Can Citizen Pressure Be Induced to Improve Public Service Provision?

Pia Raffler [*] Daniel N. Posner [†] Doug Parkerson [‡]

This version: March 23, 2022

Word count: 11,714

Abstract

Encouraging citizens to apply pressure on underperforming service providers has emerged as a prominent response to the failure of states to provide needed services. We outline three theoretical mechanisms through which bottom-up citizen pressure campaigns may affect service provision and investigate them via a large-scale field experiment in the Ugandan health sector. While we find modest positive impacts on health provider behavior, we find no effects on citizen pressure, utilization rates, or bottom-line health outcomes. Our findings cast doubt on the power of outside actors to generate improvements in development outcomes by mobilizing bottom-up pressure—at least under conditions similar to those in our study setting. Our results underscore the importance of baseline health conditions for the success of bottom-up, citizen-oriented pressure campaigns. Such conditions shape outcomes both across countries and within countries over time, with the latter finding holding important implications for countries undergoing rapid socioeconomic change.[1]

[*]Department of Government, Harvard University, praffler@gov.harvard.edu.

[†]Department of Political Science, University of California, Los Angeles, dposner@polisci.ucla.edu.

[‡]Innovations for Poverty Action, dparkerson@poverty-action.org.

[1]Authors listed in reverse alphabetical order. We thank the Ministry of Health in Uganda, in particular Dr. Anthony Mbonye; IPA Uganda, in particular Frédéric Cochinard, Martin Atyera, Joshua Bwiira, Paola Elice, Afke Jager, Kyle Holloway, Douglas Kaziiro, Steven Kizza, Ezra Rwakazooba, Laura Schmucker, and Alex Tusiime for excellent research assistance and Damien Kirchhoffer, Dickson Malunda, and Daniele Ressler for their dedicated project management. We thank our partners at GOAL Uganda, in particular Elizabeth Allen, Angela Bailey, Niamh Barry, and Fiona Mitchell; our implementing partners, Coalition for Health Promotion and Social Development, Kabarole Research and Resource Centre, and Multi-Community Based Development Initiative; and the Department for International Development for funding. Helpful comments were provided by participants at the 17th and 23rd EGAP meetings, and at seminars at Columbia, Emory, Rochester, Stanford, Texas A&M, UC Berkeley, UCLA, Vancouver School of Economics, and the World Bank. The pre-analysis plan was registered at EGAP (ID 20160611AA). The project was approved by the Internal Review Boards at IPA (Protocol ID: 0497) and the Uganda National Council for Science and Technology (UNCST) (Protocol ID: ARC157). Approval was also received from UNCST (Protocol ID: SS3559) and the Office of the President, Uganda. A working paper version of this paper was submitted for pre-publication re-analysis to the Abdul Latif Jameel Poverty Action Lab (J-PAL), where a code replication exercise was undertaken. We thank Georgiy Syunyaev and Isabelle Cohen for conducting this re-analysis.

1 Introduction

Public service provision is the most important function of the state. What happens, then, when the state fails to provide crucial public services? In many instances, citizens simply go without: public safety is not protected, clean water is unavailable or unreliable, waste is uncollected, roads are potholed, and healthcare and schooling are underprovided. Private or non-state actors sometimes emerge to substitute for the state in providing these public goods, but the services they offer are often out of reach of the poorest community members (MacLean, 2011; Cammett and MacLean, 2014). Under such circumstances, citizens can and sometimes do apply pressure on the state to improve its performance. Given the significant share of the world's population that faces severe shortfalls in public service provision and is too poor to acquire these services through the private sector, a critical question is whether such bottom-up pressure can be induced and whether—or under what conditions—it can lead to improvements in service delivery.

Bottom-up pressure can be applied either directly on underperforming frontline service providers or indirectly on the political representatives who are responsible for the providers' performance. The latter "long" route (World Bank, 2004), while central to theorizing about democratic accountability, assumes a responsiveness to electoral pressure that is frequently absent in places where public service delivery is most deficient (Chong et al., 2015; De Kadt and Lieberman, 2020; Dunning et al., 2019; Hern, 2019). The alternative, non-electoral, "short" route, whereby citizens monitor and apply pressure directly on absent or underperforming frontline service providers, may be more promising in uncompetitive, semi-democratic political systems such as those found in many low-income countries.

Generating bottom-up pressure can be challenging, however. Collective action problems (Olson, 1971), low efficacy (Kruks-Wisner, 2018; Lieberman and Zhou, 2021), low expectations of government capacity and/or responsiveness (Golooba-Mutebi, 2005; Gottlieb, 2016; Hern, 2019), and the weakness of supportive local institutions (Ostrom, 1990) all present obstacles to mobilizing citizens to demand better public services. This paper addresses whether providing citizens with information about service delivery shortfalls, along with guidance about how to mobilize and apply pressure in light of this information, is sufficient for overcoming these challenges.

We assess the viability of this information-focused, non-electoral approach through a field experiment undertaken in 187 health centers and their associated catchment areas in 16 districts in Uganda. Motivated by the theoretical literature on principal agent problems and designed in keeping with the received wisdom in development circles on the use of information to generate bottom-up pressure on service providers (Björkman and Svensson, 2009; Mansuri and Rao, 2013; Kosack and Fung, 2014), the intervention we study delivered information about patient rights and responsibilities, utilization patterns, and health outcomes at the local health center, worked with

health center staff and community members to develop action plans in light of that information, and organized meetings between members of the community and health center staff to generate a joint social contract to guide both actors' future behavior and interactions.

We randomized whether health centers received the intervention and, to assess its impact, collected three waves of annual panel data on citizen monitoring and pressure, utilization rates, treatment quality, patient satisfaction, and health outcomes at both the health center (N=187) and household (N=7,288) levels. To capture the channels through which the intervention operated, we collected data on a broad array of intermediate outcomes. We also collected data on health center, community, and household characteristics to better understand the conditions under which the intervention had the greatest impact.

While we find positive (albeit substantively small) impacts on treatment quality and patient satisfaction, these improvements are not associated with statistically significant effects on utilization rates, child mortality, or other health outcomes on average, either eight or twenty months after the intervention. These null findings are reinforced when we examine sub-populations of health centers, communities and individuals: we find persistently null effects on all of these outcomes across the vast majority of subgroups. Most important from the standpoint of the presumed mechanism that inspired the intervention, we find little evidence that the attempt to induce bottom-up pressure caused citizens to increase their monitoring or sanctioning of health care workers. Taken together, our findings cast doubt on the ability of information and citizen mobilization to generate bottom-up pressure on health workers or improvements in health outcomes in the context we study. The improvements we find in treatment quality appear to be a direct effect of exposure to the intervention on health providers' behavior, not a product of citizen pressure.

The paper makes three main contributions. First, we contribute to the literature emphasizing information and community monitoring as solutions to the problem of poor service provision. Mansuri and Rao (2013) estimate that, in the early 2000s, the World Bank invested $85 billion in programming motivated by this approach. Notwithstanding its broad embrace in development circles, this strategy has found only mixed empirical support.[2] Olken (2007), Banerjee et al. (2010), and Keefer and Khemani (2014) all report weak effects of interventions designed to generate behavioral change by frontline service providers through information provision and bottom-up grassroots monitoring. Pandey, Goyal and Sundararaman (2009), Barr et al. (2012), Pradhan et al. (2014), Andrabi, Das and Khwaja (2017), Fiala and Premand (2018), and Banerjee et al. (2018), meanwhile, find more promising results. In the health-focused studies closest to our own, Björkman and Svensson (2009) find strong positive effects on infant weights, under-five mortality, immunization rates, and other measures of health service delivery; Christensen et al. (2021) find

[2]See Fox (2015) and Tsai et al. (2019) for recent reviews.

effects on utilization, patient satisfaction, and child mortality, but not on service quality or other health outcomes; Mohanan et al. (2020) find impacts on utilization and child mortality, but not on treatment quality; and Arkedis et al. (2021), who study interventions in two different countries, find no effects on any health-related outcomes in either setting. Our study joins these efforts by providing a particularly high-powered test of the potential impact of information provision and community mobilization as tools for generating bottom-up citizen pressure and improvements in service delivery.

A second contribution is to distinguish among, and to test, three different mechanisms through which bottom-up interventions may generate improvements in development outcomes. While our findings suggest that the intervention we studied did little to improve health outcomes by mobilizing *citizen pressure* or encouraging greater *utilization* of formal health facilities, they suggest that it did have a positive, if modest, *direct impact on the behaviors of health providers*. As we show in Section 7, the distinction among these three channels is also useful for making sense of the divergent findings in the broader literature.

A third contribution is to emphasize the importance of baseline health conditions. Notwithstanding our null results on the various health outcomes we study, we do find significant treatment impacts on child mortality, as well as stronger effects on treatment quality, in the subset of catchment areas in our sample with lowest levels of development and the highest baseline child mortality rates. This may help explain the differences across the several health-related studies mentioned above. All six studies examine the impact of very similar interventions emphasizing bottom-up citizen pressure—indeed, both Christensen et al. (2021) and our own study were modeled explicitly on Björkman and Svensson (2009)'s pathbreaking intervention, and both Mohanan et al. (2020) and Arkedis et al. (2021) adapt many of its key design features. But baseline health conditions in the six settings were quite different, and, as we show, the strength of the reported treatment effects are broadly correlated with these baseline conditions. Bottom-up citizen pressure interventions appear to be more effective when baseline health conditions are lower.

This is an especially important lesson for researchers and policymakers working in countries, like Uganda, that are undergoing rapid socioeconomic change. Held up against the results of these other studies—especially Björkman and Svensson (2009), which was also implemented in Uganda, but ten years earlier when health conditions were significantly worse—our findings underscore the often underemphasized *temporal* dimension of external validity, and the extent to which interventions that may be highly effective under one set of conditions may lose their power when conditions improve.

2 How Bottom-Up Pressure Campaigns May Improve Service Delivery Outcomes

Attempts to generate bottom-up pressure for service delivery improvements have been undertaken in low income settings around the world (see Appendix A for a partial listing of such interventions). These interventions almost always involve two core components: 1) the provision of information about the relative performance of the service delivery unit and 2) the convening of meetings aimed at mobilizing communities in light of this information and helping communit members overcome collective action problems. Some efforts also add a third component involving interface meetings in which citizens and service providers come together to discuss how they might work together to jointly improve service delivery outcomes. These components may affect development outcomes through one of three distinct channels, as illustrated in Figure 1.

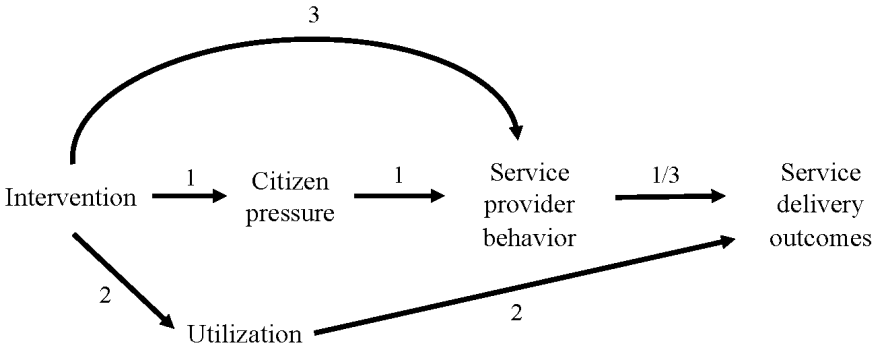


Figure 1: Three channels through which bottom-up pressure campaigns may improve service delivery outcomes

The first, *citizen pressure*, channel (pathway 1 in Figure 1) occurs in two steps. First, the distribution of information about the relative performance of the service delivery unit puts citizens in a stronger position to evaluate whether their own local service providers are performing adequately (Besley and Case, 1995; Kruks-Wisner, 2018) and creates common knowledge about the service providers' performance. Second, the holding of community meetings may generate internal efficacy (Lieberman and Zhou, 2021), foster a sense of responsibility for monitoring service providers (Pandey, Goyal and Sundararaman, 2009), help overcome free riding problems, and enable citizens to identify concrete actions they can take to improve services—all of which may be critical for generating bottom-up pressure by citizens (Lieberman, Posner and Tsai, 2014; Kruks-Wisner, 2018). The resulting bottom-up pressure may then translate into improved service delivery by inducing service providers to exert more effort, divert fewer resources, and allocate resources more efficiently. Interface meetings, where held, may augment the impact of this pressure by

allowing citizens to confront service providers directly, apply social sanctions on those revealed to be underperforming, and, where successful, bolster citizens' external efficacy and generate improvements in the relationship between community members and service providers. The resulting improvements in service delivery are then expected to generate improved outcomes. Evidence that the citizen pressure channel is operating would be found in increases in citizen monitoring of service delivery and pressure by citizens on service providers. These, in turn, should be associated with improvements in service delivery, better development outcomes, and possibly also increased utilization in response to the improvements in service delivery.

The citizen pressure channel is the mechanism that most researchers and policymakers have in mind when they think about how information and community mobilization may impact service delivery. Indeed, it motivated the intervention we study here and most of the interventions described in Appendix A. However, there are two alternative channels, less well emphasized in the literature and not involving citizen pressure, through which an ostensibly "bottom-up" intervention might also generate improvements in service delivery.

First, the intervention might affect outcomes through an increase in *utilization* (pathway 2 in Figure 1). Utilization is critically important in sectors such as health and education, as well as for some aspects of public safety (i.e., reporting crime) and sanitation (i.e., using latrines). In the health sector specifically, an information and mobilization intervention may improve health outcomes by encouraging vaccinations or by causing sick people to seek professional care at the government health center rather than self-treating or visiting traditional healers. This may occur if the dissemination of information and the holding of community meetings make the existence of the health center more salient, build trust between community members and health care providers (Christensen et al., 2021), or reduce uncertainty about the costs of seeking services at the health center. Under such circumstances, we may observe improvements in health outcomes even in the absence of citizen pressure or improvements in treatment quality. This channel should be particularly relevant in settings where baseline utilization rates in the formal health sector are low.

Second, the intervention may *directly affect the behavior of frontline service providers* (pathway 3 in Figure 1). The dissemination of unit-specific information may make workers feel monitored, which may cause them to put more effort into service provision (Duflo, Hanna and Ryan, 2012; Nagin et al., 2002; Olken, 2007). Learning about the performance of their unit relative to others may also increase workers' intrinsic motivation to provide better services—especially if the information they receive suggests that their unit is under-performing. If this channel is operating, we would expect to observe improved treatment quality (possibly resulting in increased utilization) and improved development outcomes, even in the absence of citizen pressure.

A significant advance of our study over prior work is our ability to disentangle and test these

three separate mechanisms.

3 Health Service Delivery, Citizen Participation, and Government Accountability in Rural Uganda

Public health services in Uganda are provided in a hierarchical system with national referral hospitals at the national level, regional referral hospitals at the regional level, general hospitals at the district level, and smaller scale health centers at the sub-county and parish levels—the former termed HC3s; the latter, HC2s. The intervention we study focuses on health care delivery at the HC3 and HC2 levels, the lowest levels of the public health system. HC3s, which are staffed by a trained medical worker and one or more nurses and lab technicians, provide preventative and out-patient care and have laboratory services to undertake basic tests.[3] They also generally have maternity wards and offer prenatal and antenatal services. HC2s, which comprise over half of all government-run health facilities and represent the primary source of professional medical care for many Ugandans, provide outpatient services and antenatal care (Uganda Ministry of Health, 2018). They are run by a nurse, sometimes working with a midwife and a nursing assistant. Both types of units are supported by Village Health Teams (VHTs) comprised of volunteer community health workers who undertake health education outreach, provide simple curative services, and refer patients to higher level health centers for treatment of more complicated conditions. Generally speaking, patients seek care at the facility closest to their home and are then referred to higher-level facilities as their medical condition requires.

Government-run health facilities, which are supposed to provide free services, operate alongside a growing number of private for-profit and not-for-profit (often religious) health providers, as well as traditional practitioners. In our sample at baseline, only 45% of households that reported having a health condition requiring treatment during the past year sought care at a government-run health center, whereas 17% sought care at a private or NGO-run clinic. Thirty-two percent visited traditional healers or self-treated, and 6 percent sought care from a member of the VHT.

Among the reasons cited for not visiting the government-run health center were lack of drugs, long waiting times, poor quality of services, and poor staff attitude. Factors both within and outside the health workers' control contribute to these outcomes. Understaffing, low and irregular pay, shortages of necessary medical supplies, and limited oversight by higher-level health officials are major problems (Uganda Ministry of Health, 2017; Nannyonjo and Okot, 2013; Tweheyo et al., 2019). They lead to low morale, absenteeism, and poor treatment quality, which in turn

[3]These are the government standards. At the time of our study, HC3s frequently did not have adequate staff or equipment and materials to provide the full set of services that government standards specified.

generate poor health outcomes and reduce incentives for citizens to utilize the government-run health facilities.

Despite notable improvements in health outcomes in Uganda over the past decades, health service delivery in rural areas remains poor. In 2015, Uganda ranked 177th of 196 countries in the Healthcare Access and Quality Index, which identifies mortality rates from causes that should not be fatal in the presence of effective medical care (Barber et al., 2017). In that year, 45% of Ugandan children aged 12-23 months were not fully vaccinated and the maternal mortality ratio was high, with 336 deaths per 100,000 live births (DHS 2016).

