 4,373 of which were resurveyed.

---

[47]Denote LAST as the last household in the list to be surveyed (i.e. the sampled household with the highest HHSVC). Then LAST = START + EVERY*((SHHS -1).

Table A.1. Total number of households, villages and enumeration areas in sample frame (50 units).

	Total	Within 1 km radius	Within 3 km radius excl. those within the 1 km radius	Within 5 km radius excl. those within the 3 km radius
Households	109,296	11,572	41,665	56,059
Villages	1,194	113	458	623
Enumeration areas	804			

Source: UBOS maps and census data

Table A.2. Number of households, villages and enumeration areas in sample frame (50 units)

	Mean	Median	Min	Max
Households in catchment area	2,483	2,728	490	3,938
Households within 1 km radius in CA	344	240	60	1014
Households within 3 km radius excl. those within the 1 km radius in CA	1096	991	127	2,357
Households within 5 km radius excl. those within the 1 and 3 km radius in CA	1,303	1,231	173	2,428
Villages in catchment area	29	26	7	58
Villages within 1 km radius	3	3	1	8
Villages within 3 km radius excl. those within the 1 km radius in CA	13	11	2	30
Villages within 5 km radius excl. those within the 1 and 3 km radius in CA	15	15	2	31
Enumeration areas in catchment area	20	19	4	35
Villages in enumeration area	1.9	2	0	6

Source: UBOS maps and census data.

Table A.3. Village characteristics in sample frame (50 units).

	Mean	Median	Min	Max
Number of households in village	92	84	0	273
Distance to facility	3.9	5	1	5

Source: UBOS maps and census data

Table A.4. Sampled villages according to village and strata rules and actual sample (50 units).

	According to village/strata rule	Sample
Villages (total)	295	293
Villages, average in CA	6	6
Villages in strata 1, total	64	70
Villages in strata 1, average in CA	1	2
Villages in strata 3, total	117	121
Villages in strata 3, average in CA	2	3
Villages in strata 5, total	114	102
Villages in strata 5, average in CA	2	2

Source: UBOS maps and census data.

Table A.5. Total number of households, villages and enumeration areas in actual sample

	Total	Within 1 km radius	Within 3 km radius excl. those within the 1 km radius	Within 5 km radius excl. those within the 3 km radius
Households	29,405	7,696	11,653	10,056
Villages	293	70	121	102
Enumeration areas	242			

Table A.6. Village characteristics of sample.

	Mean	Median	Min	Max
Number of households in village	102	92	22	232
Distance to facility	3.2	3	1	5

A.2 Participatory Methods

The report card was disseminated to the community using a Participatory Rural Appraisal (PRA) methodology. In the early 1990s, the participatory rural appraisal methodology was mainly used by non-government organizations in East-Africa and South-Asia but are today widely used in many different organizations all over the world.[48] Participatory rural appraisal evolved from a set of informal techniques used by development practitioners in rural areas to collect and analyze data. It emphasizes local knowledge and the importance of having beneficiaries making their own appraisal, analysis, and plans for monitoring and evaluation of service providers. It is a participatory process intended to mitigate the collective action problem by facilitating the analysis of people's environment and identification and discussion of problems. The method employs a wide range of tools and techniques such as maps, diagrams, role-plays and action planning. Next, we briefly describe the specific tools used in the Citizen Report Card project in Uganda.

Venn diagrams were used to discuss power issues in service delivery. Participants were asked to list the different stakeholders in health service delivery (i.e. health facility staff, citizens, health management committee, district officials etc). Thereafter, the participants discussed the different roles and responsibilities of these players in ensuring the quality of the service, i.e. who is accountable to whom; what is a particular stakeholder accountable for, and how can these actors account for their actions. The outcome was used in the interface meeting to identify the stakeholders who have the power to ensure that quality service is delivered. The outcome also contributed to the process of developing a shared vision of how to monitor the provider.

Focus group discussions were used to generate discussions among and across sub-groups. Participants were divided into key social groups such as women, men, youths, disabled, local leaders and elderly in order to get their perspectives over issues concerning service delivery and determine their preferences for change. Each group individually discussed the issues covered in the report card and recorded suggestions for improvements. Thereafter, each group presented the results to the other participants by using flip charts. In this way, the voice and priorities of all social groups were taken into considerations.

"Now, Soon, Later" approach is a technique aimed at helping the community identify issues they would like to address in the short term and those they would address in the longer term, considering the resource envelope at hand. Thereafter, the participants were asked to prioritize the needs according to their resource envelope and discuss which factors are important and necessary for making a change. This tool was intended to help the community analyze the resources available, the time frame for implementing the desired change and the importance of the issue.

Role play was used to illustrate community and health facility interactions as perceived by the respective parties. This tool facilitated the discussion and dialogue in the interface meeting between health staff and community members. The story of the play illustrated the participants' interpretation of an ordinary day at the health facil-

---

[48]See World Bank (1996).

ity. In the play, community members were asked to act the roles of health facility staff (In-charge; Mid-wife; Records Assistant; Watch Man; Laboratory Assistant; Senior Nurse etc) and health facility staff acted the roles of users of the facility (pregnant women; patients; poor patients; community leader; Chairman). Role plays are viewed as an effective tool for diffusing sensitive issues (such as absenteeism or weak attention of staff). It is also a tool that can be used to illustrate constraints and opportunities, enabling users and providers to forge a way forward. Not only did the role play focus on the current situation at the health facility but in a second role play, the plot exemplified what the participants would like the situation to be like in six months.

