Power to the People:

Evidence from a Randomized Experiment of a Citizen Report Card Project in Uganda*

Martina Björkman† and Jakob Svensson‡

Draft: September 20, 2006

Abstract

Strengthening the relationship of accountability between health service providers and citizens is viewed by many as critical for improving access to and quality of health care. How this is to be achieved, however, is less understood**.** This paper presents the results of a randomized evaluation of a community-based monitoring intervention (Citizen report cards) intended to enhance rural communities' ability to hold primary health care providers accountable. The Citizen report card project collected quantitative information on the quality and quantity of health service provision from citizens and public health care providers and disseminated this information in ways that create awareness and invoke participation. Both the quality and quantity of 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 providers following the intervention, but we find no evidence of increased government funding. These results suggest that the improvements in the quality and quantity of health service delivery resulted from increased effort by the health unit staff to serve the community.

*This project is a collaborative exercise involving many people. Foremost, we are deeply indebted to Frances Nsonzi and Ritva Reinikka for their contributions at all stages of the project. We would also like to acknowledge the important contributions of Gibwa Kajubi, Abel Ojoo, Anthony Wasswa, James Kanyesigye, Carolyn Winter, Ivo Njosa, Omiat Omongin, Mary Bitekerezo, and the field and data staff with whom we have worked over the years. We thank the Uganda Ministry of Health, Planning Division, the World Bank's Country Office in Uganda, and the Social Development Department, the World Bank, for their cooperation. Finally, we wish to thank the World Bank Research Committee and the Swedish International Development Agency, Department for Research Cooperation for funding this research. Björkman also thanks Jan Wallander's and Tom Hedelius' Research Foundation for funding.

†*IGIER - Bocconi University and CEPR. Email: martina.bjorkman@unibocconi.it.

‡#Institute for International Economic Studies, Stockholm University, NHH, and CEPR. Email: jakob.svensson@iies.su.se.

1 Introduction

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.[1] However, while many agree that strengthening the providers' incentives to serve the poor is crucial for addressing these failures, and in the end improving access to and quality of health care, there is little consensus on how this is to be achieved.

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 the quality of service provision. However, 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.

Partly in response to the failures of these traditional mechanisms of enforceability and answerability, it has been argued that more effort must be placed on strengthening beneficiary control, i.e. strengthen providers accountability to citizen-clients (see e.g., World Bank, 2003). However, despite the enthusiasm for beneficiary control, 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 attempt to provide some.

Empirically, the challenges when establishing whether, and if so which, institutional arrangements can foster a stronger degree of accountability between service providers and citizens re twofold. First, an intervention has to be designed that, if properly implemented, enhances citizens/clients ability to monitor and control the provider. Second, one needs to establish a credible comparison group – a group of observational units (e.g. communities) which would, in the absence of the intervention, have had outcomes similar to those exposed to it.

Our approach to deal with the first challenge is to induce variation in two important elements of the accountability relationship: access to information, and participation and local organization capacity. Improved access to information about the beneficiaries' (as a group) experiences and entitlements is critical for citizens' ability to

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

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 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 also improves users 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. Enhanced participation and local organization capacity is intended to minimize collective action problems and thus to ensure that citizens act on the information being provided.

Citizen report cards is one intervention where these elements take a central focus.[2] A citizen report card is a tool for collecting and disseminating reliable information about how the community at large views the quality and efficacy of service delivery. It also provides the community with an opportunity to compare service delivery in their community vis-à-vis other communities, or across districts in the country at large. The open comparison of outcomes and performance can also provide a spark, or focal point, for public pressure. The citizen report card methodology also emphasizes the active dissemination of information in order to create awareness and invoke participation of the community.

We rely on a randomized design to deal with the second challenge. By randomly assigning communities into a treatment group (i.e. communities in which the project was implemented) and a control group (i.e. communities in which the project was not implemented), we are relatively confident about the absence of confounding factors. In addition, the intervention we evaluate was run on a large scale - - approximately 5,000 households from 50 "communities" from nine districts in Uganda have been surveyed in two rounds, and in total there are approximately 110,000 households residing in the treatment and control communities.[3] This increases our confidence in the external validity of the results.

The community-based monitoring intervention increased the quality and quantity of primary health care provision and resulted in improved health outcomes. One year into the program, utilization (for general outpatient services) was 16 percent higher in the treatment facilities. We also find significant differences in deliveries at the treatment facilities, and in the use of antenatal care and family planning. Treatment practises, as expressed both in perception responses by households and in more quantitative indicators (immunization of children, waiting time, examination procedures)

[2]The best known examples of citizen report cards are probably those developed by the Public Affairs Centre in Bangalore, India (Paul, 2002). Citizens were asked to rate service access and quality and to report on concerns about public services, general grievances, and corruption. The information was summarized in report cards that were reported in the press and in civic forums. Citizen report cards have spread beyond Bangalore to cities in Kenya, Mozambique, the Philippines, Ukraine, and Vietnam. They have been scaled up in India to cover urban and rural services in 24 states. Overall, the citizen report card methodology has stimulated considerable media and political attention and, despite any scientific evidence to back this up, there is general acknowledgment in policy circles of their positive contribution to service improvements (see e.g. World Bank, 2003).

[3]A "community" is operationalized as the households (and villages) residing in the five-kilometer radius around the facility; see section 5.

improved significantly in the treatment communities. We find a small but significant difference in the weight of infants and a markedly lower number of deaths among children under-five in the treatment communities. No effect is found on investments or financial or in-kind support (from the government), suggesting that the changes in the quality and quantity of health care provision are due to behavioral changes of the staff. Moreover, we also find evidence that the treatment clinics started sharing information about treatment practises, availability of drugs, and service delivery in general, in response to the intervention and that the treatment communities began to monitor the health unit more extensively. 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 light of better community monitoring.

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. However, despite the enthusiasm for community participation and monitoring, there is little credible evidence on the impact of policy interventions aimed at achieving these. On the one hand, most (all) comprehensive community based monitoring initiatives have not been rigorously evaluated. On the other hand, the few (no) 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. [4] Thus, it is unclear to what extent non-elite community members were 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

[4]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 each road 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.