Poor health service delivery is a salient issue in the public discourse: 50% of Ugandans name health as one of the three most important problems that government should address, and 52% of Ugandans who had occasion to seek medical care at a public clinic or hospital in the past year report finding it difficult or very difficult to obtain medical treatment (*Afrobarometer Uganda*, 2015). Although there does exist some space for citizen demand-making on this and other issues (50% of Ugandans report having gotten together with others to raise an issue in the past year; 21% report having contacted a member of their local council directly; and 5% having participated in a protest or a demonstration (*Afrobarometer Uganda*, 2015)), many Ugandans do not feel empowered or well represented. Only 14% say that parliamentarians or local councilors "often" or "always" try their best to listen to what people like them have to say (*Afrobarometer Uganda*, 2015).

These deficiencies in government accountability have systemic roots. President Museveni, who has been in power since 1986, presides over a semi-authoritarian regime in which the political playing field is heavily—and, in the years since our study, increasingly—slanted in favor of the ruling party (Tripp, 2010; Khisa, 2019; Platas and Raffler, 2021). This party dominance extends to the local councils that are tasked with overseeing health service delivery in conjunction with the Ministry of Health (Green, 2015). With an average vote margin of 52 percentage points for subcounty councilors and 38 percentage points for district councilors, local council elections are not competitive in large parts of the country, thus undermining electoral accountability (Barro, 1973).[4] Furthermore, even when local council members may be motivated to lean on local bureaucrats, they have limited capacity to monitor service delivery (Martin and Raffler, 2021; Raffler, Forthcoming). NGOs, which play a supportive role in other contexts, are stifled by legal restrictions on their activities and the constant possibility of government sanctions (Namisi, 2009). In these respects Uganda is similar to many countries in the Global South that have "low capacity states" Hern (2019) and semi-authoritarian regimes.

The result is a situation where "long route" electoral accountability is weak but where "short route" direct pressure by citizens on frontline service providers may be viable, in part because the

[4] Authors' calculations based on 2011 data from the National Electoral Commission.

frontline service providers are less likely than elected officials to be viewed as agents of political parties. Indeed, roughly a third of household heads in our baseline survey said they had "a lot" or "some" power to improve the quality of health care at their local health center, and 81% agreed with the statement that "people like you have a say in how health facilities provide health care to your community" (see Appendix G). This is precisely the opportunity that initiatives aimed at mobilizing citizens to apply bottom-up pressure directly on service providers seek to exploit.

4 The "Accountability Can Transform (ACT) Health" Initiative

The intervention we study, Accountability Can Transform (ACT) Health, was implemented by a consortium of civil society organizations coordinated by GOAL Uganda.[5] Designed in keeping with bottom-up citizen-oriented pressure campaigns deployed in other settings, it consisted of three components (additional information about each is provided in Appendix I.4):

Information. The research team used data collected in the baseline health center and household surveys to create citizen report cards (CRCs) providing health center-specific information about citizens' knowledge of their rights and responsibilities, utilization of the various services offered at the health center, citizens' perceptions of the quality of these services, and overall satisfaction with the health care they received. For most outcomes, the health center-specific data was presented alongside district averages to provide a benchmark of relative performance. The CRCs were shared with both health care providers and community members. Information was presented with the help of visual props designed by local artists to ensure comprehension among illiterate participants.

Mobilization. Trained facilitators worked with local leaders and VHT members to organize community meetings at which the CRC results were presented and discussed. An action plan was developed to identify specific steps that could be taken by community members to improve health service delivery. Significant efforts were made to ensure that the meetings included representatives from all major social groups in the community.[6] Parallel meetings were also held separately with health center staff at which the CRC results were discussed and an action plan was formulated describing steps that the staff could take to improve health outcomes.

Interface. Facilitators brought the health center staff together with representatives of the community to discuss their respective action plans and how they might work together to improve the quality of health care in the community.[7] The output of the interface meeting was a social

[5]Further details of the implementing organizations are provided in Appendix I.2. See Bailey and Mujune (2021) for additional information.

[6]The meetings included an average of 100 attendees. Further details about the participants and intervention implementation are provided in Appendix I.3.

[7]On average, 50 community members and four health center staff members participated in the interface meetings.

contract between the citizens and health care workers laying out specific steps that each could take to contribute to improvements in health outcomes.

Implementing teams spent several days in the communities surrounding each health clinic to organize the various meetings, and they returned every six months (for a total of three follow-up visits before endline data collection) to meet with community members and health center staff to check on the progress that had been made toward the commitments stipulated in the social contract. A timeline of the intervention is provided in Figure 2. Examples of a CRC, community and health center action plans, and a joint social contract are included in Appendix I.4. We discuss ethical considerations associated with the project in Appendix B.

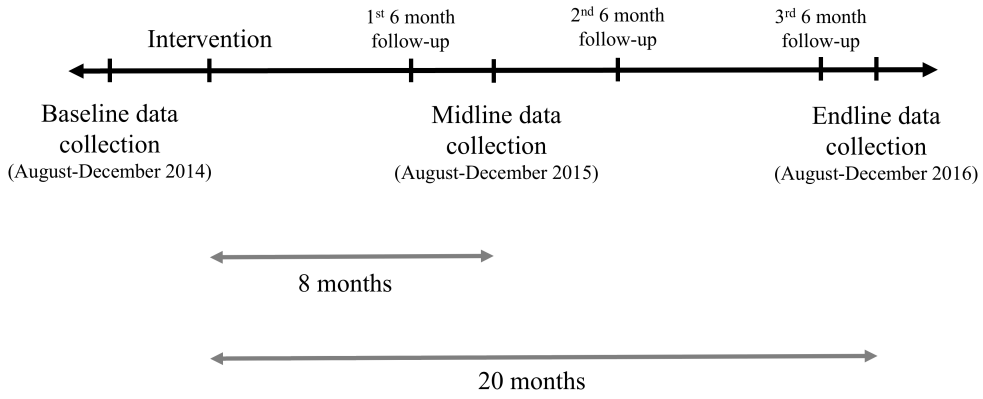


Figure 2: Timeline of the intervention

5 Data and Estimation

The unit of randomization in our study is the health center and its associated catchment area. The sample we focus on in this paper includes 187 health centers spread across sixteen districts.[8] These health centers were randomized to receive the ACT Health intervention, with blocking by district and health center (HC2 or HC3) level. We define the catchment area as the three villages that are closest in proximity to the health center in question (including the village in which the health center is located), as measured by the straight-line distance from the health center to the village

Further details are provided in Appendix I.3

[8]The full study included 376 health centers, but half of them received only partial treatments, in keeping with the factorial design described in our pre-analysis plan (see Appendix H.8). Because the effects in these additional treatment arms do not substantively change the interpretation of our results, we focus our discussion in the paper on the impact of exposure to the full treatment only. The sixteen districts are: Lira, Apac, Pader, Gulu, Lamwo, Kitgum, Agago, Katakwi, Bukedea, Manafwa, Tororo, Kabarole, Mubende, Nakaseke, Kibaale, and Bundibugyo. A map is included in Appendix I.1. We excluded health centers located in prisons or military installations.

centroid.[9] In identifying these villages, we only include villages located in the same parish (for HC2s) or sub-county (for HC3s) as the health center in question.

5.1 Data

Our data come from two main sources: a household survey and a health center survey. Both were collected at baseline, midline, and endline, with as close as possible to 12 months separating each survey round in each health center/catchment area in order to control for seasonal effects that might influence utilization rates or health outcomes. Data collection staff were completely separate from the teams that implemented the programming and had no knowledge of the treatment status of the health centers and households they visited.

Since treatment could not be administered until after the baseline data had been collected and distilled into the CRCs, the average interval between intervention and midline data collection was less than one year (8 months; SD=1.37 months). The average interval between the intervention and endline data collection was 20 months (SD=1.34 months). In the results presented below, we privilege the endline findings, but we report the full midline results in Appendix H.9.

The health center survey consisted of three components. The first was a brief questionnaire completed at the time of initial contact with the health center in each survey round. Since this visit was unannounced, it provided an opportunity for the collection of information about staff attendance, cleanliness, wait times, and other clinic characteristics before the clinic staff was able to respond to the fact that it was being evaluated. The second component was the main health center staff survey, which collected information about the variety and quality of health services provided, utilization rates, staff structure and perceptions, funding mechanisms, and drug stock-outs. This survey was conducted with the most senior health center staff member, as well as randomly drawn health workers. The third component involved the collection of administrative data on file at the health center, including monthly Health Management Information System (HMIS) forms and drug stock cards. Physical checks of drug stocks were conducted to verify the accuracy of these records.

The sample for the household survey was drawn from a sampling frame of households in the catchment area containing at least one child under five years old or a pregnant woman, based on village household lists and consultations with the village chairperson, VHT members, Health Unit Management Committee (HUMC) members, and other knowledgeable persons. We randomly sampled 40 households per catchment area from this frame, with the number of households drawn

[9]Catchment areas were determined using village-level shape files provided by the Uganda Bureau of Statistics, and health center GPS coordinates collected by GOAL. To minimize overlap of catchment areas (and spillovers), we excluded health centers that were less than 2.5 km apart or that shared a village among their three closest villages.

from each village proportional to the number of eligible households in that village.[10]

The primary respondent for the household survey was the female head of household. The survey collected information about household members' recent experiences with the local health center, knowledge about their rights and responsibilities, health status, and participation in community activities (including those directly related to monitoring the performance of their local health center), among other topics. All household surveys also included an anthropometric survey in which we recorded the weight, height, and middle-upper arm circumference (MUAC) of each child under the age of five. The ages of the children and their immunization status were verified using immunization cards, if available. At endline, we also collected retrospective information on the month of birth and, if applicable, death of all children recorded at baseline and midline in order to generate more precise estimates of child mortality rates, as described in Appendix E.

The household surveys were conducted in ten local languages with the help of 279 field staff hired and trained by IPA Uganda.[11] In all, we completed 15,295 household surveys at baseline, 14,459 at midline, and 14,609 at endline.[12] Thanks to detailed tracking protocols, we were successful in re-interviewing 95.5 percent of our study households at endline. The analyses we discuss in the paper are based on the panel of 7,288 households in the 187 health centers receiving the full treatment or in the control arm, each interviewed at minimum at baseline and endline, and the vast majority at three different points in time. As shown in Appendix F.2, attrition is balanced across treatment arms.

5.1.1 Outcomes of Interest

We estimate the impact of the ACT Health intervention on seven categories of outcomes: citizen monitoring, perceived citizen pressure, utilization rates, treatment quality, patient satisfaction, health outcomes, and child mortality.[13] Child mortality is, of course, also a health outcome, but we break it out as a separate category because of its singular importance as a bottom-line measure of health system performance. These outcomes map directly onto the components in Figure 1 and allow us to adjudicate among the three channels through which the bottom-up pressure campaign we study may have improved health service delivery.

[10] At baseline, an additional short survey was administered to another randomly selected 15 households in catchment areas assigned to the information and mobilization treatments. These additional households were included to reduce noise in the measures included in the CRC and to increase the likelihood that the community would feel that the CRC represented its views and experiences.

[11] Further details of the procedures employed to ensure data quality are discussed in Appendix C.

[12] This includes the data collected in the 189 catchment areas in the sub-treatment arms, which are discussed in Appendix H.8.

[13] The elevation of citizen monitoring from an intermediate to a main outcome, and the addition of perceived citizen pressure as an additional main outcome of interest, are departures from our pre-analysis plan. We discuss this, and other, deviations in Appendix D.

All seven outcomes are measured via an averaged z-score index (Kling, Liebman and Katz, 2007) constructed from a set of underlying components, the mean baseline values for which are presented in Table 1. In every case but child mortality, where lower values are associated with better outcomes, the index is constructed so that higher values imply a more positive outcome.

Table 1: Outcome indices and their components, baseline values

	Mean
Citizen monitoring	
Household members report having attended LC1 meetings in last year	88.9%
Household members who attended the LC1 meeting report that the local HC was discussed	65.9%
Household members think community would find out if health worker did not provide effort caring for patients ◇	73.0%
Household members think community would find out if health worker did not report for work ◇	73.5%
Perceived citizen pressure	
HC staff think community would find out if health worker did not provide effort in caring for patients	0.9%
HC staff think community members would find out if health worker did not report for work	0.5%
Any community member reported a health worker to HC staff within past 12 months ◇	38.1%
Utilization	
Vaccination rates, children <36 months for polio, DPT, BCG, and measles, by age bracket	75.3%
Share of self-reported visits to HC versus other providers	37.5%
Number of self-reported visits to HC by household members in past 12 months	14.0 visits
Treatment quality	
Whether equipment was used during most recent visit to HC	68.0%
Total time spent waiting for initial consultation and examination	104 mins
Whether person seeking care was examined by trained health staff during most recent visit	99.9%
Whether person seeking care had privacy during most recent examination	89.2%
Whether lab tests were administered during most recent visit	62.8%
Whether diagnosis was clearly explained during most recent visit	59.5%
Percent of staff in attendance during unannounced visit to HC	29.3%
Condition of HC (cleanliness of floors/walls; smell) as observed during unannounced visit	80.3%
% of months stock cards indicate availability of 6 key drugs in past 3 mos, as determined during unannounced visit	93.2%
Patient satisfaction	
Whether services currently offered at HC are judged to be of "very high" or "somewhat high" quality ◇	45.9%
Whether person seeking care was "very satisfied"/"satisfied" with quality of care received during most recent visit	67.8%
Whether person conducting exam appeared to be interested in health condition of person seeking care	90.1%
Whether person conducting exam listened to what person seeking care had to say	90.3%
Whether person seeking care felt free to express him/herself to person conducting exam	83.1%
Whether, compared to year before, availability of medical staff has improved at HC	48.8%
Health outcomes	
Weight for age among children aged 0-18 months (kg/month)	1.23
Weight for age among children aged 18-36 months (kg/month)	1.39
Middle-upper arm circumference among children aged 0-18 months (cm/month)	2.51
Child mortality	
0 to 5 years (main measure)	0.05‰
0 to 12 months	0.04‰
1 to 5 years	0.01‰

Notes. Baseline values reported here are for the full sample, across all treatment conditions. ◇ Baseline values were not collected; values shown are from the control group. ⋆ Vaccination rates are calculated at the household level as the percentage of children under 36 months who, subject to a six-week grace period, have received the full set of age-relevant vaccinations.

In addition to these seven main outcomes, we also test for treatment effects on six intermediate outcomes that map onto the mechanisms discussed in Section 2: citizen knowledge, health center staff knowledge, efficacy, community responsibility, the relationship between health workers and the community, and health center transparency. The components of these indices, along with baseline means, are listed in Appendix G. Estimating treatment effects on these intermediate outcomes can thus help us gain a deeper understanding of the mechanisms through which the intervention operates.

5.2 Estimation

To estimate the effect of exposure to ACT Health, we estimate the following intent-to-treat equation:

$$Y_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 Y_{ij}^0 + \beta_3 X_{ij} + \beta_4 X_{ij} * T_{ij} + \phi_d + u_{ij} \quad (1)$$

where Y_{ij} is the outcome measure (in our main specifications, one of our seven indices) of household i in health center catchment area j. T_{ij} is a binary variable indicating whether the health center and catchment area j was assigned to treatment. β_1 is the average treatment effect, Y_{ij}^0 is the baseline value of the outcome measure,[14] X_{ij} is a vector of demeaned controls,[15] $X_{ij} * T_{ij}$ is their interaction with the treatment indicator,[16] ϕ_d are district fixed effects, and u_{ij} are robust standard errors clustered by the health center catchment area. For child mortality, the unit of observation is the health center catchment area. Following Lin (2013), we use Huber-White heteroskedasticity-robust standard errors. We deal with missing values and outliers as described in Appendix F.1.

We also use Equation 1 to estimate the effects of treatment on the intermediate outcomes described above.

[14] We did not collect baseline values for a subset of index components, as highlighted in Tables 1 and G1. In these cases, the baseline value of the outcome index omits this component. For analyses of treatment effects on these individual components, the baseline value is omitted from the estimating equation.

[15] As specified in our pre-analysis plan, the controls include whether the health center is an HC2, provides delivery services, and has staff houses; whether household members report using the health center within the prior 12 months; the education level of the interviewed household head; and household wealth (calculated as the first component of a principal component analysis of the number of items of 17 assets—cattle, radios, bicycles etc.—owned by the household, as well as three measures of housing quality).

[16] The inclusion of the interaction between the controls and the treatment dummy was not pre-specified. We added this term for the reasons outlined in Lin (2013). Results are substantively unchanged vis-a-vis the original pre-registered specification (see Appendix H.6).

6 Results

As shown in Appendix F.2, our sample is balanced across treatment and control groups with respect to the baseline characteristics of the catchment areas and health centers. Baseline levels of our main and intermediate outcome indices are also balanced. We test for evidence of treatment spillover by comparing outcomes in control health centers that were close to and far from the nearest treated health center, and find no statistically significant differences.

6.1 Main Results

Figure 3 presents the study's main findings. The coefficient plot summarizes the effect of the ACT Health program on the seven main outcome indices as measured at endline, 20 months after the initial treatment. Corresponding regression tables for the outcome indices as well as their components (both standardized and non-standardized) are included in Appendix H.1. The dots represent the estimated treatment effect in standard deviation (sd) units; thin error bars represent the 95% confidence interval; thick error bars the 90% confidence interval.

Our findings allow us to conclude with great confidence that the effects on citizen monitoring, perceived citizen pressure, utilization, health outcomes, and child mortality are either zero or so small as to be substantively not meaningful.[17] We present a formal test for the absence of substantively meaningful treatment effects on these outcomes in Appendix H.12 by using the Two One-Sided Tests (TOST) procedure developed in Schuirmann (1987) with multiple critical values. We can reject even very small treatment effects for citizen monitoring, utilization, health outcomes, and child mortality. For citizen pressure as perceived by health center staff (n=187), we do not have sufficient statistical power to be as confident. However, the negative coefficient in combination with the precisely estimated null effect on citizen monitoring strongly suggest that perceived citizen pressure did not increase as a result of the treatment.