Roles and Responsibility Analysis is used to provide clarity as to who is responsible for what activity. In this analysis, the participants review all planned activities in the action plan and ensure that each activity becomes someone's responsibility. This tool define roles and responsibilities and helps strengthening the relationship of accountability between health service providers and citizens with regard to the activities determined in the action plan. The facilitator guides the participants to discuss the activities recorded in the action plan and help them agree on the criteria for taking up a responsibility for a particular activity. Thereafter, the participants identify who among the community or health facility staff would suit the criteria and discuss this responsibility with the person or group identified. The groups or individuals assigned to be responsible for a certain activity are then recorded in the action plan.

Action planning was a tool used in the final stage to summarize and record the community's suggestions for improvements (and how to reach them without additional resources). The action plan states the health issues/services that had been identified by the community and the staff as the most important to address; how these issues could be addressed; when they are supposed to be achieved; by whom this will be done; and how the community could monitor the improvements (or the lack thereof). The action plan is a contract between the community and the health facility. It forms the basis for local monitoring and makes it easier for the community to keep track of the implementation of agreed recommendations.

A.3 Additional results

A.3.1 Processes

Table A.7. reports additional findings on changes in processes at the community level following the intervention. As reported in column 1, community members in the treatment group are better informed about patients' rights and obligations according to the government set standard for health service delivery at the primary level.[49] The treatment communities are also more likely (although most households do not know this) to

---

[49]These data are based on simple knowledge tests administered to households. Specifically, respondents were asked to list the main "rights" (right to confidential treatment, right to polite treatment according to first come-first serve basis, right to receive information on ailment and drugs, free health care, attended with one hour) according to the Yellow Star Program (see section 5.2). The dependent

know when the project facility receives drug deliveries (regression 2) and significantly more likely to have been informed about the HUMC's role and responsibilities.

Table A.7. Program impact on processes: performance of staff discussed in village meeting and information about patients' rights

Dependent variable	Informed about patients' rights	Informed about drug deliveries	Informed about HUMC's role & responsibilities
Specification	(1)	(2)	(3)
Program impact	0.03*	0.03**	0.05***
	(0.02)	(0.01)	(0.01)
Mean in control group	0.34	0.11	0.08
District fixed effects	Yes	Yes	Yes
Observations	4996	4996	4996
R^2	0.02	0.06	0.05

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable in specifications: (1) Dummy variable indicating if the household could list at least one of the rights according to the Yellow Start program, (2) Dummy variable indicating if the household knows when the health facility receives drugs, (3) Dummy variable indicating if the household has been informed about HUMC's role and responsibilities.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

A.3.2 Treatment practices

Table A.8. reports treatment effects based on households' perception of the quality of service delivery at project facilities. Although these estimates constitute causal effects of the community monitoring project, there are several reasons why they should be interpreted with care. For all three subjective measures (overall change in the quality of services provided over the last year, change in staff politeness, change in availability of medical staff), there are positive and significant differences between the treatment and control communities' responses. Most households in the control communities (53 %) perceive that the quality of services provided at the project facility has become worse or not improved during the last year. In the treatment communities, on the other hand, a majority (564 %) [ngt fel på denna siffra!!!!] of the households surveyed report that the quality of services provided at the project facility has improved. The difference is significant and precisely estimated once controlling for district fixed effects (regression 1). We find similar patterns in households' perceptions of the politeness of staff and the availability of medical staff when visiting the clinic (regressions 2 and 3 in table A.8).[50]

---

variable (table 3, specification 2) takes the value of 1 if the respondent could list at least one of these rights and zero otherwise. We find a positive and significant effect (treatment effect) on both the extensive and intensive margin (not reported), i.e.; more informed respondents and conditional on being informed, better knowledge about patients' rights following the intervention.

[50]We find similar effects and of the same magnitude (positive and significant) using ratings on the attention given to the patient by the staff when visiting the project facility and whether the patient

Table A.8. Citizens' perception of changes in the quality of health care over the last year

Dependent variable	Overall quality	Staff politeness	Availability of medical staff
Specification	(1)	(2)	(3)
Program impact	0.09**	0.08**	0.09***
	(0.04)	(0.03)	(0.03)
Mean in control group	0.47	0.53	0.51
Controls	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes
Observations	3343	3343	3343
R^2	0.09	0.05	0.06

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable in specifications: (1) Dummy variable indicating improvement in overall quality; (2) Dummy variable indicating improvement in staff politeness; (3) Dummy variable indicating improvement in availability of medical staff.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.
d. Control variables include: Distance to nearest local council from the health facility, distance to other government health facilities in the area and electricity at the health facility.

As of 2001, public health services are free of charge. However, the survey evidence indicates that patients still encounter varying costs, although a large majority of patients do not pay (informal) user fees. In the pre-treatment data, 7 percent of the households surveyed reported having to pay user charges for outpatient services; approximately 15 percent had to pay for injections (when needed); and 67 percent paid for delivery.[51]

In table A.9, we report the program impact on these informal charges. The intervention had no significant effect on the share of households that needed to pay for drugs (regression 1) or delivery (regression 4). However, it had an impact on general outpatient services (regression 2) as well as on injections (regression 3).

---

felt he/she was free to express herself when being examined.

[51] Average payment (for those that had to pay) was UGX 1,435 (USD 0.80) for out-patient service, UGX 370 (USD 0.21) for injections, and UGX 4,955 (USD 2.75) for delivery.

Table A.9. Difference-in-difference estimates of the program impact on user charges at the health facility.