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 in 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, having one informed 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 as compared to students in traditional schools. EDUCO schools are managed autonomously by community education associations whose elected members are parents of the students. The community education associations are responsible for hiring (and firing) teachers, closely monitoring teachers' performance, and equipping and maintaining the schools. The results show 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 manage to obtain the causal treatment effect.

There is a growing empirical literature on the relationship between information dissemination (through the media) and accountability. With few exceptions (see below), 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. 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 on the probability of the incumbent winning the election of making information about corruption in the local governments public. Besley and Prat (2005) study the interdependence between media and government accountability, but focus on the reverse relationship: how the government can influence what information will be provided. 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 the disaggregated national accounts data. Finally, we identify impact using an experimen-

tal design, rather than exploiting non-experimental data. The source of identification will thus come directly from a randomized experiment.

Reinikka and Svensson (2005a) also study the relationship between information, accountability, outcomes at the provider level. They exploit a newspaper campaign aimed at reducing the capture of public funds by providing schools (parents) with information to monitor local officials' handling of a large education grant program (capitation grant). They find that the newspaper campaign was highly successful. Head teachers in schools closer to a newspaper outlet are 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 project 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 exacting accountability. It is the broad range of actions and mechanisms that citizens, communities and civil society organizations can use to hold public officials and servants accountable (Malena et al., 2004).

In practice, community-based monitoring can take a variety of forms.[5] However, the different innovations share some common features. Specifically, a key ingredient is improved access to information about the beneficiaries' (as a group) experiences and entitlements. Access to such information is viewed as critical for citizens' ability to monitor service providers. However, publicity, or access to information, is not sufficient for community-based monitoring to be effective. Community-based monitoring is subject to possibly large free-riding problems: the community would like to ensure that the staff does its job, but everyone would rather have someone else do the monitoring. Thus, information provision will not work unless there are members of the community who are willing to make use of the information. For this reason,

[5]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).

most community-based monitoring initiatives place a strong emphasis on encouraging active participation by community members as a way of minimizing collective action problems.

In theory, community-based monitoring has at least three advantages. First, 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 staff's behavior than the external agent assigned to supervise the provider. Second, they may have means to punish the provider that are not available to others, such as verbal complaints or social opprobrium (Banerjee and Duflo, 2005). Third, to the extent that the service is valuable to them, they should have strong incentives to monitor and reward or punish the provider – incentives which the external agent assigned to supervise the provider may lack. But there are also potentially large problems associated with community-based monitoring. First, the tasks of assembling information about performance and acting on this information are subject to possibly large free-riding problems: the community would like to ensure that the provider performs, but everyone would rather have someone else collecting information and monitoring performance. Second, beneficiary control is unlikely to work if citizens do not have a high demand for the service or have access to easily available (and affordable) options (private providers). In that case, the expected relative return to monitoring the public provider will be low. Third, the community must also have some direct or indirect way of sanctioning or rewarding the provider (or some higher level arm of the state). Finally, any project, and maybe community-based interventions in particular, may be subject to capture. For example, the elite may corrupt the collection or dissemination of information or may prevent citizens from speaking out or putting pressure on the provider. Thus, in the end, if and to which extent providers accountability to citizen-clients can be strengthen and if so to what extent such an institutional reform improves outcomes is an empirical question.

4 Institutional setting

Uganda, like many newly independent countries in Africa, had a functioning health care system in the early 1960's. Accessibility and affordability were relatively extensive. The 1970's 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. For example, the infant mortality rate stagnated at 88 deaths per 1,000 live births during the latter half of the 1990s (Republic of Uganda 2002, Moeller 2002) and maternal mortality and immunization rates have remained high and stagnant since the late 1990's. 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 and that such costs defer many, especially the poor, from accessing services.

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 focus of this impact evaluation is on the dispensary (level III). Dispensaries are closest to the users and 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). Dispensaries are manned by a clinical officer (who can be a medical doctor). In our sample of facilities, on average, a dispensary was staffed by a clinical officer, 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 most local actor is the Health Unit Management Committee (HUMC), which is 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 (the in-charge) and non-political representatives from the community (elected by the sub-county local council). The HUMC should monitor drugs, 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 matters 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 the health facility staff. In cases of problems at the health facility, the working practice is that the chairperson of HUMC raises the issue with the in-charge. If there is no improvement, the matter should be referred to the Health Sub-district which if it fails, will refer the errand to the Director of District Health Services. 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, as well as the Director of District Health Services, have the authority to reprimand, but not dismiss, health facility staff for indiscipline. Cases of dismissal are reported to the Chief Administrative Officer of the District who will then report such cases to the District Service Commission, which is the appointing authority for the district and has the authority to suspend or dismiss staff.

Civil Based Organizations (CBOs) are also an important actor in the health service delivery system at the local level. CBOs involved in health mainly focus on undertaking health education activities on antenatal care, family planning, HIV/AIDS prevention, etc.

5 The Program: 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 carried out by staff from the World Bank and Stockholm University, in cooperation with a number of Ugandan practitioners, 18 community-based organizations, and the Uganda Ministry of Health, Planning Division. The 50 project facilities (all in rural areas) were drawn from nine districts in Uganda (see the appendix for details). Defining the catchment area (or the "community") of each dispensary as the households (and villages) residing in the five-kilometer radius around the facility, approximately 110,000 households reside in the communities supposedly served by these 50 facilities.

The facilities were first stratified by location (districts) and then by the number of households residing in the catchment areas. From each group, half the units were randomly assigned to the treatment group and the remaining 25 units were assigned to the control group. Hence, within each district, there exist both treatment and control units.

The Citizen report card project had four components: (a) collecting quantitative information from users (citizens) and service providers using micro survey techniques; (b) assembling this information in "easy access report cards"; (c) disseminating the report cards to users and providers in such a way as to create awareness and invoke participation; (d) providing communities with practical information on how to best use the information to monitor and, in the end, improve the quality and quantity of service provision. These components are discussed next.

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 the impact. To meet these objectives, two surveys were implemented: a survey of health care providers and a survey of health care users. These surveys were implemented both prior to the intervention (data from these surveys formed the basis for the intervention) and one year after the project had been initiated.