We can also be confident that the intervention *does* have a positive effect on the quality of care provided by health care providers and on patient satisfaction, which increase by 0.070 (95% confidence interval 0.018 to 0.122) and 0.077 sd (95% confidence interval 0.029 to 0.124), respectively. The substantive sizes of these effects are not particularly large, either in absolute terms or relative to secular changes taking place on both outcomes in treatment and control units during the period we study. However, these results do suggest that exposure to treatment changed provider behavior and led to increases in patients' satisfaction with the quality of the care they received.

[17] The upper limit of the 95% confidence interval of the average treatment effect on citizen monitoring is 0.061 sd; on perceived citizen pressure, 0.061; on utilization, 0.070 sd; on health outcomes, 0.051 sd. On child mortality, the (relevant) lower limit of the 95% confidence interval is -0.026 sd. All of these limits are well below the 0.2 sd that is conventionally considered a weak effect Cohen (1992).

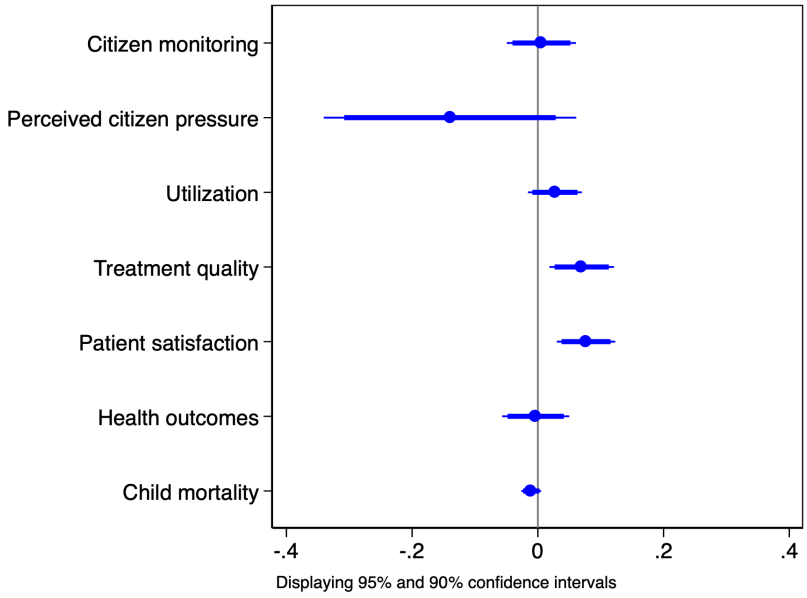


Figure 3: Effect of the full treatment at endline

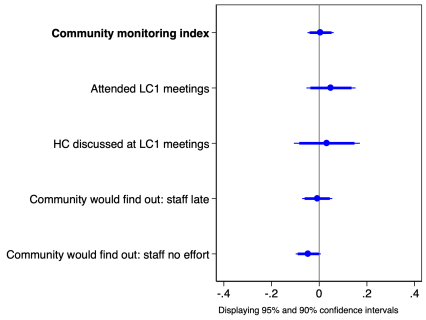
Figure 4 unpacks these index-level results into their components, showing that the null findings with respect to citizen monitoring, perceived citizen pressure, utilization, health outcomes, and child mortality are rooted in statistically insignificant coefficient estimates on nearly every index component. The patient satisfaction findings shown in panel (e), by contrast, are a product of significant, positive estimates on every component but one (which is still positive, but not statistically significant).

The significant results with respect to treatment quality are built on somewhat more mixed component-level findings (see panel (d)). Respondents in households who received their care at treated health centers were more likely to report having had privacy during their most recent exam and having had their diagnosis clearly explained to them (by 1.5 and 2.3 percentage points, respectively). Treated health centers were also 5.9 percentage points less likely to have had stockouts of key drugs during the past three months. Although these three index components are the only ones for which treatment effects reach traditional levels of statistical significance, all of the other components also have positive coefficients, resulting in a significant positive estimate for the index as a whole. This positive index-level effect is robust to several alternative specifications, including (with one exception, discussed below) dropping index components one by one and excluding the three index components measured at the health center level (observed staff presence, cleanliness, and drug availability), whose inclusion in the household-level index artificially inflates their contributions (see Appendix H.6).

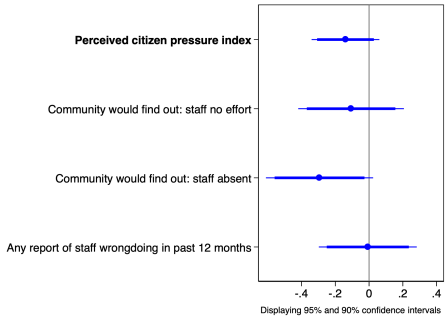
The only index component whose single omission causes the treatment quality index to lose

Figure 4: Treatment effects on outcome indices and their subcomponents at endline

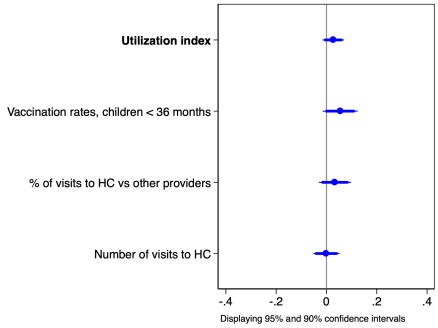
(a) Citizen monitoring



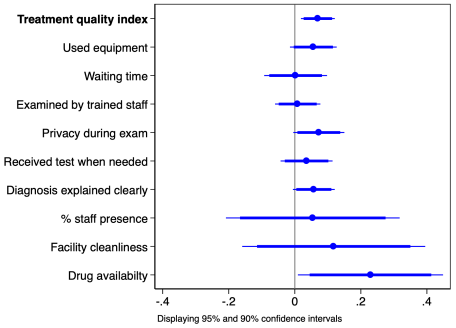
(b) Perceived citizen pressure



(c) Utilization

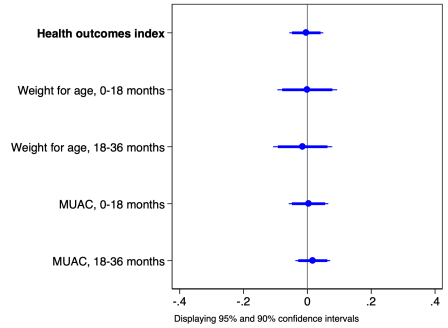


(d) Treatment quality

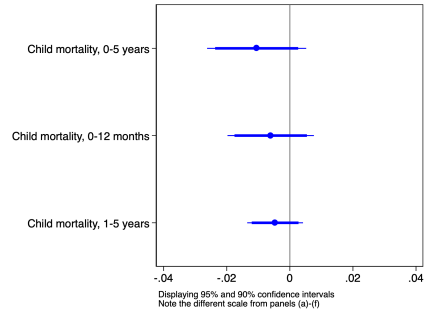


(e) Patient satisfaction

(f) Health outcomes



(g) Child mortality



its statistical significance is drug availability. Drug stockouts are more than just a statistically influential index component, however. The unavailability of essential medicines is a major source of poor health—and even death—in rural Uganda. At the time of our study, Uganda employed a hybrid "push-pull" system under which requested quantities of basic drug supplies are sent to clinics from the National Medical Store (BMAU, 2015; Rwothungeyo, 2016). Hence, exposure to the ACT Health intervention might reduce stockouts via two channels. First, health workers who might otherwise file incomplete or late paperwork requesting drugs might be impelled to project their drug needs more accurately and to request restocking in a timelier manner. Second, interacting with community members and project staff might cause health workers to resist the temptation to steal clinic drugs and sell them to patients at private pharmacies in which they have financial interests. Such drug thefts by clinic staff were a major problem in Uganda at the time of our study (Arinaitwe, 2017), as they had been for some time (Golooba-Mutebi, 2005). Eighty-eight percent of households in our sample cited health workers selling drugs on the side as an important factor in explaining poor health service delivery. The problem was so severe that in 2009 President Museveni established a special agency within State House to combat the phenomenon. The outsized contribution of drug availability to our treatment quality index can therefore be justified by pointing to the importance of reducing drug stockouts to improving health outcomes.

The significant impacts we find on patient satisfaction provide further reason to put stock in our treatment quality findings, as the increase in patients' satisfaction with their care is plausibly a response to the positive changes in health providers' behavior. Since these changes in provider behavior were not associated with measurable changes in actual health outcomes, we can infer—in keeping with a common finding in the medical literature (Kahn et al., 2015)—that patient satisfaction may be rooted in the character of patients' interactions with their health care providers rather than in improvements in health outcomes that these interactions may generate.[18]

It is also noteworthy that the three sub-outcomes exhibiting the strongest treatment effects are the ones most under the control of the health providers themselves, and not dependent on the provision of staff, equipment, or other inputs by the state. This supports the interpretation that exposure to the ACT Health intervention changed the behavior of frontline health workers.

[18] An alternative interpretation is that our findings on patient satisfaction are due less to changes in health provider behavior (which, after all, are substantively quite small) than to the participatory nature of the ACT Health intervention. Other studies have found similar increases in citizen satisfaction following community members' participation in interventions that involve consultation and/or direct participation in decision-making, even when the interventions have no tangible effects on other outcomes (Olken, 2010; Beath, Christia and Enikolopov, 2017).

6.1.1 Midline Results for Main Outcomes

Our findings at midline are generally consistent with those at endline (see Appendix H.9). When we use outcome data measured 8 months after treatment, we find no effects of exposure to the intervention on citizen monitoring, perceived citizen pressure, utilization, health outcomes, or child mortality, and a significant but substantively small (0.06 sd) effect on treatment quality. In contrast to our endline findings, we observe no treatment impacts on patient satisfaction at midline. Exposure to ACT Health thus does not appear to have had shorter-term effects that dissipated by the time of our endline data collection.

6.1.2 Robustness Tests

In addition to the main results shown in Figure 3 and Appendix H.1, we find consistent effects in t-tests (see Appendix H.10) and in various alternative models we pre-specified in our pre-analysis plan.[19] Running the models without control variables or district fixed effects, aggregating all outcome measures to the health center level, and re-specifying our outcome measures as the difference between post-treatment and pre-treatment values all leave our findings substantively unchanged. We also show that our estimated null effects on child mortality are unchanged when we re-analyze our data using at the child level using a Cox proportional hazards model, leveraging the fact that we have child-month data on survival over the course of 36 months for over 10,000 children.

To allay concerns that the number of hypotheses we test might lead us to falsely report statistically significant effects, we provide estimates of treatment impact on all indices and index components both with and without False Discovery Rate adjusted p-values (Benjamini and Hochberg, 1995), based on the comparison families described in Appendix H.11.

Quantile regressions of our seven outcome indices suggest that our estimated treatment effects (both null and positive) are not driven by just parts of the distribution. Our results on utilization, patient satisfaction, and health outcomes are also robust to substituting our main pre-registered outcome measures with alternative indices based on the first component of a principal component analysis. This is important insofar as our pre-registered indices, while deductively sensible, might not perfectly capture the underlying outcomes they were designed to summarize.

[19] These results and all of the robustness tests described in this section are presented in Appendix H.6.

6.2 Subgroup Effects

The evidence presented thus far speaks to the weak impact of ACT Health in the *average* health center and catchment area. However, it is possible that the intervention may have had significant effects in some subsets of units with particular characteristics—for example, those with different baseline levels of service provision, fewer alternative health care options, different exposure to prior NGO health programming, where health center staff are more embedded in the community, or where the community has higher (or lower) baseline levels of efficacy, collective action potential, or ongoing monitoring of health workers. Investigating such sub-group effects can be helpful for better understanding the mechanisms at work and for generating expectations about the likely external validity of the findings in other settings and populations (Banerjee, Chassang and Snowberg, 2017).

The results of our investigation into subgroup effects (described in detail in Appendix H.3) bolster our null findings with respect to citizen monitoring, perceived citizen pressure, utilization, child mortality, and other health outcomes by demonstrating that—with the exception of a handful of key findings discussed in Section 7.3 below—these statistically insignificant results hold across nearly all subsets of health centers and catchment areas.

7 Discussion

7.1 No Evidence of Citizen Pressure

As noted in Section 2, the principal rationale for providing information to citizens and mobilizing them in light of that information is that it will put them in a better position to monitor and apply pressure on underperforming service providers. Yet, our results provide no evidence that exposure to ACT Health generated either of these first stage outcomes. As shown in Figure 4, none of our measures of citizen monitoring or bottom up pressure are significantly different across treated and control communities—and in the two instances where they are (whether community members think they would find out if health workers did not provide effort in caring for patients and whether health center staff think that community members would find out if a health worker did not report for work), exposure to the ACT Health programming is associated with *lower* perceived ability to monitor and apply pressure.

The lack of evidence for the citizen pressure channel is reinforced when we investigate the intervention's impact on intermediate outcomes that theory suggests should be linked with, and may even be prerequisites for, bottom-up pressure. As shown in Figure 5, we find no evidence that exposure to ACT Health had any impact on efficacy, perceived community responsibility for

monitoring health service delivery, the relationship between health care workers and the community, or actions the health center staff may have undertaken to improve transparency vis-à-vis the community (for example, having a suggestion box or posting opening times, duty rosters, or information about services provided and patients' rights).[20]

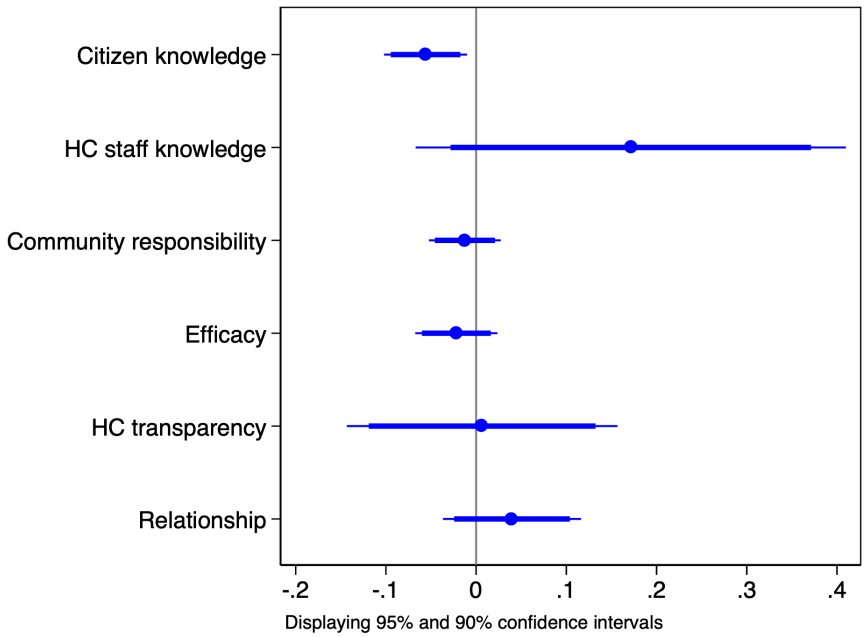


Figure 5: Treatment effect on intermediate outcomes at endline

It is worth considering whether the absence of bottom-up pressure might be due to the fact that ACT Health's mobilization efforts were too weak or reached too few community members. This explanation has at least superficial plausibility: notwithstanding the intervention's stated goal of mobilizing "the community," fewer than 20 percent of households surveyed at midline in treated villages had even heard about the community dialogues or interface meetings. In this respect, ACT Health is little different from most bottom-up community mobilization interventions, where the number of directly treated individuals constitutes only a small fraction of the people living in the treated community, and where awareness of the intervention is often quite low.[21] Moreover, it may

[20]We also find no evidence for positive treatment impacts on the other intermediate outcomes we studied: knowledge of patients' rights and responsibilities among community members and health workers' knowledge of patients' rights and responsibilities. Insofar as citizen knowledge can be thought of as a manipulation check in an information-focused intervention like ACT Health, the significant negative sign on that intermediate outcome measure may appear troubling. We note that the estimate loses statistical significance once a multiple testing adjustment is applied and that the substantive size of the coefficient is, in any case, tiny—corresponding to less than one additional correctly named right or responsibility.

[21]For example, the village meetings that are central to Olken's much celebrated study in Indonesia (Olken, 2007) contained an average of between 45 and 65 people in communities containing roughly 2,500 residents. This implies

be wrong to think that treating every member of the community with information is necessary for bottom-up pressure to materialize. All that may be required is that a sufficiently large number of community members monitor and apply pressure on underperforming service providers (Olson, 1971), and it is reasonable to think that the 100 people attending the average community meeting in ACT Health should have been sufficient to achieve this end.