Dependent variable	Drugs	General treatment	Injections	Delivery
Specification	(1)	(2)	(3)	(4)
Program impact	-0.01	-0.06*	-0.14**	-0.07
(Treatment*2005)	(0.01)	(0.029)	(0.07)	(0.11)
2005	0.002	-0.018**	0.11**	-0.13*
	(0.005)	(0.007)	(0.04)	(0.07)
Mean control group 2005	0.01	0.02	0.37	0.50
Facility fixed effects	Yes	Yes	Yes	Yes
Observations	5660	5734	2511	507
R^2	0.003	0.18	0.27	0.42

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Specification: (1)-(4) Dummy variables indicating whether the health facility charged for service provided during the last visit.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

A.3.3 Utilization

The difference-in-differences estimates on the number of outpatients and deliveries are reported in table A.10. For the number of outpatients, we present the results from estimations of difference-in-differences specifications in both levels and logarithms.

The treatment effect is positive and significantly different from zero for both outpatients served and the number of deliveries. The point estimates in the out-patient specifications suggest a substantial treatment effect.

Table A.10. Difference-in-differences estimates of the program impact on health facility utilization.

Dependent variable	Out-Patient	Log of Out-Patient	Delivery
Specification	(1)	(2)	(3)
Program impact	215.5**	0.28**	3.48*
(Treatment*2005)	(93.4)	(0.11)	(1.98)
2005	-247.3	-0.25***	1.73
	(70.1)	(0.07)	(0.89)
Mean control group in 2005	661		9.2
Facility fixed effects	Yes	Yes	Yes
Observations	100	100	100
R^2	0.77	0.82	0.90

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Robust standard errors in parenthesis.

Table 1. Average health facility and citizen characteristics, pre-treatment.

	Treatment group	Control group	Difference
Utilization:			
Out-patient care	587	908	-51
			(141)
Delivery	10.32	7.48	2.84
			(2.61)
Utilization pattern:			
Project facility	0.31	0.34	-0.03
			(0.03)
NGO health facility	0.02	0.02	-0.002
			(0.003)
Private-for-Profit health facility	0.24	0.26	-0.02
			(0.01)
Traditional healer	0.034	0.03	0.004
			(0.007)
Self treatment (drug shop)	0.36	0.32	0.04
			(0.03)
Other government health facility	0.18	0.17	0.01
			(0.05)
Other provider	0.014	0.007	0.007
			(0.005)
Quality measures:			
Waiting time	148	144	4.3
			(4.2)
Equipment usage	0.47	0.48	-0.01
			(0.02)
Funding at the facility:			
1000 shillings	4766	3429	1337
			(905)

The results are catchment area (health facility) averages. Robust standard errors in parentheses. Significantly different from zero at 99 (***), 95 (**), and 90 (*) percent confidence. Description of variables: Utilization variables are the average number of patients visiting the health facility per month; Utilization pattern is the citizens' use of different service providers in case of illness (reported in percentages); Waiting time is calculated as the difference between the time the citizen left the facility and the time the citizen arrived at the facility minus the examination time; Equipment usage is a dummy variable indicating whether the staff used any equipment during examination; Funding at the health facility is the average funds received at the health facility per month from the district and the Health Sub-district (measured in 1000 shillings).

Table 1 continued. Average health facility and citizen characteristics, pre-treatment.

	Treatment group	Control group	Difference
Catchment area statistics:			
Number of villages per health facility	23.2	24.6	-1.3
			(3.14)
Number of villages per health facility in strata 1	2.6	1.8	0.80*
			(0.45)
Number of villages per health facility in strata 3	8.9	9.5	-0.64
			(1.7)
Number of villages per health facility in strata 5	11.7	13.2	-1.5
			(1.69)
Number of households per health facility	2140	2224	-84
			(275)
Number of households per village	93.9	95.4	-1.42
			(8.2)
Health facility characteristics:			
Piped water	0.04	0.04	0
			(0.00)
Rain tank/Open well	0.52	0.36	0.16
			(0.14)
Borehole	0.44	0.60	-0.16
			(0.14)
Drinking water	1.76	1.48	0.28
			(0.20)
Separate maternity unit	0.16	0.16	0
			(0.00)
Distance to nearest Local Council I	0.72	0.85	-0.13
			(0.26)
Distance to nearest public health provider	8.68	7.76	0.92
			(1.90)
Number of days without electricity in **the** last month	18.3	20.4	-2.12
			(4.14)

The results are catchment area (health facility) averages. Robust standard errors in parentheses. Significantly different from zero at 99 (***), 95 (**), and 90 (*) percent confidence. Description of variables: Catchment area statistics are determined from UBOS maps and census data; Piped water, Rain tank and Borehole are dummy variables indicating the health facility's watersource; Drinking water is a dummy variable indicating whether the health facility has drinking water available; Separate maternity unit is a dummy variable indicating whether the health facility has a separate maternity unit; Distance to nearest Local Council I and distance to nearest public health provider is measured in kilometers; Number of days without electricity in the last month is measured out of 31 days.

Table 1 continued. Average health facility and citizen characteristics, pre-treatment.