A quantitative service delivery survey (QSDS) has been used to collect data from the health service providers. The QSDS collected detailed quantitative data on performance and outcomes from the providers. In many respects, a QSDS is similar to a standard firm-level survey. The key difference is that it explicitly recognizes that agents in the service delivery system may have a strong incentive to misreport (or not report) key data. To this end, the data are 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 local government. The former, often

available in a highly disaggregate format, were considered to suffer the least from any incentive problems in record-keeping.

The user/household survey collected quantitative, and some perception based, data on both households' health outcomes and health facility performance. It included indices of performance parameters such as availability, access, reliability, quality and satisfaction. Data were collected on all different services provided by the health facility, i.e. daily out-patient service, family planning, immunization, and antenatal care. 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 the problem of misreporting. The post-intervention household survey also included a shorter module on health outcomes, including data on under-five mortality, and all infants in the surveyed households were weighed.

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. In a typical community, households from six villages were surveyed. 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.[6] The data included in the report cards were identified as key areas subject to improvement and include utilization, services, drugs and 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 the key findings from the surveys in a format accessible to the communities.

The report cards were translated into the main language spoken in the community.[7] To support the non-literate community members, posters were specifically designed and painted by a graphical artist so that otherwise complex information and concepts were easily understood. As the information in the report card was largely statistical, the posters conveyed the principal ideas such as where people go to seek medical care, reasons for this behavior etc.[8]

5.2 Dissemination and participation

The information in the report cards was disseminated to citizens and providers using a "participatory rural appraisal approach".[9] The information dissemination process was

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

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

[8]See the appendix for prototypes of these posters.

[9]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, analy-

facilitated by staff from Community-based Organizations (CBO). These facilitators were perceived to be a good conduit through which the Citizen Report Card project could be delivered, since they were in constant interaction with the communities and had a mandate drawn from a long-term presence on the ground working with the community. In addition, they could easily make follow-up visits and provide support to the communities.[10,11]

The objective of the dissemination process was threefold. First, to allow the community members themselves to analyze and draw conclusions from the summary findings in the report cards. Second, to develop a shared view on how to monitor the provider by discussing and decomposing 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). Third, ensure that the process is not captured by the elite or any other specific sub-group of the community. To this end, a variety of methods were used, including maps, diagrams, role-play, focus group discussions and action planning.[12]

The information dissemination process was conducted in three separate meetings: a community meeting; a staff meeting; and an interface meeting.

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. The invited participants from each village consisted of a selection of representatives from different spectra of society (i.e. young, old, disabled, women, mothers, leaders). The facilitators mobilized the village members by cooperating with Local (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 shared the information in the report card with the community members using the methods detailed in the appendix. In addition to disseminating findings in the report card, the facilitators also presented information on patients' rights and entitlements.[13] At the end of the meeting, the community's

ses, 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).

[10]The CBO facilitators were trained for seven days in data interpretation and dissemination, utilisation of the participatory methodology, and conflict resolution and management. In addition, a trained enumerator recorded the findings from the CBO which facilitated intervention.

[11]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.

[12]See the appendix for a more detailed description of the various methods

[13]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 concensus that the quality of health services had been

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). After this two-day meeting, participants from each village 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 health facility 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, i.e. the report card. 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.

Following the community and the health facility meeting was an interface meeting with participants (chosen at the community meeting) from villages in the catchment area and all health facility staff. The objective of this meeting was to agree on a strategy for improved health care provision, based on the action plan developed in the community meeting and the discussions from the health facility meeting. During the interface meeting, the community and the health facility staff presented and discussed their suggestions for improvements. A role-play was used to disseminate the results from the survey and in this play, the community and the staff took reverse roles. The participants discussed their rights and entitlements and their roles and responsibilities as patients or medical staff. The outcome of this meeting was a joint action plan describing how the staff and the community collectively can best improve service delivery within the existing resource envelope. The plan contained and reflected the community's and the service provider's consensus on what needs to be done, how, when, and by whom. The joint action plan identified how the community was to monitor the provider and a time plan. Copies of the action plan were kept with the community and the health facility to support the following monitoring process.

5.2.1 Follow-up and Repeat Engagement

The Citizen Report Card process involved both follow-up and repeat engagements with the aim of institutionalizing the process. To this end, the facilitators supported the communities 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 months that followed.

After a period of six months, the communities and health facilities were revisited and a mid-term review was conducted. The mid-term review was a repeat engagement on a smaller scale which included a one-day community meeting and a one-day interface

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 and it aimed at tracking the implementation of the action plan, possibly drawing new areas for concern, and coming up with a new set of recommendations for improvement. The action plans made in the earlier intervention were printed on posters. These posters formed the ground for the discussions in the mid-term review. The facilitators presented the information on the printed action plans, followed by focus group discussions on the progress. During the interface meeting, the health facility staff and the community members jointly discussed suggestions on actions for improving or sustaining the progress of the previously determined action plan. In cases where improvements had not been made, new recommendations were agreed upon and noted in the updated action plan and in cases where improvements had been made, suggestions for sustainability were recorded in the plan. The updated action plan was kept with the community and the health facility to assist in the continued work and monitoring process.

6 Evaluation Design

Empirically, the challenge when establishing whether (and if so which) institutional arrangements can foster a stronger degree of accountability between service providers and citizens, is to establish a credible comparison group – a group of observational units (e.g. communities) which would, in the absence of the intervention, have had outcomes similar to those exposed to the intervention. To achieve this, we rely on a randomized design, i.e. facilities were first stratified by location (districts) and then by the number of households residing in the communities. From each group, half the units were randomly assigned to the treatment group and the remaining 25 units were assigned to the control group. Since treatment status was randomly assigned across health units (and their catchment areas), program participation is not correlated in expectation with either the observed or the unobserved health unit or community characteristics.

6.1 Outcomes

The main outcome of interest is whether the intervention resulted in increases in the quantity and quality of health care and thus in the end, to improved health outcomes in the treatment communities.