A more likely explanation lies in the fact that, notwithstanding the theoretical and policy appeal of the citizen pressure channel, bottom-up pressure is extremely difficult to mobilize. Baseline levels of citizen demand for better services are often low, as are expectations of government capacity and responsiveness (Gottlieb, 2016; Hern, 2019). Collective action problems are challenging to overcome (Dasgupta, 2009); citizens' efficacy and sense of responsibility for monitoring frontline providers are often weak (Kruks-Wisner, 2018; Lieberman and Zhou, 2021); formal institutions such as local councils may be moribund and/or corrupt, and therefore unable to support citizens' monitoring efforts; and, compared to the other more immediate problems people face, health care may be insufficiently important to justify the investments in time and energy that the citizen pressure channel assumes community members will be willing to make to try to effect change (Lieberman, Posner and Tsai, 2014). All of these challenges are likely magnified in a semi-authoritarian state like Uganda, where the usual obstacles to grassroots mobilization may be compounded by fear that bottom-up pressure will be interpreted as opposition to the government (Golooba-Mutebi, 2005).[22]

The low capacity of the Ugandan state also plays a role by increasing the plausibility that poor service delivery is beyond the control of the health center staff. Citizens and health care providers are engaged in a principal-agent relationship (Ross, 1973; Holmström, 1979), the crux of which is that the outcomes that the principal (the citizen) observes are also affected by factors outside of the control of the agent (the health worker). Simply supplying community members with information about the outcomes that have been achieved at the health center and how these outcomes compare with district averages (precisely the kind of information the ACT Health intervention provided) does not solve the problem of not knowing whether the observed deficiencies stem from low effort by the health center staff or, as the health workers will certainly claim, from circumstances outside

that only 2-3 percent of the community was directly treated by attendance at a village meeting. In Björkman and Svensson (2009), attendance at community meetings averaged 150 people in catchment areas that contained an average of 2,500 households, implying that only 6 percent of households were directly mobilized. Even when a larger share of the community attends the project meetings, awareness of the intervention can still be low. Banerjee et al. (2010) report the results of an intervention in India in which the average treatment village had 360 households and 108 adults attend the community meetings. Yet despite this comparatively high rate of participation, fewer than 7 percent of households in treated villages had heard about the village education committees that were the central aspect of the intervention.

[22] Although we cannot entirely rule out the impact of such political factors, the fact that we find no differences in the effect of exposure to ACT Health on citizen monitoring or pressure in places where the ruling NRM is relatively strong or weak (see Appendix H.5) suggests that political considerations are unlikely to be driving our findings.

of their control—underfunding, staff shortages, delays in the delivery of drugs and other supplies, or other factors. Such blame shifting is likely to be especially effective in a setting, like Uganda, where health providers and community members both know that the state is weak and unable to reliably provide the inputs that the providers need to care for their patients. This, in turn, may blunt the impetus for mobilizing pressure on the health providers.

These considerations are reinforced by the absence in the setting we study of another key factor stressed in principal-agent models: the ability to sanction. To the extent that information provision works, it may be that it only does where citizens have actual leverage over the frontline service providers they are being encouraged to monitor. In our study context, as in many settings where similar interventions have been deployed, it is difficult to imagine how even highly mobilized citizens would be able to sanction underperforming service providers.[23] Absent the ability to sanction, investments in monitoring may appear futile, and thus not be made. Of course, service providers may alter their behavior in *anticipation* of citizen pressure, even if such pressure never materializes.[24] But such a response is not likely to be sustainable once it is revealed that sanctions are not forthcoming.

7.2 Beyond Citizen Pressure

Although we find no evidence that the ACT Health intervention generated bottom-up pressure, we do find evidence for a modest but statistically significant impact on treatment quality. Consistent with the third channel in Figure 1, it would appear that providing citizens and health care providers with information about health service delivery and patients rights and responsibilities, and working with them to develop action plans in light of this information, had a direct effect on health workers, who responded to these activities by improving the quality of services they provided. The intervention may have been designed to generate effects on service delivery outcomes via bottom-up citizen pressure, but it appears to have bypassed citizens in generating its only measurable impacts.

Additional, suggestive evidence consistent with this *direct effect* channel is provided by chance variation in the participation of local government officials in the program activities.[25] Although ACT Health did not explicitly involve district- or subcounty-level government health officials in its programming, such officials were informed of the intervention and invited to attend the community

[23]Citizen monitoring interventions aimed at shaping the behavior of elected officials, over whom citizens in principle have sanctioning power via their votes, may be more promising. For example, Grossman and Michelitch (2018) find that Ugandan politicians about whom performance information was circulated to voters did in fact perform better, but only in competitive constituencies where citizens possessed real leverage over the politicians. For a less optimistic set of findings about voters' sanctioning power, see the studies presented in Dunning et al. (2019).

[24]Grossman and Michelitch (2018) identify precisely this type of anticipatory response as responsible for the effect of information provision on politicians' behavior in their study.

[25]These analyses were not pre-registered.

and interface meetings, and our implementing partners kept careful records of whether or not such officials did, in fact, attend these meetings (see Appendix I.3). Where they did, the effect of the intervention on treatment quality nearly doubled (see Appendix H.7). The absence of any corresponding effect of the officials' presence on citizen monitoring or perceived citizen pressure suggests that the impact on treatment quality stemmed from health workers' concern that they were being monitored by these superordinate officials, rather than by citizens. This makes sense, since subcounty government health officials have a formal oversight role and, unlike citizens, *do* have the authority to sanction underperforming health care workers. While the fact that the officials' attendance at the meetings was not randomly assigned—and also that these analyses were not pre-registered—cautions against reading too much into this finding, the result is suggestive of the power of top-down, rather than bottom-up, monitoring to improve the performance of frontline service providers.

It may be satisfying from a democratic theory perspective to think that the answer to the problem of poor service delivery lies in giving "power to the people" (to quote the title of Björkman and Svensson's influential paper). But mobilizing citizens to monitor and apply pressure on frontline providers may not be the most powerful strategy for improving the quality of health care and other services. More direct engagement with service providers (channel 3 in Figure 1), combined perhaps with top-down monitoring by government officials, may be a more promising approach.

7.3 The Salience of Baseline Conditions

Thus far, we have focused on the impact of exposure to ACT Health on treatment quality and other outcomes in the average health center. However, when we examine variation in our treatment quality findings across sub-groups of our sample, we find suggestive evidence that they are driven by the more underdeveloped catchment areas: where baseline child mortality rates are higher and baseline levels of treatment quality are lower; where community members have fewer alternative healthcare options; where the community is more rural; where fewer health NGOs are present; and where the health center is an HC2 rather than an HC3 (see Appendix H.4). Although the differences between catchment areas ranking above and below the median on these dimensions are rarely statistically significant, they all point in the direction of stronger effects on treatment quality in less developed contexts. In addition, when we re-run our main analyses in the sub-sample of health centers whose baseline child mortality rates are above the 58th percentile in our distribution (the threshold below which health centers have a child mortality rate of zero), we find that exposure to the ACT Health intervention is associated with a statistically significant reduction in child mortality (see Appendix H.4). These findings suggest that the impact of interventions like the one we study

may be conditional on baseline conditions, with poorer conditions associated with stronger effects.

The plausibility of this "advantages of a lower baseline" thesis is further supported when we compare our findings with those reported in other experimental studies of information-oriented bottom-up pressure interventions in the health sector. As shown in Table 2, which includes all five of the studies discussed in this paper's introduction, treatment impacts diminish as underlying health conditions (proxied by baseline child mortality rates) improve. Significant improvements in child mortality rates and treatment quality may simply be easier to achieve when health conditions are poor.

Table 2: Average treatment effects in information-oriented bottom-up pressure interventions in the health sector in developing countries

Study	Country	Onset of intervention	Baseline U5MR (per 1,000 live births)	Significant treatment effect on			
				Citizen pressure	Utilization	Treatment quality	Child mortality
Arkedis et al. 2021	Indonesia	2015	28	no	no	no	.
Arkedis et al. 2021	Tanzania	2015	59	no	no	no	.
This article	Uganda	2014	59	no	no	yes	no
Mohanan et al. 2020	India (Uttar Pradesh)	2016	78	.	yes	.	no
Björkman & Svensson 2009	Uganda	2004	117	yes	yes	yes	yes
Christensen et al. 2021	Sierra Leone	2012	145	no	yes	no	yes

Notes. Data on baseline U5MR is from World Development Indicators and DHS (for Uttar Pradesh). "Yes" indicates that the study reports a significant positive treatment effect (negative for child mortality) at conventional significance levels for the respective outcome; "no" indicates that the reported effect is insignificant or a precise null; and "." indicates that the study does not report effects on this outcome.

The comparisons highlighted in Table 2 provide suggestive evidence for why lower baseline health conditions may be associated with stronger treatment outcomes. Notwithstanding the emphasis in policy-making circles on the power of bottom-up *citizen pressure*, we see no evidence for this mechanism at any baseline condition—with one notable exception: the Björkman and Svensson (2009) study that initially catalyzed interest in the community monitoring approach. The channel that appears to matter is *utilization*, which increases only in the studies implemented in settings with the lowest baseline conditions and, in two cases, is associated with improvements in health outcomes. This makes sense insofar as health improvements due to increased utilization will dissipate as larger shares of sick people seek professional care at the local health facility. While only suggestive, this finding has important implications for how policymakers seeking to improve health outcomes should deploy their resources in the most disadvantaged settings.

8 Conclusion

When states fail to provide services to their citizens, and where citizens face challenges in applying pressure on the government for service delivery improvements, citizens can in principle apply pressure directly on frontline service providers. An influential idea in development circles is that such direct pressure can be induced by providing citizens with information about service delivery shortfalls and mobilizing them in light of that information. We test this hypothesis by studying a large-scale community health intervention modeled on precisely the received wisdom about the power of information and citizen mobilization to initiate this causal process. While we find evidence for small effects of the intervention on treatment quality, we find no evidence for its impact on utilization or health outcomes (including child mortality), the bottom-line outcomes that policymakers ultimately want to affect. We also find no evidence that the intervention caused citizens to more closely monitor their local health care providers or apply pressure on those who were revealed to be underperforming. Contra the literature, and the motivation for the intervention we studied, we find no evidence suggesting that the link between information provision and provider behavior runs through citizen pressure. Instead, we find evidence that, to the extent that the intervention made a difference, it impacted health providers' behavior directly.

Our finding of weak overall effects are tempered somewhat by the suggestive finding of positive impacts on both health provider behavior and child mortality in the subset of health centers that served communities with the lowest baseline levels of development. Recognizing the importance of such baseline conditions has implications for the extent to which the findings of our own study—and others like Björkman and Svensson (2009) and Christensen et al. (2021), which report somewhat more optimistic findings—are likely to travel to other settings. As shown in Figure 6, baseline health conditions in Uganda in 2004 (the time of the Björkman and Svensson (2009) study) and Sierra Leone in 2012 (the time of the Christensen et al. (2021) study) are well outside the range of most African countries today. Baseline health conditions at the time of our own study in Uganda in 2014, by contrast (and also in Tanzania in 2015, the time of the Arkedis et al. (2021) study), are much closer to the conditions in the modal African country in 2020 (the most recent data available). Our null results (and those in Arkedis et al. (2021)) may therefore be more relevant for the question of whether interventions that provide citizens with information and mobilize them to apply bottom-up pressure on frontline service providers are a viable lever for improving health service delivery in Africa today.[26]

[26]Whether our findings extend beyond the health sector to other areas of service delivery is an important question. On the one hand, the importance of health care to people's well being may make it easier for communities to mobilize to demand service delivery improvements. This would suggest that bottom-up mobilization may be even more challenging to achieve in other sectors. On the other hand, the status differentials between health care providers and citizens may make it harder for citizens to apply bottom-up pressure—an effect that is likely magnified by the

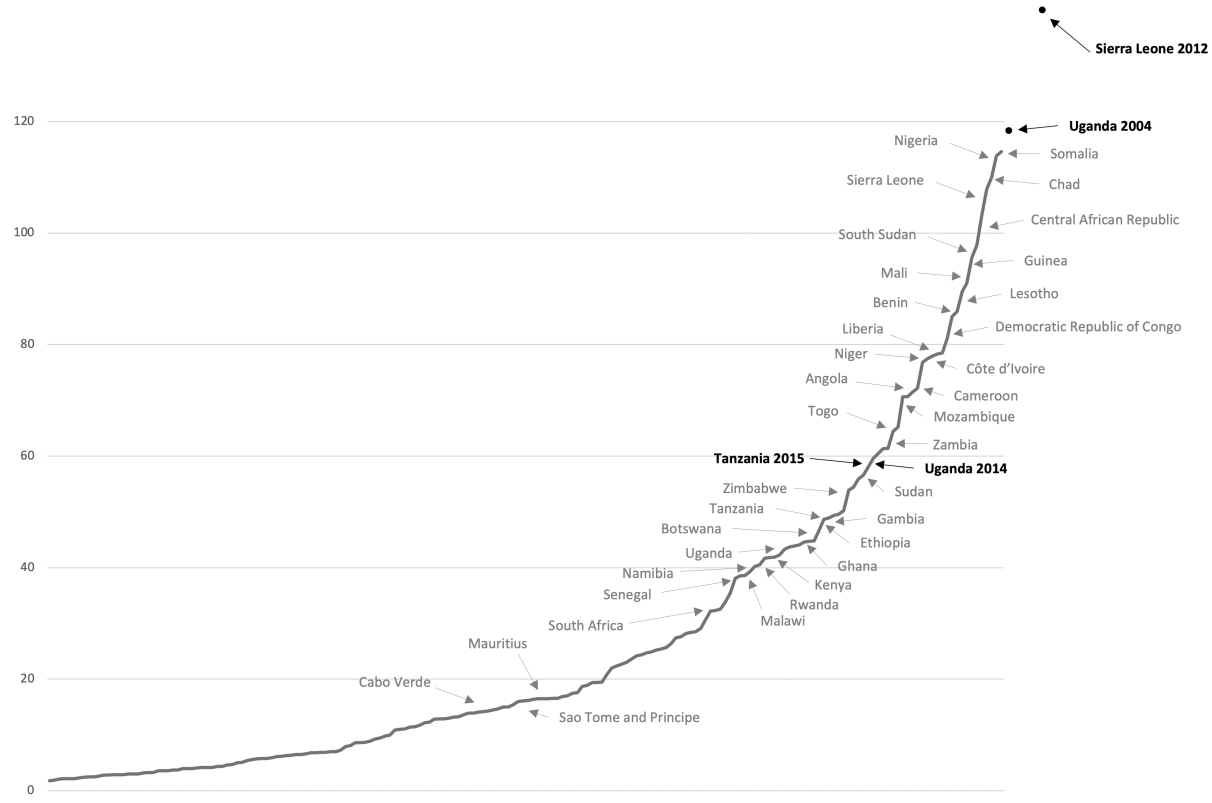


Figure 6: Global under-five mortality rates, ordered lowest to highest, with selected African countries labeled. All figures are from 2020, except those labeled in bold, which correspond to the implementation dates of the studies discussed in the text.

The more focused comparison between our study and Björkman and Svensson (2009) holds a second lesson as well: that interventions that generate strong positive effects at one moment in a country's development may become less effective as conditions improve. We tend to think about external validity as relating to the generalizability of research findings across space. Our findings, viewed alongside those of Björkman and Svensson (2009), underscore the importance of also considering the generalizability of research findings in the same setting *over time*—especially during periods of rapid socioeconomic change, such as those experienced by many low-income countries during the past decade.

importance of health care to people's lives and their fear of losing access to it if they complain too much. This would suggest that prospects for mobilizing citizen pressure for service delivery improvements might be greater in other sectors than is suggested by the results presented here.

References

Afrobarometer Uganda. 2015.

Andrabi, Tahir, Jishnu Das and Asim Ijaz Khwaja. 2017. "Report cards: The impact of providing school and child test scores on educational markets." *American Economic Review* 107(6):1535–1563.

Arinaitwe, Solomon. 2017. "Drug thefts mar hospitals as patients bear brunt." *Daily Monitor, December 3* .

Arkedis, Jean, Jessica Creighton, Dixit Akshay, Archon Fung, Stephen Kosack, Dan Levy and Courtney Tolmie. 2021. "Can transparency and accountability programs improve health? Experimental evidence from Indonesia and Tanzania." *World Development* 142:1 –14.

Bailey, Angela and Vincent Mujune. 2021. "Bottom-up accountability in Uganda: Learning from people-centered, multi-level health advocacy campaigns." Accountability Working Paper no. 8.

Banerjee, Abhijit, Rema Hanna, Jordan Kyle, Benjamin A. Olken and Sudarno Sumarto. 2018. "Tangible information and citizen empowerment: Identification cards and food subsidy programs in Indonesia." *Journal of Political Economy* 126(2):451–491.

Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, Rachel Glennerster and Stuti Khemani. 2010. "Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India." *American Economic Journal: Economic Policy* 2(1):1–30.

Banerjee, Abhijit, Sylvian Chassang and Erik Snowberg. 2017. Decision theoretic approaches to experiment design and external validity. In *Handbook of economic field experiments*, ed. Abhijit Banerjee and Esther Duflo. Elsevier pp. 141–174.

Barber, Ryan M et al. 2017. "Healthcare Access and Quality Index 1990–2015: A novel analysis from the Global Burden of Disease Study 2015." *The Lancet* 390(10091):231–266.

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

Barro, Robert J. 1973. "The control of politicians: an economic model." *Public Choice* pp. 19–42.

Beath, Andrew, Fotini Christia and Ruben Enikolopov. 2017. "Direct democracy and resource allocation: Experimental evidence from Afghanistan." *Journal of Development Economics* 124:199–213.

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the false discovery rate: A practical and powerful approach to multiple testing." *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 289–300.

Besley, Timothy and Anne Case. 1995. "Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition." *The American Economic Review* 85(1):25–45.

Björkman, Martina and Jakob Svensson. 2009. "Power to the people: Evidence from a randomized field experiment of a community-based monitoring project in Uganda." *Quarterly Journal of Economics* 124(2):735–769.

BMAU. 2015. "Continuous stock-outs of medical supplies in Uganda: What are the root causes?" Budget Monitoring and Accountability Unit (BMAU) Briefing Paper 15/15.

Cammett, Melani and Lauren M. MacLean. 2014. *The politics of non-state social welfare*. Cornell University Press.