	Treatment group	Control group	Difference
Citizen perceptions:			
Polite behavior	3.06	3.02	0.04
			(0.04)
Attention	3.17	3.16	0.01
			(0.03)
Free to express	3.8	3.77	0.03
			(0.02)
Citizens' informations about drug deliveries	0.14	0.16	-0.02
			(0.05)
Supply of drug deliveries at the health facility:			
Erythromycin	420	346	74
			(131)
Chloroquine	3410	2915	495
			(567)
Septrine	2690	2430	260
			(623)
Quinine	573	335	238*
			(130)
Mebendazole	1597	1500	97
			(230)
User charges:			
Drugs	0.024	0.011	0.013
			(0.012)
General treatment	0.10	0.03	0.07*
			(0.04)
Delivery	0.50	0.58	0.08
			(0.10)
Injection	0.24	0.20	0.04
			(0.06)

The results are catchment area (health facility) averages. Robust standard errors in parentheses. Significantly different from zero at 99 (***), 95 (**), and 90 (*) percent confidence. Description of variables: Citizen's perceptions describes his/her experience during the last visit at the health facility and are measured on a scale from 1 to 4 where a higher value represents higher satisfaction; Citizen's information about drug deliveries is a dummy variable indicating if the citizen knows when the health facility receives drugs from the district/Health Sub-district; Supply of drug deliveries per month is measured as the average number of tablets received at the health facility per month from the district/Health Sub-district; User charges are a dummy variable indicating if the household had to pay for the service provided at the health facility.

Table 2. Program impact on monitoring tools at the health facility.

Dependent variable	Suggestion box	Numbered waiting cards	Poster informing of free services	Poster on patients' rights and obligations	At least one monitoring tool	Discuss the health facility in LC meetings
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Program impact	0.38***	0.20**	0.19**	0.12	0.56**	0.13***
	(0.10)	(0.10)	(0.09)	(0.10)	(0.11)	(0.02)
Mean control group	0	0.04	0.12	0.12	0.16	0.33
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	50	50	50	50	50	3119
R²	0.35	0.30	0.47	0.26	0.70	0.11

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Dependent variables in specifications (1)-(5) are based on data collected through visual checks by the enumerators: (1) Dummy variable indicating if the health facility has a suggestion box for complaints and recommendations; 2) Dummy variable indicating if the health facility has numbered waiting cards for its patients; (3) Dummy variable indicating if the health facility has a poster informing about free health services; (4) Dummy variable indicating if the health facility has a poster on patients' rights and obligations; (5) Dummy variable indicating if the health facility has a least one of the "monitoring tools" (suggestion boxes, numbered waiting cards, posters on free-services), (6) Dummy variable indicating if the household discusses the functioning of the health facility at Local council meetings.

c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas in regression 6.

44

Table 3. Difference-in-difference estimates of the program impact on treatment practices at the health facility.

Dependent variable	Equipment usage	Waiting time
Specification	(1)	(2)
Program impact	0.08**	-14.0*
(Treatment*2005)	(0.03)	(7.7)
2005	-0.07***	-9.6*
	(0.02)	(5.3)
Mean control group 2005	0.41	133
Facility fixed effects	Yes	Yes
Observations	5280	5148
R^2	0.15	0.12

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Specification: (1) Dummy variable indicated whether the staff used any equipment during examination when the citizen visited the health facility; (2) Waiting time is calculated as the difference between the time the citizen left the facility and the time the citizen arrived at the facility minus the examination time.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

Table 4. Program impact on management

Dependent variable	Absence rate	Staff present	Condition of clinic
Specification	(1)	(2)	(3)
Program impact	-0.10*	0.78*	1.13***
	(0.058)	(0.46)	(0.31)
Mean control group	0.53	2.3	-0.52
District fixed effects	Yes	Yes	Yes
Observations	50	50	50
R^2	0.35	0.30	0.47

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Specification: (1) Absence rate is the ratio of workers not physically present at the time of the post-intervention survey to the number of workers employed; (2) Staff present is the number of workers (verified by the enumerators) to be present at the time of the (surprise) post-intervention survey; (3) Condition of clinic is the first component from a principal components analysis of the variables "condition of the floors of the health center", "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.
c. Robust standard errors in parenthesis.

Table 5a. Program impact on measles immunization of children.

Dependent variable	Measles					Measles				
Age	0	1	2	3	4	0	1	2	3	4
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Program impact	-0.027	0.056*	0.008	0.032**	0.006	-0.037	0.040*	0.004	0.033***	-0.004
	(0.05)	(0.029)	(0.023)	(0.014)	(0.019)	(0.04)	(0.023)	(0.020)	(0.011)	(0.016)
Constant	0.35***	0.79***	0.88***	0.91***	0.94***					
	(0.04)	(0.02)	(0.02)	(0.011)	(0.01)					
Controls	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	558	1022	1035	1229	563	558	1022	1035	1229	563
R^2	0.001	0.005	0.002	0.004	0.001	0.03	0.03	0.01	0.01	0.02

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

c. Age: 0 is below 12 months; 1 is between 13-24 months; 2 is between 25-36 months; 3 is between 37-48 months; 4 is between 49-60 months..

d. Control variables: see note (d) in Table A.8.

47

Table 5b. Program impact on polio immunization of children.

Dependent variable	Polio					Polio				
Age	0	1	2	3	4	0	1	2	3	4
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Program impact	0.082	0.053*	0.034	0.042*	0.024	0.095**	0.047*	0.030*	0.037**	0.017
	(0.051)	(0.031)	(0.026)	(0.023)	(0.028)	(0.039)	(0.026)	(0.017)	(0.019)	(0.029)
Constant	0.59***	0.79***	0.84***	0.87***	0.89***					
	(0.04)	(0.02)	(0.02)	(0.018)	(0.020)					
Controls	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	560	1010	1021	1213	570	560	1010	1021	1213	570
R^2	0.007	0.005	0.002	0.004	0.002	0.07	0.03	0.04	0.02	0.02

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

c. Age: 0 is below 12 months; 1 is between 13-24 months; 2 is between 25-36 months; 3 is between 37-48 months; 4 is between 49-60 months..

d. Control variables: see note (d) in Table A.8.