Another outcome of interest is the extent to which the intervention had an impact on (informal) user-charges, investments (by the clinic), or financial support (from the government). Yet another outcome of interest is whether the intervention changed the way in which the community interacts with the facility and whether mechanisms were implemented to facilitate monitoring and participation by the community.

6.2 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_{jdt} the outcome variable of 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:

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

where $t = PRE$ denotes the pre-intervention period, T_{id} is a dummy indicating whether health facility j is in the treatment group and ε_{jdPRE} is the error term. When using household data, the dependent variable is y_{ijdPRE}, where subscript i denotes household.

The standard errors are differently adjusted for regressions on health facility data and household data. For regressions on health facility data, the disturbance terms are assumed to be independent across districts, but are allowed to be correlated across health facilities within the same district. In regressions using household data, the disturbance term is adjusted to allow for correlations within catchment areas. We also estimate a version of equation (1):

$$y_{jdPRE} = \alpha + \beta T_{jd} + X_{jdPRE}\pi + \theta_d + \varepsilon_{jdPRE} \ . \tag{2}$$

Specification (2) 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.

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

$$y_{jdPOST} = \alpha + \beta T_{jd} + \varepsilon_{jdPOST} \ . \tag{3}$$

As in the control experiment, we estimate (3) both with and without district fixed effects θ_d and the vector of control variables X.

For a subset of variables, we can also stack the pre and post data and explore the difference-in-differences in outcomes,[14] i.e., we estimate:

$$y_{jdt} = \alpha + \gamma T_{jd} + \delta POST + \beta(T_{jd} * POST) + \varepsilon_{jdt}, \tag{4}$$

where $POST$ is a post period dummy and β is the difference-in-differences estimate (program impact).

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

7 Results

7.1 Pre-intervention differences

Prior to the intervention, the treatment and the control group were similar in most characteristics. Thus, the randomization appears to have been successful. As depicted in table 1, there are no statistically significant differences across the two groups in utilization (number of outpatient treated and deliveries per month), use of different service providers (including drug shops) in case of illness, waiting time, equipment usage, government funding, 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.

7.2 Utilization

Tables 2 and 3 present estimates of the effect of the community-based monitoring intervention on the quantity of health care provision. We collected detailed data 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.[15]

We estimate equations (3) and (4) for the four different quantity outcomes. The average treatment effects are reported in table 2. The community monitoring project was to have an impact. The differences in utilization between treatment and control facilities are positive 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 and a small set of facility controls, the treatment effect is also 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 3) and fairly precisely estimated. There are also positive and significant differences in the number of patients seeking antenatal care (20 percent, regression 5) and family planning (63 percent, regression 8).

These findings are reinforced by the difference-in-differences results depicted in table 3. The treatment effect is positive and significantly different from zero for both

[15] 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.

out-patients served and the number of deliveries.[16] The point estimates are also similar to the difference estimates in table 2.

Table 4 reports changes in utilization patterns based on household data. We collected each household member's decision of where to seek care 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 (both 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 tables 2 and 3, we find a positive and significant difference in the use of the project facility between the treatment and control facilities following the intervention (regression 1). The increase, 15 percent higher in the treatment group as compared to the control group, is almost identical to that reported in table 2 (using facility records).

Table 4 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.

7.3 Treatment practises

Measuring the quality of care is difficult. In this sub-section, we report the results on treatment practises, both as expressed in perception responses by households and in more quantitative indicators such as immunization of children, waiting time, and examination procedures. In the following sub-section, we study health outcomes (which is a function both on the quality and quantity of care).

We start by looking at household perception of how service delivery is carried out at the project facility. Although these estimates constitute causal effects of the community monitoring project, there are several reasons why they should be interpreted with care.[17] As reported in table 5, 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 not improved or has become worse. However, in the treatment communities, a majority (54 %) of the households surveyed report that the quality of services provided at the project facility has improved in the first year of the project. The difference is significant and precisely estimated once controlling for district fixed effects (table 5, regression 1). We find similar patterns with respect to perceptions of the politeness of

[16]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.

[17]For example, these perception variables are ordinal indices but here, they are treated as cardinal measures.

staff and the availability of medical staff when visiting the clinic (regressions 2 and 3 in table 5).[18]

Table 6 reports evidence of improvements in treatment practises. There is no easily measured indicator that can be used to evaluate whether and how patients in the project facilities receive better treatment. 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 be systematically different across groups.[19] Regression 1, table 6, shows 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. 59 percent of the patients in the control community reported that no equipment was used the last time the respondent (or the respondent's child) visited the project clinic. This number is lower (50 percent) for the treatment community. The difference-in-differences estimate is highly significant.

In regression 2, table 6, we look at a more indirect measure of quality, the waiting time, defined as the difference between the time the user left the facility and the time the user arrived at the facility minus the examination time. On average, the waiting time was 133 minutes in the control facilities and 117 in the treatment facilities. This difference is significant.

The findings on immunization of children under-five are reported in tables 7a-7d.[20] We have information on how many times (doses) in total each child has been immunized with polio, DPT, BCG, and measles. To the extent that this is possible, these data were collected from immunization cards.

According to the Uganda National Expanded Program on Immunisation (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 first contact with health facility); and polio (three doses, or four if delivery takes place at the facility, at 6 weeks, 10 weeks, 14 weeks). To account for these immunization requirements, we create dummy variables taking the value 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

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

[19]It is possible that, due to the intervention, patients with more severe illnesses seek care at the project facilities in the treatment area and that this, in turn, can have a direct impact on observed treatment practises. However, the evidence does not support this claim. We have information on reported symptoms for which the patient seeks care (from the household survey). There are no systematic differences in reported symptoms across treatment and control communities.

[20]We report results of estimating (3) 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.

(3), 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 7a-7d.

There are significant positive differences between the treatment and the control community for all four vaccines, although not for all cohorts. Program impact on measles vaccination is depict in table 7a. Approximately 40 percent of 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, 83 percent of the children have been immunized, while the corresponding number in the treatment group is 5.2 percentage points higher. A smaller, but significant, difference also shows up in the cohort of three year old children (37-48 months). Table 7b shows the results on immunization against polio, we find positive and significant differences in all but the oldest age group (regressions 6-9). The difference is largest for the youngest cohort (4.7 percent points). This corresponds to a 13 percent increase in the treatment group compared to the control group. For DPT, in table 7c, we find a significant positive difference in two out of five cohorts and for BCG, in table 7d, we find a positive and significant difference (7 percentage points) in the youngest cohort (regression 1).