Chong, Alberto, Ana L. De La O, Dean Karlan and Leonard Wantchekon. 2015. "Does corruption information inspire the fight or quash the hope? A field experiment in Mexico on voter turnout, choice, and party identification." *Journal of Politics* 77(1):55–71.

Christensen, Darin, Oeindrila Dube, Johannes Haushofer, Bilal Siddiqi and Maarten Voors. 2021. "Building resilient health systems: Experimental evidence from Sierra Leone and the 2014 Ebola outbreak." *Quarterly Journal of Economics* 136:1145–1198.

Cohen, Jacob. 1992. "A power primer." *Psychological Bulletin* 112(1):155.

Dasgupta, Partha. 2009. "Trust and Cooperation Among Economic Agents." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1533):3301–3309.

De Kadt, Daniel and Evan S Lieberman. 2020. "Nuanced accountability: Voter responses to service delivery in southern Africa." *British Journal of Political Science* 50(1):185–215.

Duflo, Esther, Rema Hanna and Stephen P Ryan. 2012. "Incentives work: Getting teachers to come to school." *American Economic Review* 102(4):1241–78.

Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, Craig McIntosh and Gareth Nellis. 2019. *Information, accountability, and cumulative learning: Lessons from Metaketa 1*. Cambridge University Press.

Fiala, Nathan and Patrick Premand. 2018. "Social accountability and service delivery: Evidence from a large-scale experiment in Uganda." Working Paper.

Fox, Jonathan. 2015. "Social accountability: What does the evidence really say?" *World Development* 72:346–361.

Golooba-Mutebi, Frederick. 2005. "When popular participation won't improve service provision: Primary health care in Uganda." *Development Policy Review* 23(2):165–182.

Gottlieb, Jessica. 2016. "Great expectations? A field experiment to improve accountability in Mali." *American Journal of Political Science* 60(1):143–157.

Green, Elliott. 2015. "Decentralization and development in contemporary Uganda." *Regional & Federal Studies* 25(5):491–508.

Grossman, Guy and Kristin Michelitch. 2018. "Information dissemination, competitive pressure, and politician performance between elections: A field experiment in Uganda." *American Political Science Review* 112(2):280–301.

Hern, Erin. 2019. *Developing states, shaping citizenship: Service delivery and political participation in Zambia*. University of Michigan Press.

Holmström, Bengt. 1979. "Moral hazard and observability." *The Bell Journal of Economics* 10:74–91.

Kahn, Steven A., James C. Iannuzzi, Nicole A. Stassen, Paul E. Bankey and Mark Gestring. 2015. "Measuring satisfaction: Factors that drive hospital consumer assessment of healthcare providers and systems survey responses in a trauma and acute care surgery population." *The American Surgeon* 81(5):537–543.

Keefer, Philip and Stuti Khemani. 2014. "Mass media and public education: The effects of access to community radio in Benin." *Journal of Development Economics* 109:57–72.

Khisa, Moses. 2019. "Shrinking democratic space? Crisis of consensus and contentious politics in Uganda." *Commonwealth & Comparative Politics* 57(3):343–362.

Kling, Jeffrey R., Jeffrey B. Liebman and Lawrence F. Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica* 75(1):83–119.

Kosack, Stephen and Archon Fung. 2014. "Does transparency improve governance?" *Annual Review of Political Science* 17:65–87.

Kruks-Wisner, Gabrielle. 2018. *Claiming the state: Active citizenship and social welfare in rural India*. Cambridge University Press.

Lieberman, Evan S., Daniel N. Posner and Lily L. Tsai. 2014. "Does information lead to more active citizenship? Evidence from an education intervention in rural Kenya." *World Development* 60:69–83.

Lieberman, Evan S. and Yang-Yang Zhou. 2021. "Self-efficacy and citizen engagement in development: Experimental evidence from Tanzania." *Journal of Experimental Political Science* .

Lin, Winston et al. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics* 7(1):295–318.

MacLean, Lauren M. 2011. "The paradox of state retrenchment in Sub-Saharan Africa." *World Development* 39(7):1155–1165.

Mansuri, Ghazala and Vijayendra Rao. 2013. *Localizing development: Does participation work?* Washington, DC: World Bank.

Martin, Lucy and Pia J Raffler. 2021. "Fault lines: The effects of bureaucratic power on electoral accountability." *American Journal of Political Science* 65(1):210–224.

Mohanan, Manoj, Vikram S. Rajan, Kendal M. Swanson and Harsha Thirumurthy. 2020. "Information and facilitation interventions for accountability in health and nutrition: Evidence from a randomized trial in India." ERID working paper no. 295.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders and Lowell J. Taylor. 2002. "Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment." *American Economic Review* 92(4):850–873.

Namisi, Harriet. 2009. Building pressure from below: Lessons from Uganda. In *From political won't to political will: Building support for participatory governance*, ed. Carmen Malena. Kumarian Press pp. 113–134.

Nannyonjo, Justine and Nicholas Okot. 2013. "Decentralization, local government capacity and efficiency of health service delivery in Uganda." *Journal of African Development* 15(1):125–158.

Olken, Benjamin. 2007. "Monitoring corruption: Evidence from a field experiment in Indonesia." *Journal of Political Economy* 115(2):200–249.

Olken, Benjamin. 2010. "Political institutions and local public goods: Evidence from a field experiment in Indonesia." *American Political Science Review* 104(7):243–267.

Olson, Mancur. 1971. *The logic of collective action*. Harvard University Press.

Ostrom, Elinor. 1990. *Governing the commons*. Cambridge University Press.

Pandey, Priyanka, Sangeeta Goyal and Venkatesh Sundararaman. 2009. "Community participation in public schools: Impact of information campaigns in three Indian states." *Education Economics* 17(3):355–75.

Platas, Melina R and Pia J Raffler. 2021. "Closing the gap: Information and mass support in a dominant party regime." *Journal of Politics* 83(4):1619–1634.

Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, Arminda Alisjahbana and Rima Prama Artha. 2014. "Improving educational quality through enhancing community participation: Results from a randomized field experiment in Indonesia." *American Economic Journal: Applied Economics* 6(2):105–126.

Raffler, Pia J. Forthcoming. "Does political oversight of the bureaucracy increase accountability? Field experimental evidence from an electoral autocracy." *American Political Science Review* .

Ross, Stephen A. 1973. "The economic theory of agency: The principal's problem." *The American Economic Review* 63(2):134–139.

Rwothungeyo, Billy. 2016. "Procurement reforms: Has health service delivery improved?" New Vision, June 1.

Schuirmann, Donald J. 1987. "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability." *Journal of Pharmacokinetics and Biopharmaceutics* 15(6):657–680.

Tripp, Aili Mari. 2010. *Museveni's Uganda: Paradoxes of power in a hybrid regime*. Lynne Rienner.

Tsai, Lily, Ben Morse, Guillermo Toral and Lipovsek Varja. 2019. "Information and accountability: Evidence syntheses of within-government and citizen-government accountability pathways." Transparency and Accountability Initiative.

Tweheyo, Raymond, Catherine Reed, Stephen Campbell, Linda Davies and Gavin Daker-White. 2019. "'I have no love for such people, because they leave us to suffer': A qualitative study of health workers' responses and institutional adaptations to absenteeism in rural Uganda." *BMJ Global Health* 2.

Uganda Ministry of Health. 2017. *Annual health sector performance report, 2016-2017*.

Uganda Ministry of Health. 2018. *National health facility master list: A complete list of all health facilities in Uganda*.

World Bank. 2004. *World Development Report 2004: Making services work for poor people*. Washington, DC: World Bank.

Online Appendix for:
Can Citizen Pressure Be Induced to Improve Public Service Provision?

The following appendices are not intended for print publication.

Appendices

A Interventions Seeking to Generate Bottom-Up Pressure for Service Delivery Improvements

Interventions aimed at encouraging citizens to apply bottom-up pressure for service delivery improvements have been implemented in low income countries around the world. Below, we list a selection of such interventions, along with the research study that evaluated their effectiveness.

- Björkman and Svensson (2009) study the *Power to the People* intervention in Uganda, which aimed to improve health outcomes by providing information to citizens about the performance of local health providers.
- Arkedis et al. (2021) evaluate the impact of the *Transparency for Development* initiative which sought to improve maternal and newborn health outcomes in Indonesia and Tanzania by encouraging civic participation.
- Christensen et al. (2021) report the results of a community monitoring intervention in Sierra Leone designed to improve health outcomes at government-run clinics.
- Fiala and Premand (2018) study the impact of a program in Uganda designed to empower citizens to demand better quality public services by providing information about project performance and social accountability training.
- Banerjee et al. (2010) report the results of the *Pratham* program in India, which attempted to reinvigorate pre-existing citizen-led village education committees to monitor teachers and apply pressure on them to improve their performance. Pandey, Goyal and Sundararaman (2009) study a similar intervention in three Indian states. Pradhan et al. (2014) study a similar initiative in Indonesia.
- Mohanan et al. (2020) study an intervention in Uttar Pradesh, India designed to provide information to citizens about health outcomes and health service entitlements as a means of promoting accountability of health providers and improved service delivery outcomes.
- Lieberman, Posner and Tsai (2014) study the impact of distributing information about students' test scores on the pressure that parents put on educational providers in Kenya. Andrabi, Das and Khwaja (2017) study a similar initiative in Pakistan, as do Keefer and Khemani (2014) in Benin, where the information is provided via community radio broadcasts.
- Freire, Galdino and Miznozzetti (2020) describe the *Tá de Pé* intervention, which aimed to empower citizens to monitor school construction projects in Brazilian municipalities via a mobile phone application.
- Banerjee et al. (2018) report the results of an intervention designed to reduce leakage in a food subsidy program by mailing cards with program information to targeted beneficiaries.
- Ravallion et al. (2015) describe an intervention aiming to improve outcomes in India's NREGA anti-poverty program by showing people a movie that teaches them about their rights under the program.
- Reinikka and Svensson (2005) describe an intervention aiming to reduce leakage of education funds by publicizing information about the size of grants that schools are meant to receive.
- Olken (2007) studies a program designed to put citizens in a stronger position to monitor infrastructure spending in Indonesia.

- Barr et al. (2012) study the impact of a program in Uganda designed to improve the accountability of teachers and school administrators by facilitating community-based monitoring.

B Ethical Considerations

IRB approvals for our research were secured at IPA (Protocol ID: 0497) and at the Uganda National Council for Science and Technology (UNCST) (Protocol ID: ARC157). More general approval for the project was also received from UNCST itself (Protocol ID: SS3559) and from the Office of the President, Uganda.

As indicated in the materials we submitted to these bodies, we took steps to address the major ethical challenges associated with our research. We took these challenges to be a) that providing information about poor service delivery and mobilizing citizens in light of that information might generate conflict between health care providers and citizens, b) that interviewing health care providers might have taken them away from their duties and delayed or otherwise compromised the health care of their patients, and c) that participants might feel pressured to participate in the study unwillingly and/or that their anonymity might be compromised by our data storage protocols.

With respect to the first concern, our implementation team underwent extensive training regarding how to present the information and conduct the dissemination and interface meetings in a manner that generated constructive discussion rather than conflict. With respect to the concern that our questioning of health workers might distract them from performing their duties, enumerators were instructed to interrupt the survey when a health worker was busy and to resume when she was again available.

With respect to concerns about survey respondents' informed consent and privacy, participation in the study was voluntary and all respondents need to have given their informed consent in order to participate. Respondents received a small compensation for their time in the form of a long bar of soap (worth 1,500 UGX, which about corresponds to the average rural income for the 90-minute duration of the survey).[27] All data was collected electronically on password protected PDAs, using SurveyCTO, an ODK based platform. Data was uploaded to an encrypted server on a daily basis, networks permitting, and then stored on password protected computers using encryption and removing all personally identifying information from the datasets.

Our protocols for the organization of the community and interface meetings were designed to ensure that participants included a diverse cross-section of citizens served by the health center being studied, as described in Appendix I.

We have no reason to believe, and our data reveal no evidence, that the intervention differentially benefited or harmed particular groups.

Quite apart from these considerations, one ethics concern sometimes raised with RCTs is lack of consultation with people on the ground. Our goal was to ensure exchange throughout the lifespan of the project—both for ethical reasons and because we believe that local knowledge is critical for ensuring research quality. Towards this end, throughout the study period, from initiation to dissemination of results, we conducted many consultations with the implementing organizations and the Ministry of Health.

We also conducted two rounds of qualitative interviews. In 2014, we conducted key informant interviews and focus group discussions in seven health centers, fourteen villages in their catchment areas, and district officials in three districts that had all been part of the original P2P intervention. Our objective was to better understand the mechanisms through which the original intervention may have affected health outcomes, thus informing the design of the baseline questionnaire. Our research team separately interviewed the District Health Officer or Chief Administrative Officer, health center staff, as well as community members

[27] According to the Ugandan National Household Survey from 2016, the average monthly household income in rural areas was 303,000 UGX with an average household size of five persons (including children) at the time.

(ensuring balance with regard to age and gender), members of the village health team, and elected village chairpersons. In 2016, we conducted another round of focus group discussions in eight treatment health center catchment areas across three different districts. Our objective was to understand how participants had experienced the GOAL intervention and the reasons they identified for its success or lack thereof and to use these insights to inform the preanalysis plan and questionnaire for the endline.

A two-day workshop with representatives from all implementing organizations in 2016 was held to make sense of the midline findings and inform the pre-analysis plan and design of the questionnaire for the endline.

This focus on local knowledge is also inherent to design of the intervention, which aimed to create space for dialogue and negotiation between citizens, service providers, and potentially local government officials.

C Procedures to Ensure Data Quality

The ACT Health project team followed a standard set of procedures and processes developed by IPA over the years to manage large-scale academic research projects. These protocols include specific requirements for adhering to human subjects' regulations, developing survey instruments, fielding data collection teams, implementing data quality reviews, and producing and cleaning datasets for analysis.

The ACT Health project submitted research protocols for the three waves of data collection to both a local IRB committee (Mildmay/MUREC and UNCST) and to IPA's internal IRB review committee (#2127). The project team worked closely with local authorities and received approvals for its work from the Office of the President, the Ugandan Ministry of Health, and from the chief administrative officers and district health officers in each of the 16 districts in which the project was implemented.

All personnel who handled the data and identified surveys in the field were required to obtain IRB training certificates. All field officers (including surveyors) signed confidentiality forms and administered informed consent to every respondent.

To minimize concerns over social desirability bias and Hawthorne effects, we took great care to decouple the intervention and the data collection exercise in the perception of respondents and blinded survey team members to treatment status. We can thus rule out the concern that members of the survey team might have sought to validate the program's objectives through the way they asked questions or recorded observations about the clinics they visited.

Data collection was undertaken in four distinct steps by teams headed by a research associate and consisting of field managers, team leaders, enumerators (health center, household, and anthropometric), mobilizers, trackers, and auditors.

1. Mobilization: A team of trained mobilizers contacted targeted households a day prior to the start of data collection to alert them to the survey work to come and to document the locations in which surveys would be administered. The conditions of the studied health center was assessed, staff attendance was recorded, and drug supplies were checked during a surprise visit to the health center on the same day. In-charges were notified to prepare the relevant records for the enumeration team's visit the following day.

2. Enumeration: The enumeration team completed four different surveys.

   (a) Household Survey: Household enumerators completed an average three to four surveys a day. The household survey took about one and a half hours to complete. Enumerators were instructed to interview the female head of the household. In the event that the female household head was

no longer living in the house (at midline or endline), enumerators were instructed to follow the decision tree below to interview the correct person. During the interview, enumerators were asked to assess the number of children under five present in the household and to complete a form that household members would later give to anthropometric enumerators during their visits. At the end of the interview, contact forms were given to the household with instructions on how to report any comments to the HR management or IRB committee.

(b) Anthropometric Survey: Anthropometric enumerators were specially trained to measure the height, weight and middle-upper arm circumference of all children in the household under 5 years old. Anthropometric enumerators were in charge of collecting the form left by household enumerators to ensure that the household survey was administered.

(c) Health Center Survey: Each health center enumerator (three per team) completed one survey a day. The survey consisted of interviewing the in-charges (at endline, a survey of other staff members was also added), assessing the quality of the health center, and collecting administrative reports (HMIS, PHC funds, etc.). Health center enumerators' visits were announced but could not take place the day of an immunization campaign, when health center staff were occupied. Health center enumerators were instructed to take pictures of administrative report pages with their PDA to prevent misreporting. These pictures were deleted each evening by the field team leaders.

(d) LC1 Survey: Team Leaders were responsible for the LC1 surveys, which involved interviewing the LC1 chairman about the characteristic of the village (rural/urban), its social cohesion, the political affiliations of officials, and other topics.

3. Tracking: Household that could not be found on the day of the enumeration were tracked by a team of trackers who were also trained to do the anthropometric survey. Tracking sheets were given to trackers by field managers after receiving approval from the research associate, following the decision tree below.

4. Auditing: Auditors performed back checks and spot checks (with field managers) on daily basis. Auditors received auditing sheets from the research associate once household data collection was finalized. They reported the findings of their investigations and handled their surveys directly to the research associate.

The ACT Health survey team followed a set of standard operating procedures to ensure high quality data collection. These included:

- High quality training for everyone involved in the data collection: A total of four different teams of 75 enumerators worked in the 16 different districts. Mobilizers, enumerators, auditors and team leaders went through one-week trainings before being selected, including a soft launch to put in practice what they had learned.

- High-frequency checks: Specific survey questions that were susceptible to typos or incoherence were audited every evening by the research associates and field managers to ensure data quality. Daily feedback was provided to enumerators based on the findings from the monitoring, back checks, and high-frequency checks.