48

Table 5c. Program impact on DPT immunization of children.

Dependent variable	DPT					DPT				
Age	0	1	2	3	4	0	1	2	3	4
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Program impact	0.017	0.056	0.023	0.081**	0.011	0.033	0.051**	0.027	0.070***	-0.001
	(0.050)	(0.044)	(0.041)	(0.039)	(0.037)	(0.042)	(0.025)	(0.022)	(0.023)	(0.031)
Constant	0.56***	0.73***	0.78***	0.79***	0.87***					
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)					
Controls	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	561	1016	1037	1230	566	561	1016	1037	1230	566
R^2	0.001	0.004	0.001	0.011	0.001	0.05	0.10	0.08	0.08	0.03

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

c. Age: 0 is below 12 months; 1 is between 13-24 months; 2 is between 25-36 months; 3 is between 37-48 months; 4 is between 49-60 months..

d. Control variables: see note (d) in Table A.8.

Table 5d. Program impact on BCG immunization of children.

Dependent variable	BCG					BCG				
Age	0	1	2	3	4	0	1	2	3	4
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Program impact	0.078**	0.006	-0.007	0.021	0.013	0.066**	0.007	-0.008	0.015	0.006
	(0.035)	(0.018)	(0.014)	(0.013)	(0.019)	(0.031)	(0.015)	(0.011)	(0.010)	(0.016)
Constant	0.76***	0.92***	0.94***	0.93***	0.95***					
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)					
Controls	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
District fixed effects	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	970	1011	1029	1212	556	970	1011	1029	1212	556
R^2	0.01	0.000	0.000	0.002	0.001	0.04	0.01	0.01	0.02	0.02

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

c. Age: 0 is below 12 months; 1 is between 13-24 months; 2 is between 25-36 months; 3 is between 37-48 months; 4 is between 49-60 months..

d. Control variables: see note (d) in Table A.8.

Table 6. Program impact on citizens' information.

Dependent variable	Health information	Importance of family planning
Specification	(1)	(2)
Program impact	0.09***	0.07***
	(0.02)	(0.02)
District fixed effects	Yes	Yes
Observations	4996	4996
R^2	0.16	0.10

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable in specifications: (1) Dummy variable indicating if the household receives information about the importance of visiting the health facility and the danger of self-treatment, (2) Dummy variable indicating if the household receives information about family planning.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

Table 7. Program impact on health facility utilization

Dependent variable	Out-Patient			Delivery			Antenatal		Family Planning	
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Program impact	107.4*	137.0**	123.3***	6.3*	7.3**	3.45*	16.1	17.4*	5.5	9.1**
	(65.0)	(54.4)	(47.8)	(3.3)	(2.5)	(2.03)	(16.1)	(9.6)	(4.9)	(4.1)
Pre-utilization			0.31***			1.01***				
			(0.05)			(0.09)				
Constant	660.8***			9.2***			78.9***		15.2***	
	(35.0)			(1.6)			(11.9)		(3.5)	
Controls	No	Yes	No	No	Yes	No	No	Yes	No	Yes
District fixed effects	No	Yes	No	No	Yes	No	No	Yes	No	Yes
Observations	50	50	50	50	50	50	50	50	50	50
R^2	0.05	0.40	0.49	0.07	0.61	0.66	0.02	0.65	0.03	0.44

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Robust standard errors in parenthesis.

c. Control variables include: type of watersource at the health facility, availability of drinking water at the health facility, number of villages in catchment area and whether the health facility has a separate maternity unit.

Table 8. Difference-in-differences estimates of the program impact on citizens' health seeking pattern.

Dependent variable	Project facility	NGO	Private-for-profit	Traditional healer	Self-treatment	Other government health facility	Other
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Program impact (Treatment*2005)	0.034* (0.019)	-0.004 (0.006)	0.03 (0.02)	-0.013* (0.008)	-0.03* (0.02)	-0.005 (0.05)	-0.01 (0.02)
2005	-0.08*** (0.013)	0.008 (0.005)	-0.02 (0.014)	-0.002 (0.005)	0.03** (0.01)	-0.043** (0.02)	0.05*** (0.01)
Mean control group 2005	0.26	0.03	0.25	0.028	0.37	0.12	0.06
Facility fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9200	9200	9200	9200	9200	9200	9200
R^2	0.08	0.09	0.04	0.03	0.09	0.02	0.08

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.
c. Dependent variable is citizens' use of different service providers in case of illness (reported in percentages).

Table 9. Program impact on health outcomes: Under-five child deaths.

Dependent variable	Child death (children < 5 year)		Under-5 mortality rate	
Specification	(1)	(2)	(3)	(4)
Program impact	-0.016*	-0.017**	-48.0**	-46.0*
	(0.01)	(0.009)	(24.2)	(25.8)
Constant	0.049***		144.9***	
	(0.006)		(16.9)	
Controls	No	Yes	No	Yes
District fixed effects	No	Yes	No	Yes
Observations	2922	2922	50	50
R^2	0.002	0.01	0.08	0.18

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Dependent variable (columns 1-2) is a dummy variable indicating whether any children under five in the household have died during the last year and estimated under-5 mortality rate in the community expressed per 1,000 live births (columns 3-4).

c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

d. Control variables: see note (d) in Table A.8.

Table 10. Program impact on health outcomes: Child weight of infants (weight-for-age z-scores).