7.4 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 the 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, quality of services. 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, all infants (i.e. all children under 18 months of age) and children (between 18 and 36 months of age) in the surveyed households were weighed.[21]

Table 8 reports 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.

[21] All infants (aged above one month and below eighteen months) and children (aged between eighteen and thirty-six months) were weighed. The weighing scale was a regular hanging baby scale with trousers (Salter type). Two trained enumerators assisted in the task and during the weighing process, the enumerators took help from family members, mostly mothers, in order to not scare the infant. This made the weighing process less scary for the baby and minimized the difficulties, and hence also the measurement errors, when weighing the infant. When the infant/child was hanging calmly on the scale, the enumerators recorded the weight.

The corresponding number in the control community is 4.9 percent. The difference, as reported in regression 1, is significant and fairly precisely estimated, when controlling for district fixed effects and a small set of facility and household specific variables (regression 2).[22] 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 following the intervention.[23]

The program impact on the weight of infants is reported in table 9. Given the sample size, we pool the data and study the differences in the average weight of infants between 1-18 months of age. As reported in regression 1, the average weight increased by 0.13 kilograms. For the average child (age nine months and weight 7.9 kilograms), this represents a 1.6 percent increase in weight between the treatment and the control group one year into the program. Albeit small, the coefficient is precisely estimated when controlling for district fixed effects (regression 2). Column 3 in table 9 reports the program impact on child weight for children between 18-36 months of age. The results show that the program did not have any effect on older children's weight.

7.5 Informal user charges

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 out patient services; approximately 15 percent had to pay for injections (when needed); and 67 percent paid for delivery.[24]

In table 10, 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 out patient services (regression 2) as well as on injections (regression 3).

7.6 Processes

Although the Citizen report card project was a structured intervention, it left plenty of room for the communities to choose if and how to react to the information being disseminated. In this section, we present some evidence of the initiatives and processes initiated or strengthened as a result of the intervention. The aim of the project was that the report card information was to provide a spark for community action

[22] These data most likely underestimate the number of under-five deaths. An indication of this being the case is that most under-five deaths (36 %) occur within the first week of life (World Health Organization, 2006). In the survey data, 19 % of the deaths occurred in the first month of life.

[23] Note, though, that 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.

[24] 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.

but also provide citizens with hard data through which the health facility could be evaluated/monitored.

To avoid influencing local initiatives, we chose not to have enumerators spending time in the field after the first round of meetings. Therefore, we were 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 changed following the intervention. First, the CBOs involved in disseminating the report card information submitted reports on what type of changes they observed. The evidence from these reports suggests that the project influenced the way in which the providers were being monitored. This evidence is supported by facility and household survey data as well as data assembled through a Local (village) council survey.

According to the CBO reports, the community-based monitoring process that followed the first set of meetings (community, facility and interface meetings) was a joint effort mainly managed by the village local councils, HUMC (Health Unit Management Committee) and the community members. In the communities, the performance of the health facility was discussed during village meetings. The Local Council survey confirms this claim. A typical village had, on average, six village meetings in 2005. In those meetings, 89 percent of the villages discussed issues concerning the project health facility. The main subjects of discussion in the villages concerned the action plan (30 percent of the villages) or parts of it such as behavior of the staff (49 percent of the villages), drug deliveries at the health facility (48 percent of the villages), and that government health services are supposed to be free of charge (68 percent of the villages).

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 was viewed as 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 re-elected, while others felt the pressure from the community to act and follow up on the issues covered in the action plan. Once more, these reports are confirmed in the survey data: more than one third of the HUMCs in the treatment communities were reelected or received new members following the initial intervention. 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 action plan which had or had not been addressed. 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 put in place in several treatment facilities.

In tables 11-13, we formally look at the program impact of these processes. As reported in table 11, one year into the project, treatment facilities are 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). There are also differences between the treatment

and control facilities in the extent to which information is posted on free-services and patient's rights and obligations.[25]

Table 12 shows that households in the treatment communities are better informed about various aspects of service provision following the intervention. A significantly larger number of households have received information about the importance of visiting a health clinic for medial treatment and the dangers of self-treatment (regression 1), and for family planning (regression 2). The treatment community is slightly more likely (although most households do not know this) to know when the project facility receives drug deliveries (regression 3).

There are also differences between the treatment and control group in the extent to which the performance of the staff at the project facility is discussed in Local Council meetings (table 13, regression 1), and whether, one year into the project, community members know anyone who is a member of the HUMC (regression 2) and have knowledge of the HUMC's roles and responsibilities (regression 3).

7.7 Robustness

One concern with the evaluation design, given that within each district there are both treatment and control units, 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.

In practise, there are reasons to believe 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 have 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 (3) for the sample of control facilities.[26] That is, we estimate

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

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 14. In all specifications, the estimate of λ is insignificantly different from zero.

Table 15 reports a difference-in-differences version of (5). Once more, we do not find any impact.

[25] The data in table 11 are collected through visual checks by the enumerators.

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

Another concern, which does not influence the casual effect of the project but the interpretation, is if the district or sub-district management changed its behavior or support in response to the intervention. For example, the Health Sub-district or the District Health Services may have provided additional funding or other support to the treatment facilities. However, the results in tables 16-19 do not provide any evidence of this being the case. In the first year of the project, the treatment facilities did not receive more funding from the sub-district or district (table 16) as compared to the control facilities. The difference-in-differences estimate is in fact even negative. Difference-in-differences estimates of the monthly supply of drugs also indicate that the treatment and control facilities are similar. If anything, drug supplies are smaller in the treatment clinics (table 17). There are no differences in constructions or infrastructure during the first project year (table 18), and with the exception of microscopes, there are no differences in the availability of equipment at the health facility (table 19).

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 the quantity of 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 into 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 are a direct hindrance 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 to partly close 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 process can studied over time. It is also possible that even better results can be achieved by combining bottom-up monitoring (community based 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). The evaluation of such a project is currently underway.