- Back checks: During the survey itself, data auditors re-surveyed a random sub-sample of survey participants (on a portion of the survey) to monitor enumerators' performance and to confirm that enumerators were interviewing the correct respondents. Field managers monitored their teams and accompanied each enumerator at least once every week.

- Monitoring and supervision of data collection: Research associates traveled with the survey team throughout the five months of the data collection and across the 16 different districts to supervise the data collection process. Field managers were present in the villages in which enumeration was ongoing on daily basis in order to monitor the data collection and perform spot checks on randomly selected enumerators. Principal investigators were updated in weekly calls about the data collection and were consulted to solve problems as they arose.

D Deviations from the Pre-Analysis Plan

We registered the pre-analysis plan (PAP) in 2016 prior to having access to endline data, and registered an updated version in 2017 prior to analyzing endline data. Below is a list of deviations from the preregistered updated pre-analysis plan. Edits in 2017 were relatively minor and are clearly indicated in the updated PAP.

1. The PAP listed just two anthropometric outcome measures: weight-for-age and MUAC. However, subsequent discussions with public health experts suggested that we should also collect child height to be able to measure height-for-age (stunting), which we now report in the paper.

2. Following Lin, Green and Coppock (2016), we include the interaction of our standardized covariates with the treatment indicator.

3. In deviation from our pre-analysis plan, we added an eleventh subgroup. While we had pre-specified replicating our analyses in the subsample of health centers within one standard deviation of the child mortality level in (Björkman and Svensson, 2009); for greater generalizability, we are instead assessing heterogeneous treatment effects by different cutoffs of child mortality.

4. Further, heterogeneous effects by presence of local government officials (as discussed in section 7.2) were not pre-specified.

5. We dropped the indicator variable for whether a health center has staff houses from the vector of controls since it co-varied with another control, health center level, and thus dropped from the analyses.

6. In response to reviewer comments, we elevated citizen monitoring from an intermediate to a main outcome and added perceived citizen pressure as an additional main outcome of interest.

7. We had pre-specified a further "tertiary analysis" which we do not deem sensible in light of the results: to test the hypothesis of non-effect on any of the four outcomes using the nonparametric combination approach, as proposed by Caughey, Dafoe and Seawright (2017). Given our mostly null findings with the pre-specified primary specification, we no longer see this as a value addition.

8. As a tertiary analysis and further robustness check, we had prespecified running the main specification dropping units for which we have evidence that the quality of implementation was severely compromised. This is not applicable, since the detailed monitoring data from the implementation team suggests that no such health centers exist.

9. Anticipating treatment effects on our main intervention and the desirability of teasing apart pieces of the bundled treatment, we pre-registered and implemented a factorial design, which we report on in Appendix XX.

E Measuring Child Mortality

Child mortality is calculated at the health center level using the synthetic cohort life table approach, as used by the DHS (Croft, Marshall and Allen, 2018). Based on data on the months of birth and death of all children under the age of five in our sample households, we are able to calculate the probability of death for children aged 0-12 months, 1-5 years, and under 5 years. The mortality rates we report here are the estimated probability of a child in our sample dying before reaching a given age, expressed as a rate per 1,000 live births. The method of calculation is detailed in Appendix E.[28] For this measure alone, lower values imply a more positive outcome. The components of the five main outcome indices, along with their mean values at baseline, are presented in Table 1.

We use three different approaches to measure child mortality rates: a *synthetic cohort approach* to calculate mortality rates per health center catchment area, which is similar to the method used in the Demographic and Health Surveys (DHS), a *child-level indicator* for whether a specific child is alive or dead in a given month, and the share of children who died in a catchment area, which mirrors the *vital statistics approach* used in Björkman and Svensson (2009). We describe each in turn.

Originally, we planned to use a vital statistics approach to measure mortality rates, since this was the main approach used in Björkman and Svensson (2009).[29] The vital statistics method uses a simple ratio of deaths under a certain age to live births during a recall period (UNDG, 2003). However, we updated our pre-analysis plan to prioritize the synthetic cohort life table approach because it offered a more precise measure of mortality. The difference in the data required for each method is that the synthetic life table approach requires the dates (month and year) of birth and death for every child that died during the recall period. In contrast, the vital statistics approach only requires asking if any child under the age of five had died in the last 12 months and the age they were when they died.

At endline, we collected the month of birth and, if applicable, death, also retrospectively for all children recorded during baseline and midline. Since the birth and death of children in the family is a very salient event, we are in this case not concerned about recall bias. To the contrary, the second, retrospective round of data collection of the month and age of death proved to be a helpful verification exercise, during which it became evident that a considerable share of the children that had been reported as having died in the past 12 months during baseline or midline had in fact died much earlier.

With this life table data, we are also able to use an even more nuanced measure of child mortality *at the child level*. Since we have the month of birth and, if applicable, death, for all 20,598 children in our sampled households who were ever under the age of five or unborn at baseline and still lived in the household (if alive) at either midline or endline, we are able to create a panel dataset indicating whether each child is dead or alive in a given month over the course of the 36 month study period.[30] This dataset, in turn, allows us to run child-level survival analyses using a Cox proportional hazards model (Cox, 1972), an estimation approach widely used for the analysis of survival rates (Rosner, 2015). We show results from this approach in the appendix.

Synthetic cohort life table

The *synthetic cohort life table* approach is used in the Demographic and Health Surveys (DHS), as described

[28] As also discussed in Appendix E, we supplemented this health center-level synthetic cohort data with a child-level measure that leverages the detailed child-month level retrospective data we collected at endline. Results for these child-level estimates are shown in Appendix H.

[29] Björkman and Svensson (2009) also use an alternative measure, a binary indicator for child death during the recall period.

[30] Twelve month recall period prior to the baseline, 12 months between baseline and midline, and 12 months between midline and endline.

in Croft, Marshall and Allen (2018) and Rowland (2003).[31] The approach calculates the probability of dying before a certain age (expressed per 1,000 births) by dividing the total number of deaths under that age by the total number of child years of exposure to the risk of dying.

We begin by calculating the age (in months) of each child in our sample, for each calendar month in the period of investigation. In our case, that is 12 months prior to the date of the first baseline survey (August 2013) until the date of the last endline survey (December 2016). In the case of death, the age counter stops on the date of death. For each child that died during the study period, we create a binary variable indicating the month during which the child died. On this basis, we can calculate for each calendar month and age (in months) the number of children per health center catchment area who died in a given calendar month, and who were alive. We then sum the number of children who were alive at a given age in a given HC catchment area across calendar months (the denominator); as well as the number of children who died at a given age in a given HC catchment area across calendar months (the numerator). We calculate the ratio to arrive at the age and health center specific mortality rate. We then calculate the age and health center specific survival rate by subtracting the mortality rate from one. Finally, we calculate the overall survival rate in a given age bracket by multiplying the individual age-specific survival rates across the relevant ages (e.g. 0-59 months for the under-five survival rate), and arrive at the overall mortality rate by subtracting the survival rate from one.

Vital statistics approach

The vital statistics approach calculates child mortality as the ratio of dead children in a given age bracket over all children in a given age bracket, dead or alive, calculated per catchment area.

Child-level indicator

A 36-month panel dataset for 20,598 children (some of whom are born into the panel at a later stage or age out of it), indicating for each month whether a child is dead or alive.

F Technical Details and Validity Checks

F.1 Missing Values and Outliers

As specified in our pre-analysis plan, we remove outliers by capping (top-coding) unbounded variables at the 99th percentile of the observed values in our data. To deal with missing values on our covariates, we adopt the following approach as pre-specified and proposed by Lin, Green and Coppock (2016): If no more than 10% of the covariate's values are missing, we recode the missing values to the overall mean. If more than 10% of the covariate's values are missing, we include a missingness dummy as an additional covariate and recode the missing values to zero. We deal with missing values on our outcome measures by setting them equal to the means of the respective treatment arms (as pre-specified and proposed by Kling, Liebman and Katz (2007)).

[31] The approach has also been used to measure child mortality rates in a randomized evaluation of a community health promoter program in Uganda implemented by Living Goods and BRAC (Björkman Nyqvist et al., 2019)

F.2 Attrition, Balance, and Spillover

Table F1: Attrition Across Treatment Arms

	(1) Full treatment	(2) Information treatment	(3) Interface treatment	(4) Control	(5) P-value difference (1)-(4)	(6) P-value difference (1)&(2)-(3)&(4)	(7) P-value difference (1)&(3)-(2)&(4)
Reinterview rate, baseline to midline	0.947	0.940	0.950	0.944	0.68	0.44	0.16
Reinterview rate, baseline to endline	0.956	0.949	0.958	0.957	0.88	0.29	0.40

Notes. Numbers reported correspond to the average of household participation at health center level for the four different arms. The unit of observation is the health center catchment area. Columns (5)-(7) report the p-values of two-sided t-tests comparing re-interview rates in columns (1) and (4), (2) and (4), and (3) and (4), respectively. The full sample is composed of 376 health centers. At baseline, 379 health centers were surveyed but 3 dropped of the selected sample due to external factors (moved to another location, closed due to district decision or structural damages from flood).

Table F2: Spillover analysis

	Mean of close control HC	Mean of far control HC	P-value difference
A. Difference Midline - Baseline levels of key outcome indices			
Citizen monitoring	-0.023	0.044	0.371
Perceived citizen pressure	0.036	-0.075	0.653
Utilization	0.006	-0.011	0.652
Treatment quality	-0.082	0.171	0.000
Patient satisfaction	0.012	-0.021	0.525
Health outcomes (HH level averaged by HC catchment area)	-0.070	-0.034	0.621
Health outcomes (Child level averaged by HC catchment area)	0.009	0.010	0.988
Child Mortality	0.010	0.029	0.286
B. Difference Endline - Baseline levels of key outcome indices			
Citizen monitoring	-0.024	0.048	0.358
Perceived citizen pressure	0.046	-0.080	0.538
Utilization	-0.009	-0.008	0.995
Treatment quality	-0.049	0.111	0.008
Patient satisfaction	-0.002	0.004	0.899
Health outcomes (HH level averaged by HC catchment area)	0.033	0.008	0.714
Health outcomes (Child level averaged by HC catchment area)	0.045	-0.074	0.430
Child Mortality	0.011	0.049	0.023
N	64	31	

Notes. Difference in means test comparing changes in main outcomes among control health centers that are relatively close to a treatment health center and those that are relatively far away. *Close* indicates that the distance to the nearest treatment health center is below the 67th percentile (5.2 miles) among all closest control/treated pairs in our sample, *far* indicates that it is above. To increase the number of relatively close pairs, we define "treated" health centers as those receiving either the full ACT Health intervention or one of its subcomponents for this analysis. The dependent variable is defined as the change from baseline to midline (panel A) and the change from baseline to endline (panel B), respectively.

Table F3: Balance Across Treatment Arms

	(1) Full treatment	(2) Information only	(3) Interface only	(4) Control	(5) P-value difference (1) - (4)	(6) P-value difference (1) & (2) - (3) & (4)	(7) P-value difference (1) & (3) - (2) & (4)
A. Characteristics of catchment area							
Avg. distance of households to HC	0.98	0.97	0.85	1.34	0.28	0.52	0.17
Avg. household wealth in catchment area	0.01	-0.02	-0.00	0.02	0.86	0.79	1.00
Log pop. density in 3km radius around HC	8.46	8.62	8.64	8.64	0.23	0.34	0.45
Avg. level of education of household head	7.59	7.45	7.43	7.66	0.73	0.86	0.73
Share of households that received a VHT visit in the last 12 months	0.47	0.48	0.48	0.48	0.82	0.76	0.98
Share of households that declared NGOs activity in their village	0.15	0.13	0.14	0.15	0.87	0.80	0.63
B. Characteristics of HC							
Share providing delivery services	0.64	0.54	0.62	0.66	0.76	0.34	0.61
Share having staff houses	0.82	0.80	0.78	0.82	0.92	0.85	0.73
# of trained medical staff	6.26	6.45	6.34	7.17	0.18	0.40	0.28
Share having piped water	0.14	0.12	0.13	0.15	0.91	0.77	0.91
Share having electricity (grid or solar)	0.75	0.67	0.71	0.75	0.97	0.71	0.68
Avg. distance to nearest other government HC in district	3.90	3.63	3.97	4.09	0.49	0.17	0.70
C. Baseline levels of key outcome indices							
Community monitoring	0.79	0.77	0.78	0.77	0.23	0.93	0.10
Perceived citizen pressure	0.68	0.72	0.73	0.66	0.54	0.71	0.59
Utilization	5.24	5.27	5.52	5.43	0.41	0.17	0.86
Treatment quality	12.23	12.68	12.11	11.93	0.59	0.29	0.76
Patient satisfaction	0.78	0.79	0.78	0.78	0.61	0.82	0.35
Health outcomes (HH level)	1.07	1.12	1.07	1.05	0.38	0.10	0.58
Health outcomes (Child level)	2.01	2.04	2.07	1.97	0.61	0.96	0.53
Child Mortality (deaths per 1,000 live births)	50.34	50.26	41.40	39.39	0.26	0.13	0.88
D. Baseline levels of intermediate outcome indices							
Citizen knowledge	0.20	0.20	0.21	0.20	0.54	0.14	0.55
Health worker knowledge	0.45	0.44	0.42	0.43	0.39	0.26	0.98
Efficacy	0.58	0.56	0.57	0.59	0.11	0.09	0.61
Community responsibility	0.23	0.23	0.23	0.24	0.24	0.23	0.68
Relationship between health workers and community	0.80	0.80	0.80	0.82	0.16	0.24	0.47
Health center transparency	0.14	0.13	0.12	0.14	0.88	0.52	0.66
N	92	92	97	95			

G Intermediate Outcomes

Table G1: Intermediate outcome indices and their components

	Mean
Citizen knowledge	
Share of patients' rights that household head is able to name correctly	10.20%
Share of patients' responsibilities that household head is able to name correctly	30.12%
Share of services offered at health center that household head is able to name correctly	64.76%
Health center staff knowledge	
Share of patients' rights that health center staff is able to name correctly	31.97%
Share of patients' responsibilities that health center staff is able to name correctly	54.93%
Efficacy	
Whether household head thinks she has "a lot"/"some" power to improve quality of health care at local HC	33.98%
Whether household head thinks she would be able to pressure a health worker to exert better effort	62.79%
Whether household head thinks she would be able to pressure a health worker to report to work on time ◇	61.93%
Whether household head thinks she has "a lot"/"some" influence in making village a better place to live ◇	33.65%
Whether household head agrees that "people like you have a say about how the government provides health care to your community"	82.32%
Whether household head agrees that "people like you have a say about how health facilities provide health care to your community"	81.18%
Community responsibility	
Whether household head thinks she is responsible for making sure health workers come to work and provide high quality health services	45.15%
Whether household head thinks community members are responsible for making sure health workers come to work and provide high quality health services	1.24%
Relationship between health care workers and the community	
Whether household members report being "satisfied"/"very satisfied" with relationship with health center staff	73.67%
Whether household members say they trust the workers at the health center	60.10%
Whether health center staff report being "satisfied"/"very satisfied" with their relationship with the community	90.62%
Whether household members did not say that the health center staff would "refuse to see me" or "behave hostilely toward me" if they had a complaint about the quality of services at the health center and decided to talk to the facility staff	97.69%
Health center transparency	
Whether a poster showing health center's opening/closing hours was visible during unannounced visit	2.78%
Whether a staff duty roster was displayed publicly during unannounced visit	20.31%
Whether a suggestion box was present during unannounced visit	6.14%
Whether information was posted listing services provided at the health center during unannounced visit	33.14%
Whether information was posted about patients' rights and responsibilities during the announced visit	3.46%

◇ Baseline values for this variable were not collected; values shown are from the control group. The baseline index omits these components.

H Supporting Tables

H.1 Main Outcomes

The following tables show regression results for the main outcome indices (as summarized in Figure 3) and their components. The index components are shown first in standardized and then in non-standardized forms.