Dependent variable	weight-for-age z-scores				
	1-18 months			19-36 months	
Specification	(1)	(2)	(3)	(4)	(5)
Program impact	0.16*	0.17**	0.15**	0.012	0.04
	(0.09)	(0.08)	(0.07)	(0.09)	(0.06)
Child age (log)			-1.28***		0.07
			(0.07)		(0.17)
Female			0.26***		0.08
			(0.09)		(0.06)
Constant	-0.64***	-0.71***		-0.95***	
	(0.07)	(0.06)		(0.08)	
Controls	No	No	Yes	No	Yes
District fixed effects	No	No	Yes	No	Yes
Observations	1167	1135	1135	1300	1300
R^2	0.002	0.003	0.22	0.00	0.04

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Dependent variable is weight-for-age z-scores.

c. Specification: (1) Includes all children under 18 months, (2) Includes all children under 18 months with observations with recorded weight above the 90th percentile in the growth chart reported in Cortinovis et al (1997) dropped. (3) Includes all children under 18 months plus controls. (4) Includes all children between 18 and 36 months plus controls. (5) Includes all children between 18 and 36 months plus controls.

d. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

e. Control variables: see note (d) in Table A.8.

Table 11. Robustness test: The effect on utilization at the control facilities when controlling for proximity to project facility.

Dependent variable	Out-Patient Services	Delivery	Family planning	Antenatal care
Specification	(1)	(2)	(3)	(4)
Distance to nearest project facility	-1.13	-0.10	0.07	-0.56
	(2.11)	(0.07)	(0.22)	(0.52)
Constant	696***	12.4***	13*	96***
	(66)	(2.75)	(7)	(18.1)
District fixed effects	No	No	No	No
Observations	25	25	25	25
R^2	0.02	0.06	0.01	0.03

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Robust standard errors in parenthesis.

Table 12. Robustness test: Difference-in-difference estimates on the effect on utilization at the control facilities when controlling for proximity to project facility.

Dependent variable	Out-Patient Services	Delivery
Specification	(1)	(2)
Distance to closest project	-3.41	0.04
facility in 2005	(4.72)	(0.04)
2005	-142.1	0.54
	(154.0)	(1.3)
Facility fixed effects	Yes	Yes
Observations	50	50
R^2	0.75	0.91

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Robust standard errors in parenthesis.

Table 13. Difference-in-difference estimates of drugs supply and funding received

Dependent variable	Erythromycin	Chlorquine	Septrine	Quinine	Mebendazole	Funding
Specification	(1)	(2)	(3)	(4)	(5)	(6)
Program impact	151.8	-176	-2.9	-237	114	-248
(Treatment*2005)	(127)	(688)	(682)	(154)	(533)	(1224)
2005	-145**	-531	-457	-30	984**	1261
	(65)	(415)	(415)	(101)	(412)	(965)
Mean control group 2005	257	2384	1973	305	2483	4471
Facility fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	96	100	100	99	100	94
R^2	0.73	0.55	0.60	0.57	0.66	0.72

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable is the average number of tablets received at the health facility per month from the district and Health Sub-district during the last year (columns 1-5); Average amount of public health care funds received at the health facility per month from the district and Health Sub-district during the last year (measured in 1000 Uganda shillings).
c. Robust standard errors in parenthesis.

Table 14. Program impact on infrastructure at the health facility.

Dependent variable	New units	Toilets	Water source	Electricity
Specification	(1)	(2)	(3)	(4)
Program impact	-0.09	0.09	0.05	0.05
	(0.12)	(0.11)	(0.10)	(0.10)
Controls	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes
Observations	50	50	50	50
R^2	0.50	0.34	0.29	0.42

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Dependent variable is a dummy variable indicating whether any constructions or renovations of infrastructure have been done at the health facility during the last year.

c. Robust standard errors in parenthesis.

d. Control variables: see note (c) in Table 6.

Table 15. Difference-in-difference estimates on equipment at the health facility.

Dependent variable	Bicycles	Examination beds	Blood pressure equipment	Weighing scale	Microscope
Specification	(1)	(2)	(3)	(4)	(5)
Program impact	0.04	0.20	-0.08	0.08	0.22
(Treatment*2005)	(0.19)	(0.25)	(0.19)	(0.11)	(0.15)
2005	0.40***	0.20	0.36**	0.12*	-0.001
	(0.14)	(0.11)	(0.14)	(0.06)	(0.001)
Mean control group 2005	2.92	2	2.04	2.6	0.44
Facility fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	100	100	100	100	100
R^2	0.97	0.84	0.92	0.99	0.99

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Dependent variable is the number of each equipment available at the health facility.

c. Robust standard errors in parenthesis. .

Table 16. Program impact on monitoring of upper-level authorities and dismissals and transfers of staff at the health facility

Dependent variable	Sub-county official	Parish officials	Dismissals	Transferred	Left
Specification	(1)	(2)	(3)	(4)	(5)
Program impact	0.12	0.06	-0.054	-0.083	0.079
	(0.12)	(0.11)	(0.07)	(0.23)	(0.15)
District fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	50	50	50	50	50
R^2	0.43	0.45	0.22	0.16	0.17

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.

b. Dependent variable in specification (1) dummy indicating if the facility has received a monitoring/support visit from any Sub-county officials in 2005; (2) dummy indicating if the facility has received a monitoring/support visit from any Parish officials in 2005; (3) number of staff that has been dismissed in 2005; (4) number of staff that has been transferred from the facility in 2005; (5) number of staff that voluntarily left the facility in 2005.