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 (reported in table 8). The intervention resulted in 1.7 percentage points fewer child deaths during the first project year in the treatment communities. To the extent that this number is representative of the total treatment population, this would imply that approximately 500 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 $500 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).[27]

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 outcomes of social services. This is an area where at present, research on what works and what does not work is clearly lagging behind policy.

[27]These numbers should be viewed with extreme caution. For the cost-benefit estimates of the Citizen report card project, it should be noted that the sample is, by construction, not fully representative of the population (since villages closer to the facility were oversampled). 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 used partly 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.

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 Accesibility 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.

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. The sample 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.[28] ¿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.[29] Third, for some facilities, a significant part of the catchment area lies outside the facilities' administrative boundaries. These facilities/catchment areas

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

[29]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 village as the smallest unit. The maps were drawn in September 2001 (some earlier) as a preparation for the 2001/2002 Census.

were therefore dropped from the sample.[30] 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".[31] 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.

[30] 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.

[31] 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.

Third, to ensure that the villages surveyed are clustered together and that the 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).[32]

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.[33]

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 in which the total number of sampled households is equal to SHHS. The variable EVERY is defined as the maximum integer such that

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

[33]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.

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

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.[34]

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. Since it was likely that there would 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.

[34]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 delivered to the community by using a Participatory Rural Appraisal (PRA) methodology which guides the community on how to best use the information in the report cards. 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.[35] 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 enables local people to make their own appraisal, analysis, plans and monitor and evaluate the results. It is a participatory learning process aiming at solving the collective action problem by facilitating the critical 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 subgroups. Participants were divided into key social groups such as women, men, youths, disabled, local leaders and elderly in order to get their different perspectives over similar issues around service delivery and also to determine their desire for change according to their different priorities. Each group individually discussed the issues covered in the report card and recorded suggestions for improvements and prioritized these issues. 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 to identify those 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. To put this technique into the context of the participants, they were asked to consider the different domestic needs and resources they have available. 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. These factors included funding, resources, time frame, how pressing the need was, and whether other partners were needed in the implementation process. This tool helped the community analyze

[35] See World Bank (1996).

the resources available, the time frame for implementing the desired change and the seriousness of the problem that is to be addressed.

Role play was used to illustrate community and health facility interactions as perceived by the respective parties and facilitate the discussion and dialogue in the interface meeting between health facility staff and the community members. The story of the play illustrated the participants' interpretation of an ordinary day at the health facility. In the play, community members were asked to act 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). This was a highly effective and enjoyable tool. It vividly depicted all the hidden ills as they happen at the health facility and it was very effective in bringing out the voice of the users in the face of the providers so that they can 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 how the participants would like the situation to be in six months.

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 clearly states the health issues/services that had been identified by the community and the health facility 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 kept both by the community and the facility staff and forms the basis for local monitoring and helps keeping track of the status of the recommendations.

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. It also highlights those areas where external support and advice might be needed. 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.

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

	Treatment group	Control group	Difference
Utilization:			
Out-patient care	857	908	-51
			(141)
Delivery	11.74	7.89	3.85
			(2.68)
Utilization pattern:			
Project facility	0.32	0.34	-0.02
			(0.03)
NGO health facility	0.02	0.02	-0.001
			(0.009)
Private-for-Profit health facility	0.24	0.27	-0.03
			(0.02)
Traditional healer	0.034	0.03	0.004
			(0.006)
Self treatment (drug shop)	0.36	0.32	0.04
			(0.03)
Other government health facility	0.17	0.18	-0.01
			(0.05)
Other provider	0.012	0.007	0.005
			(0.004)
Quality measures:			
Waiting time	151	140	9
			(9.9)
Equipment usage	0.44	0.49	-0.05
			(0.06)
Funding at the facility:			
1000 shillings	4766	3429	1337
			(905)

The results are catchment area (health facility) averages. 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	23.1	24.7	-1.6
			(3.12)
Number of villages per health facility in strata 1	2.6	2.0	0.60
			(0.45)
Number of villages per health facility in strata 3	8.8	9.6	-0.80
			(1.7)
Number of villages per health facility in strata 5	11.6	13.2	-1.6
			(1.69)
Number of households per health facility	2139	2232	-93
			(274)
Number of households per village	94.6	95.4	-0.80
			(8.3)
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
			(1.91)
Number of days without electricity in last month	18.3	20.4	-2.12
			(4.14)

The results are catchment area (health facility) averages. 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 at the health facility 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	3.16	3.17	-0.01
			(0.04)
Free to express	3.8	3.78	0.01
			(0.04)
Citizens' informations about drug deliveries	0.14	0.16	-0.02
			(0.04)
Supply of drug deliveries at the health facility:			
Erythromycin	420	346	74
			(134)
Chloroquine	3410	2915	495
			(567)
Septrine	2690	2430	260
			(623)
Quinine	573	335	238*
			(129)
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. 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 and 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 and 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 health facility utilization

Dependent variable	Out-Patient		Delivery		Antenatal		Family Planning	
Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Program impact	102.3*	128.5**	6.3*	7.1**	16.1	16.7*	5.5	9.5*
	(52.6)	(56.2)	(3.5)	(2.8)	(13.3)	(8.7)	(4.8)	(5.1)
Constant	659.1***	19.1	9.2***	-28.9**	78.8***	-112.8***	15.2***	9.8
	(32.5)	(197.9)	(2.5)	(10.7)	(14.0)	(50.7)	(3.9)	(8.8)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
District fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	50	50	50	50	50	50	50	50
R^2	0.05	0.39	0.07	0.62	0.02	0.60	0.03	0.41

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Robust standard errors in parenthesis. Disturbance terms are clustered within districts.
c. Control variables include: type of watersource at the health facility, availability of drinkingwater at the health facility, size of catchment area and whether the health facility has a separate maternity unit.

Table 3. Difference-in-differences estimates of the program impact on health facility utilization.