Table H1: Main outcomes: Averaged z-score indices

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**	(7) **Child mortality**
Full treatment	0.006	-0.140	0.027	0.070***	0.077***	-0.003	-0.011
	(0.028)	(0.102)	(0.022)	(0.026)	(0.024)	(0.027)	(0.008)
Constant	0.003	0.015	-0.020	0.000	-0.002	-0.510***	0.061***
	(0.019)	(0.064)	(0.016)	(0.020)	(0.018)	(0.022)	0.006
N	7,288	187	7,288	7,288	7,288	4,930	187
R^2	0.097	0.152	0.230	0.102	0.043	0.112	0.197
P-value (Full treatment = 0)	0.838	0.171	0.213	0.008	0.001	0.900	0.188
Adjusted p-value (FT)	0.900	0.298	0.298	0.028	0.010	0.900	0.298

Notes. Estimates from Equation 1 comparing outcomes between the full treatment arm and the control group. The unit of observation in columns (1) and (3)-(5) is the household, in column (6) it is child, and in columns (1) and (7) the health center catchment area. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at health center level. The row *P-value (Full treatment = 0)* shows p-values for a Wald test of the null hypothesis that the coefficient on *full treatment* is equal to zero. Adjusted p-values (FT) refer to p-values of the coefficient on *Full treatment* which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table H2: Citizen monitoring index – Subcomponents

	(1) **Citizen monitoring index**	(2) Attended LC1 meetings	(3) HC discussed at LC1 meetings	(4) Community would find out: staff late	(5) Community would find out: staff no effort
Full treatment	0.006	0.049	0.032	-0.008	-0.046*
	(0.028)	(0.052)	(0.070)	(0.032)	(0.027)
Constant	0.003	0.044	0.026	0.008	0.005
	(0.019)	(0.041)	(0.049)	(0.024)	(0.019)
N	7,288	7,288	7,288	7,288	7,288
R^2	0.097	0.105	0.086	0.035	0.026
Adjusted p-value (FT)	0.900	0.692	0.796	0.796	0.354

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2)-(5), which are z-scores of the following variables: (2) a dummy variable whether household members report having attended at least one LC1 meeting during the last 12 months; (3) a dummy variable whether the local health center was discussed at the most recent LC1 meeting; (4) a Likert-scale variable of whether the community would find out if a staff were regularly late or (5) extended no effort. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table H3: Citizen monitoring index non-standardized – Subcomponents

	(1) **Citizen monitoring index**	(2) Attended LC1 meetings	(3) HC discussed at LC1 meetings	(4) Community would find out: staff late	(5) Community would find out: staff no effort
Full treatment	0.006	0.013	0.011	-0.003	-0.019*
	(0.028)	(0.014)	(0.023)	(0.013)	(0.011)
Constant	0.003	0.764***	0.426***	0.792***	0.796***
	(0.019)	(0.022)	(0.021)	(0.010)	(0.007)
N	7,288	7,288	7,288	7,288	7,288
R^2	0.097	0.105	0.086	0.035	0.026
Mean control group endline	0.000	0.833	0.498	0.789	0.793
Mean control group baseline	0.000	0.888	0.648		

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2)-(5), which are z-scores of the following variables: (2) a dummy variable whether household members report having attended at least one LC1 meeting during the last 12 months; (3) a dummy variable whether the local health center was discussed at the most recent LC1 meeting; (4) a Likert-scale variable of whether the community would find out if a staff were regularly late or (5) extended no effort. *** p<0.01; ** p<0.05; * p<0.10

Table H4: Perceived citizen pressure index – Subcomponents

	(1) **Perceived citizen pressure index**	(2) Community would find out: staff no effort	(3) Community would find out: staff absent	(4) Any report of staff wrongdoing in past 12 months
Full treatment	-0.140	-0.107	-0.294*	-0.008
	(0.102)	(0.159)	(0.161)	(0.147)
Constant	0.015	0.028	-0.005	0.020
	(0.064)	(0.107)	(0.103)	(0.101)
N	187	187	187	187
R^2	0.152	0.142	0.185	0.140
Adjusted p-value (FT)	0.298	0.753	0.210	0.959

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. The dependent variable in column (1) is the averaged z-score index of the outcomes presented in columns (2)-(4). The latter are (2) the perceived probability that community members would find out if a health worker did not provide effort in caring for patients, or (3) did not report for work, and (4) whether there had been any community complaints about wrongdoing by health center staff in the past 12 months. Index components are measured through the health center survey. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** $p<0.01$; ** $p<0.05$; * $p<0.10$

Table H5: Perceived citizen pressure index non-standardized – Subcomponents

	(1) **Perceived citizen pressure index**	(2) Community would find out: staff no effort	(3) Community would find out: staff absent	(4) Any report of staff wrongdoing in past 12 months
Full treatment	-0.140	-0.024	-0.108*	-0.004
	(0.102)	(0.036)	(0.059)	(0.071)
Constant	0.015	0.977***	0.830***	0.382***
	(0.064)	(0.059)	(0.052)	(0.049)
N	187	187	187	187
R^2	0.152	0.142	0.185	0.140
Mean control group endline	0.009	0.952	0.842	0.381
Mean control group baseline	0.002	0.840	0.470	

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. The dependent variable in column (1) is the averaged z-score index of the outcomes presented in columns (2)-(4). The latter are (2) the perceived probability that community members would find out if a health worker did not provide effort in caring for patients, or (3) did not report for work, and (4) whether there had been any community complaints about wrongdoing by health center staff in the past 12 months. Index components are measured through the health center survey. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

Table H6: Utilization index – Subcomponents

	(1) **Utilization index**	(2) Vaccination rates, children<36 months	(3) % of visits to HC, vs. other providers	(4) Number of visits to HC
Full treatment	0.027	0.054	0.034	-0.001
	(0.022)	(0.035)	(0.032)	(0.027)
Constant	-0.020	-0.011	-0.001	-0.017
	(0.016)	(0.025)	(0.023)	(0.020)
N	7,288	4,212	7,288	7,288
R^2	0.230	0.057	0.178	0.284
Adjusted p-value (FT)	0.298	0.370	0.446	0.981

Notes. Estimates are derived from from Equation 1, comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects as well as demeaned baseline covariates and their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the outcomes presented in columns (2)-(4). The latter are z-scores of (2) vaccination rates of children under 36 months, (3) share of visits to the designated health center versus other providers, (4) number of visits to the designated health center. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10.

Table H7: Utilization index – Non-standardized subcomponents

	(1) **Utilization index**	(2) Vaccination rates, children<36 months	(3) % of visits to HC, vs. other providers	(4) Number of visits to HC
Full treatment	0.027	0.022	0.007	-0.008
	(0.022)	(0.014)	(0.007)	(0.342)
Constant	-0.020	0.739***	0.231***	9.128***
	(0.016)	(0.017)	(0.007)	(0.316)
N	7,288	4,212	7,288	7,288
R^2	0.230	0.057	0.178	0.284
Mean control group endline	-0.013	0.787	0.326	15.327
Mean control group baseline	-0.008	0.755	0.377	14.186

Notes. Estimates are derived from from Equation 1, comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects as well as demeaned baseline covariates and their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the outcomes presented in columns (2)-(4). The latter are (2) vaccination rates of children under 36 months, (3) share of visits to the designated health center versus other providers, (4) number of visits to the designated health center. *** p<0.01; ** p<0.05; * p<0.10.

Table H8: Treatment quality index – Subcomponents

	(1) **Treatment quality index**	(2) Used equipment	(3) Waiting time	(4) Examined by trained staff	(5) Privacy during exam	(6) Received test when needed	(7) Diagnosis explained clearly	(8) % staff presence	(9) Facility cleanliness	(10) Drug availability
Full treatment	0.070***	0.056	0.002	0.009	0.073*	0.036	0.058*	0.054	0.118	0.229**
	(0.026)	(0.036)	(0.048)	(0.035)	(0.039)	(0.040)	(0.032)	(0.133)	(0.140)	(0.111)
Constant	0.000	0.031	-0.030	-0.000	-0.000	0.010	0.005	0.010	0.012	0.016
	(0.020)	(0.029)	(0.033)	(0.026)	(0.035)	(0.032)	(0.025)	(0.107)	(0.112)	(0.092)
N	7,288	7,288	7,288	7,288	7,288	7,288	7,288	187	187	187
R^2	0.102	0.023	0.084	0.026	0.039	0.030	0.019	0.299	0.166	0.421
Adjusted p-value (FT)	0.028	0.269	0.960	0.901	0.224	0.604	0.224	0.878	0.604	0.224

Notes. Estimates are derived from from Equation 1, comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects as well as demeaned baseline covariates and their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2)-(10). The latter are z-scores of (2) whether household members reported that, during their most recent visit to the health center, equipment was used during examination, (3) waiting time consisting of the total amount of time spent by the household members waiting for the initial consultation and the examination; whether household members declared that, during their most recent visit to the health center, (4) they were examined by trained health care staff, (5) they had privacy during their examination, (6) lab tests were administered, (7) their diagnosis was clearly explained to them; (8) percent of staff in attendance during an unannounced visit to the health center, measured at the health center level, (9) condition of the clinic (cleanliness of floors and walls, whether the clinic smelled as observed during unannounced visit to health center), measured at the health center level, (10) share of months in which stock cards indicated availability of six key tracer drugs in the past three months, measured at the health center level. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10.

Table H9: Treatment quality index – Non-standardized subcomponents

	(1) **Treatment quality index**	(2) Used equipment	(3) Waiting time	(4) Examined by trained staff	(5) Privacy during exam	(6) Received test when needed	(7) Diagnosis explained clearly	(8) % staff presence	(9) Facility cleanliness	(10) Drug availability
Full treatment	0.070***	0.021	0.142	0.000	0.015*	0.013	0.023*	0.017	0.016	0.059**
	(0.026)	(0.013)	(2.829)	(0.002)	(0.008)	(0.014)	(0.013)	(0.041)	(0.019)	(0.029)
Constant	0.000	0.759***	69.736***	0.831***	0.910***	0.742***	0.697***	0.379***	0.691***	0.731***
	(0.020)	(0.014)	(2.228)	(0.152)	(0.012)	(0.015)	(0.013)	(0.043)	(0.056)	(0.099)
N	7,288	7,288	7,288	7,288	7,288	7,288	7,288	187	187	187
R^2	0.102	0.023	0.084	0.026	0.039	0.030	0.019	0.299	0.166	0.421

Notes. Estimates are derived from from Equation 1, comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects as well as demeaned baseline covariates and their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2)-(10). The latter are (2) whether household members reported that, during their most recent visit to the health center, equipment was used during examination, (3) waiting time consisting of the total amount of time spent by the household members waiting for the initial consultation and the examination; whether household members declared that, during their most recent visit to the health center, (4) they were examined by trained health care staff, (5) they had privacy during their examination, (6) lab tests were administered, (7) their diagnosis was clearly explained to them; (8) percent of staff in attendance during an unannounced visit to the health center, measured at the health center level, (9) condition of the clinic (cleanliness of floors and walls, whether the clinic smelled as observed during unannounced visit to health center), measured at the health center level, (10) share of months in which stock cards indicated availability of six key tracer drugs in the past three months, measured at the health center level. *** p<0.01; ** p<0.05; * p<0.10.

Table H10: Patient satisfaction index – Subcomponents

	(1) **Patient satisfaction index**	(2) Satisfied by HC quality	(3) Satisfied with quality of care	(4) Polite staff	(5) Staff interested in health	(6) Staff listening	(7) Free to express clearly	(8) Availability of staff improving
Full treatment	0.077***	0.105***	0.061*	0.074**	0.101***	0.071**	0.040	0.078*
	(0.024)	(0.035)	(0.032)	(0.035)	(0.030)	(0.031)	(0.039)	(0.040)
Constant	-0.002	-0.007	0.017	0.016	0.020	0.013	0.020	-0.024
	(0.018)	(0.023)	(0.023)	(0.027)	(0.024)	(0.025)	(0.029)	(0.032)
N	7,288	7,288	7,288	7,288	7,288	7,288	7,288	7,288
R^2	0.043	0.066	0.044	0.024	0.018	0.023	0.019	0.053
Adjusted p-value (FT)	0.010	0.010	0.065	0.060	0.006	0.055	0.310	0.065

Notes. Estimates are derived from from Equation 1, comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the outcomes presented in columns (2)-(8). The latter are z-scores of indicator variables of whether household members declared that (2) the services currently offered at the health center are of "very high quality" or "somewhat high quality", (3) they were "very satisfied" or "satisfied" with the quality of care received during their most recent visits to the health center, (4) during their most recent visit to the health center, the person conducting the examination behaved politely/showed respect, (5) during their most recent visit to the health center, the person conducting the examination appeared to be interested in their health condition, (6) during their most recent visit to the health center, the person conducting the examination listened to what they had to say, (7) during their most recent visit to the health center, they felt free to express themselves to the person conducting the examination, (8) compared to the year before, the availability of medical staff had improved at the health center. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10.

Table H11: Patient satisfaction index – Non-standardized subcomponents

	(1) **Patient satisfaction index**	(2) Satisfied by HC quality	(3) Satisfied with quality of care	(4) Polite staff	(5) Staff interested in health	(6) Staff listening	(7) Free to express clearly	(8) Availability of staff improving
Full treatment	0.077***	0.052***	0.023*	0.016**	0.023***	0.016**	0.011	0.035*
	(0.024)	(0.017)	(0.012)	(0.007)	(0.007)	(0.007)	(0.011)	(0.018)
Constant	-0.002	0.423***	0.712***	0.878***	0.883***	0.869***	0.840***	0.394***
	(0.018)	(0.011)	(0.012)	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)
N	7,288	7,288	7,288	7,288	7,288	7,288	7,288	7,288
R^2	0.043	0.066	0.044	0.024	0.018	0.023	0.019	0.053

Notes. Estimates are derived from from Equation 1, comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the outcomes presented in columns (2)-(8). The latter are indicator variables of whether household members declared that (2) the services currently offered at the health center are of "very high quality" or "somewhat high quality", (3) they were "very satisfied" or "satisfied" with the quality of care received during their most recent visits to the health center, (4) during their most recent visit to the health center, the person conducting the examination behaved politely/showed respect, (5) during their most recent visit to the health center, the person conducting the examination appeared to be interested in their health condition, (6) during their most recent visit to the health center, the person conducting the examination listened to what they had to say, (7) during their most recent visit to the health center, they felt free to express themselves to the person conducting the examination, (8) compared to the year before, the availability of medical staff had improved at the health center. *** $p<0.01$; ** $p<0.05$; * $p<0.10$.

Table H12: Health outcomes index at the child level – Subcomponents

	(1) **Health outcomes index**	(2) **Weight/Age 0-18 months**	(3) **MUAC 0-18 months**	(4) **Weight/Age 18-36 months**	(5) **MUAC 18-36 months**
Full treatment	-0.003	-0.000	-0.015	0.004	0.017
	(0.027)	(0.048)	(0.047)	(0.031)	(0.028)
Constant	-0.510***	0.006	0.007	-0.463***	-0.640***
	(0.022)	(0.037)	(0.037)	(0.027)	(0.019)
N	4,930	2,140	2,140	2,790	2,790
R^2	0.112	0.018	0.018	0.225	0.346
Adjusted p-value (FT)	0.900	0.993	0.993	0.993	0.993

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2)-(5). The latter are z-scores of (2) the average ratio of weight over number of months for children under 18 months, (3) the average ratio of weight over number of months for children 18-36 months old, (4) the average ratio of upper arm circumference over number of months for children under 18 months, (5) the average ratio of upper arm circumference over number of months for children 18-36 months old. The unit of analysis is the child. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table H13: Health outcomes index at the child level–Non-standardized subcomponents

	(1) **Health outcome overall**	(2) **Weight/Age 0-18 months**	(3) **MUAC 0-18 months**	(4) **Weight/Age 18-36 months**	(5) **MUAC 18-36 months**
Full treatment	-0.003	-0.000	-0.041	0.000	0.002
	(0.027)	(0.048)	(0.132)	(0.002)	(0.003)
Constant	-0.510***	1.351***	2.874***	0.363***	0.452***
	(0.022)	(0.139)	(0.327)	(0.003)	(0.003)
N	4,930	2,140	2,140	2,790	2,790
R^2	0.112	0.018	0.018	0.225	0.346

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2)-(5). The latter are (2) the average ratio of weight over number of months for children under 18 months, (3) the average ratio of weight over number of months for children 18-36 months old, (4) the average ratio of upper arm circumference over number of months for children under 18 months, (5) the average ratio of upper arm circumference over number of months for children 18-36 months old. The unit of analysis is the child. *** p<0.01; ** p<0.05; * p<0.10

Table H14: Child mortality at the HC level

	(1) Child mortality 0-5 years old	(2) Child mortality 0-1 year old	(3) Child mortality 1-5 years old
Full treatment	-0.011	-0.006	-0.005
	(0.008)	(0.007)	(0.004)
Constant	0.061***	0.041***	0.020***
	(0.006)	(0.005)	(0.003)
N	187	187	187
R^2	0.197	0.211	0.184
Adjusted p-value (FT)	0.298	0.383	0.383

Notes. Estimates from equation 1 comparing the full treatment to the control group. The unit of observation is health center catchment area. The dependent variable is the mortality rate calculated using a synthetic cohort approach for the age brackets 0-5 years (1), 0-12 months (2), and 1-5 years (3), respectively. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

H.2 Intermediate Outcomes

The following tables show regression results for the seven intermediate outcome indices (as summarized in Figure 5) and their components.