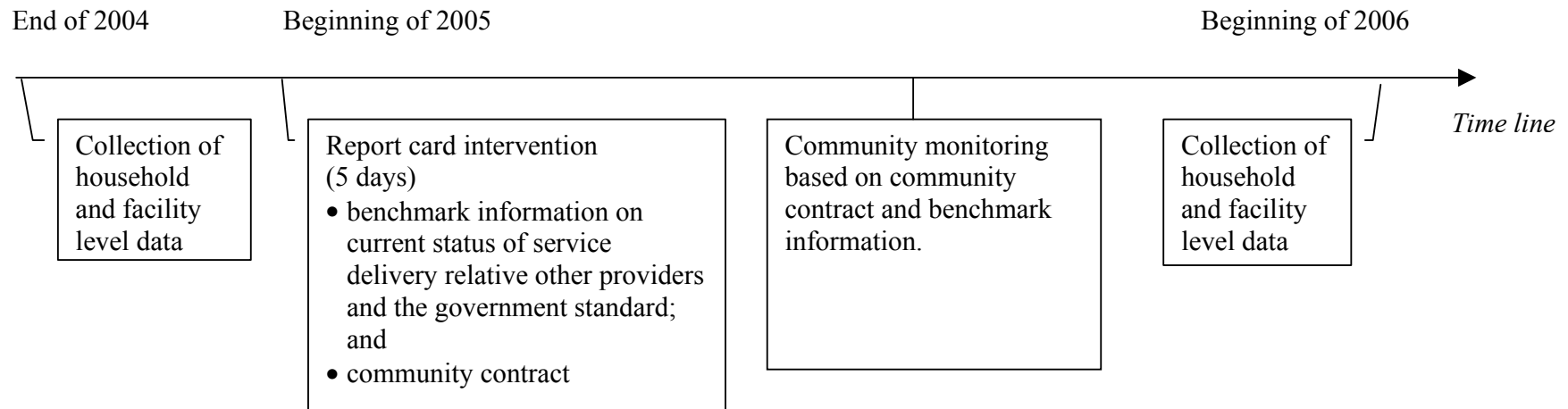c. Robust standard errors in parenthesis.

Figure 1: Timing of project

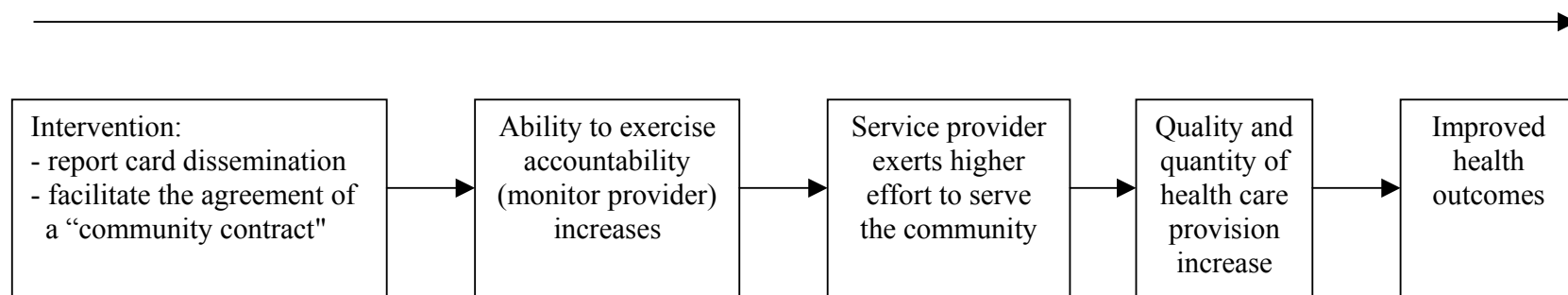End of 2004   Beginning of 2005   Beginning of 2006

Time line

| Collection of household and facility level data | Report card intervention (5 days)
• benchmark information on current status of service delivery relative other providers and the government standard; and
• community contract | Community monitoring based on community contract and benchmark information. | Collection of household and facility level data |