Dependent variable	Out-Patient Services	Delivery
Specification	(1)	(2)
Treatment group	-50.7	2.84
	(107.9)	(3.11)
Program impact	153.1*	3.48*
(Treatment*2005)	(76.2)	(1.87)
2005	-249.02***	1.74
	(66.2)	(1.28)
Constant	908.1***	7.48***
	(82.5)	(1.73)
Observations	100	100
R^2	0.06	0.08

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Robust standard errors in parenthesis. Disturbance terms are clustered within districts.

Table 4. 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)
Treatment group	-0.03	-0.003	-0.02	0.004	0.045	0.015	0.007
	(0.03)	(0.007)	(0.03)	(0.007)	(0.03)	(0.05)	(0.004)
Program impact	0.04**	-0.004	0.03	-0.013*	-0.03*	0.0002	-0.01
(Treatment*2005)	(0.018)	(0.006)	(0.02)	(0.007)	(0.02)	(0.05)	(0.02)
2005	-0.08***	0.008	-0.02	-0.003	0.03***	-0.045**	0.05***
	(0.013)	(0.005)	(0.014)	(0.005)	(0.01)	(0.02)	(0.01)
Constant	0.34***	0.02***	0.26***	0.03***	0.34***	0.17***	0.007***
	(0.02)	(0.004)	(0.02)	(0.004)	(0.02)	(0.03)	(0.002)
Observations	9200	9200	9200	9200	9200	9200	9200
R^2	0.01	0.001	0.001	0.003	0.003	0.001	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. Dependent variable is citizens' use of different service providers in case of illness (reported in percentages).

Table 5. Citizens' perception of changes in 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)
Constant	0.36***	0.48***	0.49***
	(0.09)	(0.08)	(0.07)
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 changes in overall quality; (2) Dummy variable indicating changes in staff politeness; (3) Dummy variable indicating changes 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.

Table 6. Difference-in-difference estimates of the program impact on treatment practices at the health facility.

Dependent variable	Equipment usage	Waiting time
Specification	(1)	(2)
Treatment group	-0.01	4.3
	(0.06)	(9.9)
Program impact	0.08**	-15.5**
(Treatment*2005)	(0.03)	(7.3)
2005	-0.07***	143.6***
	(0.02)	(6.9)
Constant	0.48***	-10.6**
	(0.04)	(5.3)
Observations	5280	5148
R^2	0.003	0.01

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 7a. 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	-0.037	0.052*	0.002	0.027**	0.009	-0.044	0.042**	0.002	0.031***	-0.001
	(0.05)	(0.029)	(0.019)	(0.012)	(0.014)	(0.04)	(0.021)	(0.018)	(0.010)	(0.011)
Constant	0.38***	0.83***	0.93***	0.95***	0.96***	0.39***	0.81***	0.93***	0.91***	0.97***
	(0.04)	(0.02)	(0.01)	(0.010)	(0.01)	(0.09)	(0.04)	(0.04)	(0.02)	(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	526	983	991	1176	547	526	983	991	1176	547
R^2	0.001	0.005	0.000	0.006	0.001	0.03	0.04	0.01	0.02	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 5.

Table 7b. 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	0.056	0.050*	0.029	0.037*	0.027	0.047*	0.045*	0.030**	0.035*	0.020
	(0.036)	(0.029)	(0.021)	(0.020)	(0.026)	(0.028)	(0.025)	(0.015)	(0.019)	(0.026)
Constant	0.41***	0.82***	0.88***	0.91***	0.91***	0.45***	0.75***	0.80***	0.88***	0.93***
	(0.03)	(0.02)	(0.02)	(0.015)	(0.018)	(0.06)	(0.06)	(0.04)	(0.04)	(0.06)
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	939	971	977	1160	554	939	971	977	1160	554
R^2	0.003	0.005	0.002	0.005	0.003	0.03	0.03	0.04	0.02	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 5.

Table 7c. 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	0.053	0.018	0.078**	0.013	0.012	0.053**	0.024	0.066***	0.002
	(0.039)	(0.045)	(0.039)	(0.039)	(0.034)	(0.032)	(0.025)	(0.020)	(0.023)	(0.030)
Constant	0.38***	0.76***	0.82***	0.83***	0.90***	0.39***	0.41***	0.58***	0.61***	0.83***
	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.07)	(0.06)	(0.05)	(0.04)	(0.06)
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	940	977	993	1177	550	940	977	993	1177	550
R^2	0.003	0.004	0.001	0.010	0.001	0.03	0.13	0.09	0.10	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 5.

Table 7d. 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	0.070**	0.001	-0.014	0.015	0.016	0.065**	0.002	-0.013*	0.013	0.011
	(0.035)	(0.013)	(0.008)	(0.009)	(0.013)	(0.031)	(0.011)	(0.007)	(0.009)	(0.010)
Constant	0.79***	0.96***	0.99***	0.97***	0.98***	0.84***	1.00***	1.00***	0.97***	0.95***
	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.05)	(0.03)	(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	938	972	985	1159	540	938	972	985	1159	540
R^2	0.009	0.000	0.003	0.003	0.004	0.03	0.01	0.02	0.01	0.04

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 5.

Table 8. Program impact on health outcomes: Under-five child deaths.

Dependent variable	Child death (children < 5 year)	
Specification	(1)	(2)
Program impact	-0.016*	-0.017**
	(0.01)	(0.009)
Constant	0.049***	0.025
	(0.006)	(0.017)
Controls	No	Yes
District fixed effects	No	Yes
Observations	2922	2922
R^2	0.002	0.01

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable is a dummy variable indicating whether any children under-five in the household have died during the last year.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.
d. Control variables: see note (d) in Table 5.

Table 9. Program impact on health outcomes: Child weight of infants.

Dependent variable	Child weight of infants		
Specification	(1)	(2)	(3)
Program impact	0.13*	0.14**	-0.092
	(0.07)	(0.07)	(0.09)
Child age	0.25***	0.25***	0.15***
	(0.01)	(0.01)	(0.01)
Constant	5.63***	5.62***	7.52***
	(0.11)	(0.19)	(0.35)
Controls	No	Yes	Yes
District fixed effects	No	Yes	Yes
Observations	1152	1152	1422
R^2	0.45	0.45	0.22

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable is child weight in kilograms of infants younger than 18 months.
c. Specification: (1) Includes all children under 18 months, (2) Includes all children under 18 months plus controls, (3) 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 5.