Figure 2: Schematic view of intervention and expected outcome

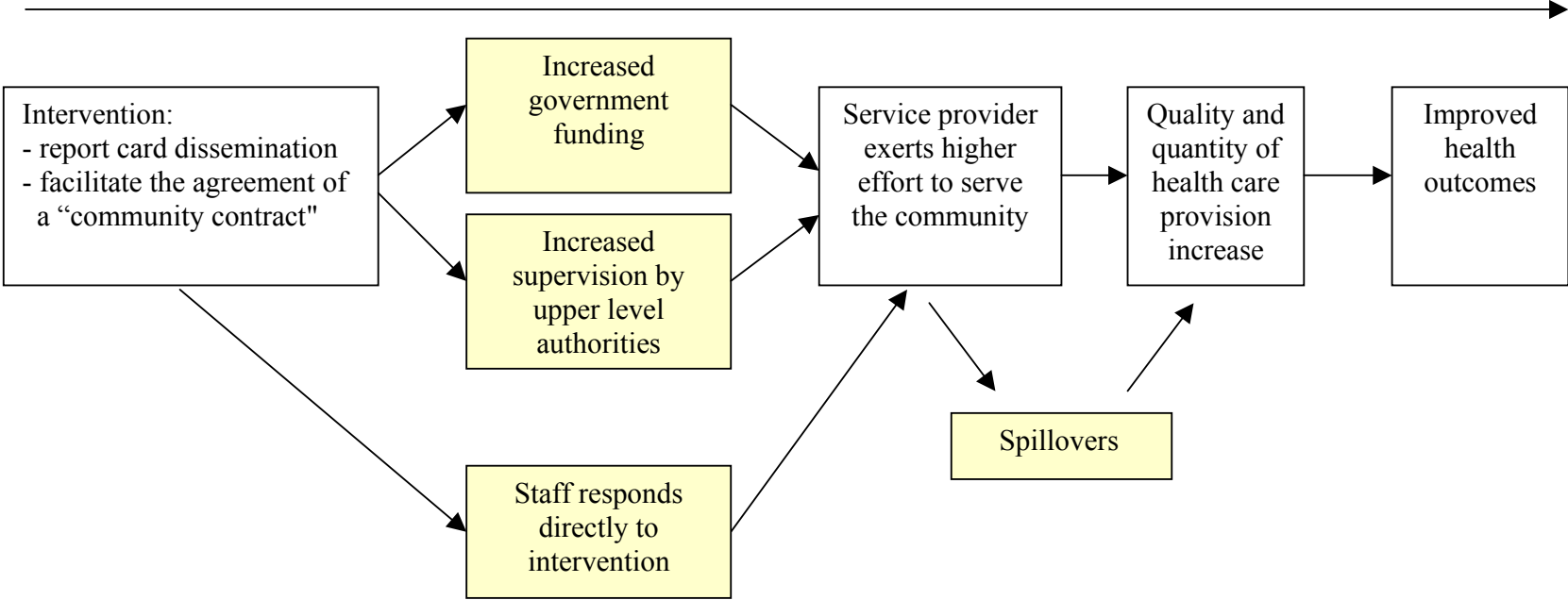| Intervention:
- report card dissemination
- facilitate the agreement of a "community contract" | → | Ability to exercise accountability (monitor provider) increases | → | Service provider exerts higher effort to serve the community | → | Quality and quantity of health care provision increase | → | Improved health outcomes |

Figure 3: Alternative hypotheses
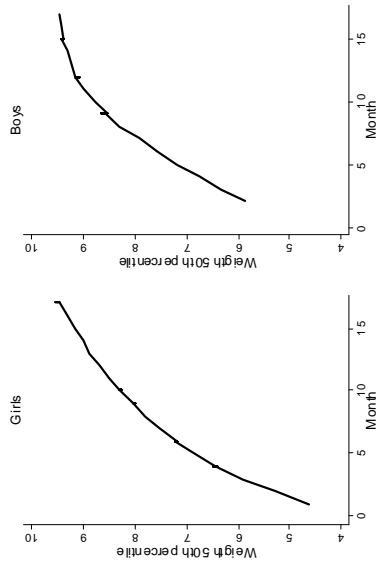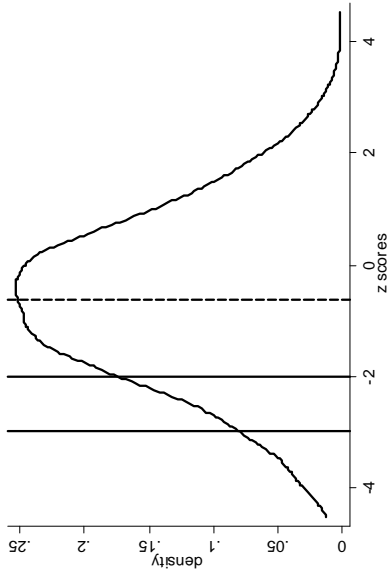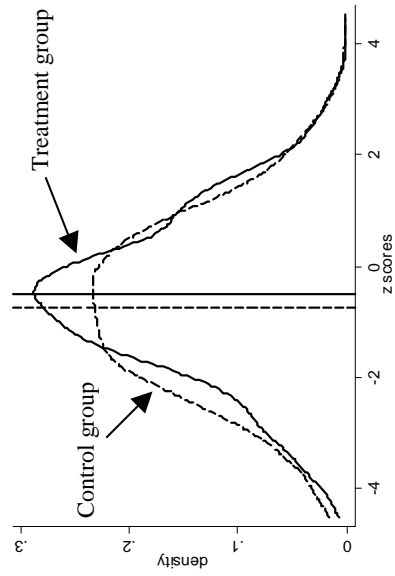
Figure 4: Growth charts

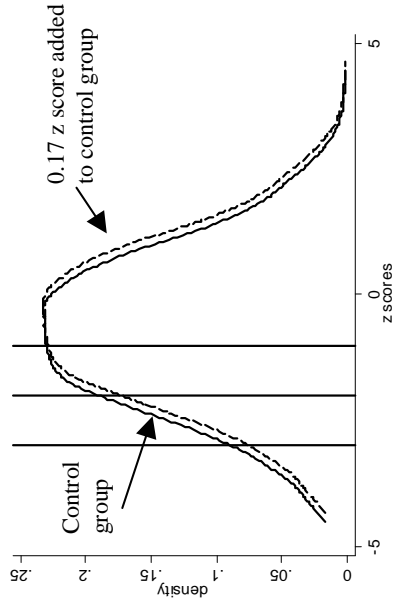**Figure 5:** Distribution of weight-for-age z scores for infants (1-18 months)

Notes: Vertical solid lines denote -3 and -2 z scores, dashed line denotes mean z-score

Figure 6: Distributions of z scores for treatment and control groups

Notes: Vertical solid line denotes mean in treatment group, dashed line denotes mean in control group.

Figure 7: Treatment effect

Notes: Vertical solid lines denote -3, -2, -1 z scores.

64