Table 10. 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)
Treatment group	0.02	0.07**	0.05	0.06
	(0.01)	(0.036)	(0.06)	(0.09)
Program impact	-0.01	-0.05*	-0.15**	-0.09
(Treatment*2005)	(0.01)	(0.026)	(0.07)	(0.10)
2005	0.002	-0.016**	0.15**	-0.14***
	(0.005)	(0.006)	(0.05)	(0.05)
Constant	0.009**	0.03***	0.22***	0.64***
	(0.004)	(0.007)	(0.05)	(0.07)
Observations	5660	5734	2511	507
R^2	0.003	0.03	0.01	0.04

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

Table 11. 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
Specification	(1)	(2)	(3)	(4)
Program impact	0.38***	0.20***	0.19**	0.12
	(0.12)	(0.06)	(0.08)	(0.15)
Constant	-0.25***	0.20***	-0.12**	-0.08
	(0.08)	(0.04)	(0.05)	(0.09)
District fixed effects	Yes	Yes	Yes	Yes
Observations	50	49	50	50
R^2	0.35	0.30	0.47	0.26

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable in specifications is 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.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within districts.

Table 12. Program impact on citizens' information.

Dependent variable	Health information	Importance of family planning	Drug deliveries
Specification	(1)	(2)	(3)
Program impact	0.09***	0.07***	0.03**
	(0.02)	(0.02)	(0.013)
Constant	0.28***	0.31***	0.09***
	(0.07)	(0.06)	(0.02)
District fixed effects	Yes	Yes	Yes
Observations	4996	4996	4996
R^2	0.16	0.10	0.17

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, (3) Dummy variable indicating whether the household knows when the health facility receives drugs.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

Table 13. Program impact on citizens' information about HUMC and local council meetings.

Dependent variable	Discuss the health facility in LC meetings	Know members of HUMC	Knowledge of HUMC
Specification	(1)	(2)	(3)
Program impact	0.13***	0.09***	0.05***
	(0.02)	(0.02)	(0.01)
Constant	0.40***	0.22***	0.09***
	(0.07)	(0.03)	(0.01)
District fixed effects	Yes	Yes	Yes
Observations	3119	4996	4996
R^2	0.11	0.05	0.05

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable in specifications: (1) Dummy variable indicating if the household discusses the functioning of the health facility at Local Council meetings, (2) Dummy variable indicating if the household knows any member of the HUMC, (3) Dummy variable indicating if the household has received information about HUMC's roles and responsibilities during the last year.
c. Robust standard errors in parenthesis. Disturbance terms are clustered within catchment areas.

Table 14. 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	-1.39	-0.10	0.07	-0.56
	(1.23)	(0.08)	(0.16)	(0.66)
Constant	702***	12.4***	13*	96***
	(62)	(3.1)	(7)	(22.9)
Observations	25	25	25	25
District fixed effects	No	No	No	No
R^2	0.03	0.06	0.01	0.03

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Robust standard errors in parenthesis. Disturbance terms are clustered within districts.

Table 15. Robustness test: 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 facility	2.28	-0.14**
	(3.88)	(0.06)
Distance to closest project	-3.67	0.04
facility in 2005	(3.85)	(0.05)
2005	-135.8	0.53
	(139.5)	(1.52)
Constant	837.7***	11.8***
	(176.9)	(2.85)
Observations	50	50
District fixed effects	No	Yes
R^2	0.15	0.12

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Robust standard errors in parenthesis. Disturbance terms are clustered within districts.

Table 16. Program impact on funding to the health facility.

Dependent variable	Funding (1000 shilling)
Treatment group	1337
	(893)
Program impact	-453
(Treatment*2005)	(866)
2005	988
	(980)
Constant	3429***
	(501)
Controls	No
District fixed effects	No
Observations	94
R^2	0.04

a. *** [**] (*) denote significance at the 1 [5] (10) percent level.
b. Dependent variable: 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. Disturbance terms are clustered within districts.
d. Control variables: see note (c) in Table 2.

Table 17. Difference-in-difference estimates of drugs supply received at the health facility.

Dependent variable	Erythromycin	Chlorquine	Septrine	Quinine	Mebendazole
Specification	(1)	(2)	(3)	(4)	(5)
Treatment group	74	496	260	238**	97
	(122)	(399)	(438)	(79)	(150)
Program impact	92	-176	-3	-243*	114
(Treatment*2005)	(102)	(515)	(659)	(93)	(590)
2005	-89	-531	-457	-30	984*
	(112)	(546)	(575)	(144)	(560)
Constant	346***	2915***	2430***	335***	1500***
	(86)	(357)	(468)	(101)	(186)
Observations	96	100	100	99	100
R^2	0.03	0.03	0.02	0.08	0.11

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

Table 18. 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	0.10	0.05	0.05
	(0.14)	(0.11)	(0.12)	(0.08)
Constant	1.53***	0.83**	-0.41	0.28
	(0.23)	(0.34)	(0.13)	(0.33)
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. Disturbance terms are clustered within districts.
d. Control variables: see note (c) in Table 2.

Table 19. Difference-in-difference estimates on equipment at the health facility.

Dependent variable	Bicycles	Examination beds	Blood pressure machine	Weighing scale	Microscope
Specification	(1)	(2)	(3)	(4)	(5)
Program impact	0.24	0.16	-0.24	0.12	0.28
	(0.44)	(0.17)	(0.32)	(0.47)	(0.22)
Program impact	0	0.20	-0.08	0.08	0.20*
	(0.12)	(0.24)	(0.15)	(0.11)	(0.09)
2005	0.40**	0.20	0.36*	0.12	0.04
	(0.15)	(0.20)	(0.12)	(0.47)	(0.04)
Constant	2.52***	1.8***	1.68***	2.6***	0.44**
	(0.53)	(0.14)	(0.24)	(0.30)	(0.16)
Observations	100	100	100	100	100
R^2	0.01	0.03	0.04	0.006	0.09

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. Disturbance terms are clustered within districts.