Table H15: Intermediate outcomes – Averaged z-score indices

	(1) **Citizen knowledge**	(2) **HC staff knowledge**	(3) **Efficacy**	(4) **Community responsibility**	(5) **Relationship**	(6) **HC transparency**
Full treatment	-0.056**	0.171	-0.022	-0.012	0.040	0.007
	(0.023)	(0.121)	(0.023)	(0.020)	(0.039)	(0.076)
Constant	-0.009	-0.016	-0.002	-0.002	0.004	-0.006
	(0.016)	(0.080)	(0.015)	(0.014)	(0.028)	(0.053)
N	7,288	187	7,288	7,288	7,288	187
R^2	0.205	0.276	0.045	0.054	0.047	0.481
P-value (Full treatment = 0)	0.017	0.158	0.343	0.538	0.307	0.930
Adjusted p-value (FT)	0.104	0.473	0.515	0.646	0.515	0.930

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control for intermediate outcome indices. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The row *P-value (Full treatment = 0)* shows p-values for a Wald test of the null hypothesis that the coefficient on *full treatment* is equal to zero. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table H16: Citizen knowledge index – Subcomponents

	(1) **Citizen knowledge index**	(2) # of patients rights correctly named	(3) # of patients resp. correctly named	(4) # of HC services correctly named
Full treatment	-0.056**	-0.094***	-0.118***	0.042
	(0.023)	(0.028)	(0.032)	(0.037)
Constant	-0.009	-0.009	-0.015	-0.003
	(0.016)	(0.019)	(0.024)	(0.023)
N	7,288	7,288	7,288	7,288
R^2	0.205	0.091	0.166	0.286
Adjusted p-value (FT)	0.104	0.001	0.001	0.246

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the dependent variables in columns (2)-(4). The dependent variable in column (2) is the z-score of the number of patient *rights*, listed in the patient's charter of the Ministry of Health, correctly named by community members, in column (3) is is the z-score of the number of patient *responsibilities*, listed in the patient's charter of the Ministry of Health, correctly named by households members, in column (4) it is the z-score of the number of health center services correctly named by community members. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table H17: HC staff knowledge index – Subcomponents

	(1) **HC staff knowledge index**	(2) # of patients rights correctly named	(3) # of patients resp. correctly named
Full treatment	0.171	0.222	0.128
	(0.121)	(0.139)	(0.151)
Constant	-0.016	-0.019	-0.013
	(0.080)	(0.095)	(0.093)
N	187	187	187
R^2	0.276	0.290	0.211
Adjusted p-value (FT)	0.473	0.223	0.397

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2) and (3). The dependent variable in column (2) is the z-score of the number of patient *rights*, listed in the patient's charter of the Ministry of Health, correctly named by the health center in-charge, in column (3) it is the z-score of the number of patient *responsibilities*, listed in the patient's charter of the Ministry of Health, correctly named by the health center in-charge. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table H18: Efficacy index – Subcomponents

	(1) **Efficacy index**	(2) Have power to improve HC services	(3) Can pressure health worker (effort)	(4) Can pressure health worker (timely)	(5) Can make village a better place to live	(6) Influence over gov. about health services	(7) Influence over HC about services provided
Full treatment	-0.022	-0.034	-0.014	-0.032	-0.040	-0.025	0.008
	(0.023)	(0.028)	(0.033)	(0.038)	(0.032)	(0.031)	(0.028)
Constant	-0.002	0.001	-0.009	-0.008	0.007	-0.001	0.001
	(0.015)	(0.021)	(0.023)	(0.026)	(0.023)	(0.020)	(0.018)
N	7,288	7,288	7,288	7,288	7,288	7,288	7,288
R^2	0.045	0.060	0.025	0.018	0.040	0.031	0.037
Adjusted p-value (FT)	0.515	0.632	0.787	0.632	0.632	0.632	0.787

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2)-(7). The remaining dependent variables are z-scores of the following indicator variables for whether community members think they: (2) have power to improve the quality of health care at the designated health facility, (3) they can pressure health worker to exert better effort in caring for patients by reporting them, (4) they can pressure health worker to work on time by reporting them, (5) they have influence in making the designated village a better place to live, (6) they have a say about how authorities provide health care to their community, (7) they have a say about how health facilities provide health care to their community. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table H19: Community responsibility index – Subcomponents

	(1) **Community responsibility index**	(2) Community resp. for monitoring HC	(3) Community members also responsible
Full treatment	-0.012	0.014	-0.037
	(0.020)	(0.027)	(0.026)
Constant	-0.002	0.001	-0.006
	(0.014)	(0.019)	(0.018)
N	7,288	7,288	7,288
R^2	0.054	0.059	0.039
Adjusted p-value (FT)	0.646	0.614	0.303

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2) and (3). The dependent variable in column (2) is the z-score of a dummy variable indicating whether respondents think that they themselves are responsible for making sure that health workers come to work and provide high-quality health services, in column (3) it is the z-score of a dummy variable indicating whether respondents think *community members* are responsible for making sure that health workers come to work and provide high-quality health services. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

*XXX FIX

Table H20: Relationship index – Subcomponents

	(1) **Relationship index**	(2) Community satisfied with relations	(3) Trust health workers	(4) HC staff satisfied with relations	(5) Health workers will listen to complaints
Full treatment	0.040	0.060*	0.076**	-0.042	0.001
	(0.039)	(0.031)	(0.035)	(0.146)	(0.030)
Constant	0.004	-0.000	0.006	-0.002	0.000
	(0.028)	(0.021)	(0.024)	(0.102)	(0.024)
N	7,288	7,288	7,288	187	7,288
R^2	0.047	0.043	0.046	0.167	0.009
Adjusted p-value (FT)	0.515	0.112	0.112	0.969	0.969

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2) and (5). FIX The latter are indicators variables of whether households think that (2) they are responsible for making sure that health workers come to work and provide high quality health services and (3) that health center staff would listen to their complaints and would not refuse to see them or behave hostilely. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table H21: HC transparency index – Subcomponents

	(1) **HC transparency index**	(2) Poster with opening times	(3) Duty roster displayed	(4) Suggestion box	(5) Info on services provided	(6) Info on patient rights
Full treatment	0.007	0.126	-0.142	0.010	-0.093	0.125
	(0.076)	(0.139)	(0.145)	(0.124)	(0.142)	(0.121)
Constant	-0.006	-0.018	-0.014	0.053	0.002	0.081
	(0.053)	(0.093)	(0.114)	(0.099)	(0.101)	(0.091)
N	187	187	187	187	187	187
R^2	0.481	0.328	0.210	0.398	0.276	0.462
Adjusted p-value (FT)	0.930	0.610	0.610	0.937	0.643	0.610

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the components presented in columns (2)-(6). The remaining dependent variables are z-scores of the following indicator variables whether the designated health facility has: (2) a poster with opening times (3) a duty roster table displayed; (4) a suggestion box; (5) a list of services provided displayed; (6) patient's rights displayed. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

H.3 Treatment Effects on Subgroups of Health Centers

The following tables provide further details on the subgroup results discussed in Section 6.2. To test for subgroup treatment effects, we undertake a number of tests for treatment effects on the five main outcome indices in particular subsets of our sample. We estimate the standard equation:

$$Y_{ij} = \beta_0 + \beta_1 T_{ij}^k + \beta_2 T_{ij}^k * Sub_{ij} + \beta_3 Sub_{ij} + \beta_4 Y_{ij}^0 + \beta_5 X_{ij} + \beta_6 X_{ij} * T_{ij}^k + \phi_d + u_{ij} \qquad (2)$$

where Sub_{ij} is an indicator variable of the subgroup for which we are testing for treatment effects, which for this purpose is not included in the vector of covariates X_{ij}.[32]

We conduct analyses on 10 pre-specified subgroups.[33] Based on our theoretical priors, we divide them into two sets of subgroups in the tables below based on baseline levels of development and proxies of community engagement.

In the tables below, each set of subgroup effects is thus derived from a separate regression, estimated using Equation 2. The tables display the coefficient on Treatment for the base subgroup, β_1 (for example, HC3), the coefficient on the interaction between Treatment and an indicator variable describing the subgroup of interest, β_2, which is indicating the marginal increase in the treatment effect in the health centers/catchment areas in this subgroup (for example, Treat * HC2), and the linear combination of the two coefficients, $\beta_1 + \beta_2$ (for example, Treat + Treat * HC2). Further, we show the difference between subgroups in the control group, β_3. The corresponding standard errors are shown in parentheses. Thus, β_1 indicates the average treatment effect for the base subgroup, β_2 indicates the difference between subgroups, and $\beta_1 + \beta_2$ indicates the average treatment effect for the subgroup of interest.

[32] For specifications looking at subgroup effects by health center level we exclude the three health center level covariates from the vector since they have limited variation, leading to concerns about multicollinearity if they are not excluded.

[33] In deviation from our pre-analysis plan we added an eleventh subgroup. While we had pre-specified replicating our analyses in the subsample of health centers within one standard deviation of the child mortality level in (Björkman and Svensson, 2009); for greater generalizability, we are instead assessing heterogeneous treatment effects by different cutoffs of child mortality, as also shown in Table H25.

Table H22: Subgroup effects on main outcomes – Proxies of low levels of development

	(1) Citizen monitoring	(2) Perceived citizen pressure	(3) Utilization	(4) Treatment quality	(5) Patient satisfaction	(6) Health outcomes	(7) Child mortality
Treat	0.009	-0.321*	0.026	0.038	0.100***	0.024	-0.013
	(0.044)	(0.189)	(0.036)	(0.041)	(0.035)	(0.045)	(0.013)
Treat*HC2	-0.006	0.325	0.003	0.054	-0.039	-0.047	0.005
	(0.059)	(0.243)	(0.046)	(0.055)	(0.049)	(0.059)	(0.017)
HC2	0.081**	-0.321*	-0.005	-0.086**	-0.041	0.042	0.003
	(0.039)	(0.168)	(0.034)	(0.041)	(0.037)	(0.043)	(0.013)
Treat+Treat*HC2	0.003	0.004	0.029	0.093***	0.061*	-0.024	-0.008
	(0.038)	(0.135)	(0.028)	(0.035)	(0.033)	(0.036)	(0.011)
Treat	-0.004	-0.116	0.035	0.050*	0.051*	0.049	-0.018
	(0.032)	(0.135)	(0.029)	(0.028)	(0.028)	(0.038)	(0.012)
Treat*Low treatment quality	0.017	-0.045	-0.017	0.039	0.048	-0.103*	0.018
	(0.037)	(0.258)	(0.034)	(0.026)	(0.038)	(0.055)	(0.017)
Low treatment quality	-0.046*	0.045	-0.006	-0.008	-0.052*	0.070*	-0.007
	(0.024)	(0.144)	(0.023)	(0.020)	(0.029)	(0.040)	(0.013)
Treat+Treat*Low treatment quality	0.013	-0.161	0.018	0.090***	0.100***	-0.054	-0.000
	(0.035)	(0.198)	(0.026)	(0.030)	(0.032)	(0.038)	(0.012)
Treat	-0.028	0.392*	-0.034	0.053	0.128*	0.101	-0.007
	(0.074)	(0.202)	(0.059)	(0.067)	(0.074)	(0.075)	(0.021)
Treat*High U5MR (median)	0.040	-0.612**	0.073	0.023	-0.059	-0.127	-0.007
	(0.081)	(0.246)	(0.064)	(0.073)	(0.078)	(0.080)	(0.023)
High U5MR (median)	-0.017	0.247	-0.009	0.119*	0.069	-0.026	-0.038**
	(0.079)	(0.230)	(0.051)	(0.064)	(0.078)	(0.081)	(0.018)
Treat+Treat*High U5MR (median)	0.013	-0.219*	0.039*	0.076***	0.069***	-0.026	-0.014
	(0.031)	(0.121)	(0.023)	(0.028)	(0.026)	(0.029)	(0.009)
Treat	-0.044	-0.066	0.022	0.096***	0.063*	0.003	-0.006
	(0.034)	(0.149)	(0.028)	(0.032)	(0.033)	(0.035)	(0.011)
Treat*High U5MR (w/i 1 SD of P2P)	0.140**	-0.183	0.018	-0.076	0.036	-0.016	-0.014
	(0.063)	(0.244)	(0.046)	(0.054)	(0.053)	(0.058)	(0.016)
High U5MR (w/i 1 SD of P2P)	-0.075*	0.174	-0.017	0.046	-0.021	-0.021	-0.013
	(0.044)	(0.155)	(0.035)	(0.042)	(0.041)	(0.044)	(0.015)
Treat+Treat*High U5MR (w/i 1 SD of P2P)	0.096*	-0.249	0.039	0.021	0.099**	-0.014	-0.020*
	(0.052)	(0.172)	(0.036)	(0.044)	(0.040)	(0.045)	(0.012)
Treat	0.019	-0.056	0.011	0.044	0.109***	0.008	-0.015
	(0.041)	(0.167)	(0.029)	(0.037)	(0.036)	(0.037)	(0.012)
Treat*Few alternative options	-0.020	-0.147	0.048	0.052	-0.069	-0.044	0.012
	(0.062)	(0.233)	(0.045)	(0.053)	(0.051)	(0.058)	(0.017)
Few alternative options	-0.016	0.055	-0.090***	-0.027	0.039	0.094**	-0.015
	(0.041)	(0.150)	(0.032)	(0.042)	(0.038)	(0.043)	(0.012)
Treat+Treat*Few alternative options	-0.000	-0.204	0.059*	0.096**	0.040	-0.035	-0.003
	(0.044)	(0.149)	(0.033)	(0.038)	(0.035)	(0.042)	(0.012)
Treat	0.020	-0.322	0.109**	-0.030	0.066	-0.009	0.001
	(0.054)	(0.297)	(0.042)	(0.053)	(0.051)	(0.065)	(0.018)
Treat*Rural	-0.008	0.272	-0.099**	0.138**	0.007	0.002	-0.013
	(0.067)	(0.333)	(0.050)	(0.061)	(0.059)	(0.075)	(0.021)
Rural	0.089**	-0.060	0.149***	-0.039	-0.055	-0.053	0.023
	(0.040)	(0.248)	(0.036)	(0.051)	(0.047)	(0.055)	(0.017)
Treat+Treat*Rural	0.012	-0.050	0.010	0.108***	0.073**	-0.008	-0.012
	(0.035)	(0.115)	(0.025)	(0.030)	(0.029)	(0.031)	(0.009)
Treat	0.043	0.047	0.048	0.024	0.112***	-0.043	-0.015
	(0.041)	(0.149)	(0.034)	(0.042)	(0.038)	(0.091)	(0.013)
Treat*No health NGO present	-0.068	-0.315	-0.035	0.073	-0.064	0.035	0.007
	(0.062)	(0.231)	(0.049)	(0.057)	(0.053)	(0.093)	(0.017)
No health NGO present	-0.006	0.158	0.024	-0.083*	0.033	-0.074	-0.014
	(0.043)	(0.143)	(0.034)	(0.045)	(0.041)	(0.066)	(0.013)
Treat+Treat*No health NGO present	-0.026	-0.268	0.013	0.098***	0.048	-0.007	-0.008
	(0.042)	(0.162)	(0.032)	(0.036)	(0.035)	(0.043)	(0.010)

Notes. This table shows estimated average treatment effects for subgroups of health centers. Each set of subgroup effects is derived from a separate regression, estimated using Equation 2. For continuous variables, *High* indicates that a health center/catchment area's value for the given variable is at or above the median; *Low* indicates that it is below the median. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

Table H23: Subgroup effects on main outcomes – Proxies of community engagement

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**	(7) **Child mortality**
Treat	0.003	-0.188	0.017	0.068*	0.052	0.021	-0.012
	(0.042)	(0.157)	(0.033)	(0.038)	(0.036)	(0.038)	(0.009)
Treat*High embeddedness	0.004	0.121	0.022	0.004	0.047	-0.050	0.004
	(0.057)	(0.234)	(0.046)	(0.053)	(0.052)	(0.055)	(0.016)
High embeddedness	-0.072*	0.112	-0.004	0.041	-0.015	0.040	-0.000
	(0.038)	(0.151)	(0.031)	(0.042)	(0.037)	(0.041)	(0.012)
Treat+Treat*High embeddedness	0.008	-0.066	0.039	0.072*	0.100***	-0.029	-0.008
	(0.037)	(0.158)	(0.030)	(0.037)	(0.036)	(0.040)	(0.013)
Treat	0.037	-0.192	0.044	0.093**	0.103**	0.041	-0.007
	(0.049)	(0.149)	(0.033)	(0.043)	(0.041)	(0.045)	(0.012)
Treat*High collective action potential	-0.051	0.104	-0.030	-0.041	-0.048	-0.075	-0.005
	(0.061)	(0.201)	(0.044)	(0.056)	(0.050)	(0.061)	(0.017)
High collective action potential	0.067*	-0.186	0.080**	0.035	0.113***	0.079*	0.005
	(0.040)	(0.136)	(0.032)	(0.043)	(0.037)	(0.044)	(0.013)
Treat+Treat*High collective action potential	-0.014	-0.087	0.015	0.052	0.055*	-0.034	-0.012
	(0.035)	(0.142)	(0.029)	(0.034)	(0.029)	(0.037)	(0.011)
Treat	0.024	-0.155	0.041	0.115***	0.069**	0.002	-0.015
	(0.037)	(0.142)	(0.031)	(0.030)	(0.032)	(0.036)	(0.011)
Treat*High community monitoring	-0.030	0.029	-0.026	-0.083**	0.012	-0.010	0.009
	(0.044)	(0.222)	(0.037)	(0.034)	(0.037)	(0.053)	(0.018)
High community monitoring	0.116***	-0.011	0.065**	0.053**	0.072**	-0.007	-0.008
	(0.044)	(0.156)	(0.027)	(0.024)	(0.029)	(0.041)	(0.013)
Treat+Treat*High community monitoring	-0.006	-0.126	0.015	0.032	0.082***	-0.008	-0.006
	(0.034)	(0.158)	(0.025)	(0.032)	(0.027)	(0.039)	(0.012)
Treat	0.007	-0.127	0.003	0.076***	0.062**	0.045	-0.025**
	(0.032)	(0.127)	(0.028)	(0.028)	(0.028)	(0.038)	(0.011)
Treat*High efficacy	0.000	-0.011	0.053*	-0.013	0.036	-0.108*	0.028*
	(0.028)	(0.209)	(0.032)	(0.020)	(0.030)	(0.057)	(0.017)
High efficacy	0.043**	-0.148	-0.002	0.012	0.037*	0.047	-0.013
	(0.020)	(0.142)	(0.023)	(0.015)	(0.022)	(0.042)	(0.013)
Treat+Treat*High efficacy	0.008	-0.138	0.056**	0.063**	0.098***	-0.062	0.004
	(0.029)	(0.164)	(0.025)	(0.028)	(0.028)	(0.040)	(0.012)
Treat	-0.002	0.098	-0.004	0.060*	0.038	-0.008	-0.006
	(0.039)	(0.129)	(0.030)	(0.034)	(0.032)	(0.042)	(0.013)
Treat*HC near	0.021	-0.478**	0.067	0.017	0.083	0.010	-0.009
	(0.059)	(0.200)	(0.044)	(0.054)	(0.052)	(0.059)	(0.017)
HC near	-0.035	0.138	-0.036	0.023	-0.069*	0.033	-0.006
	(0.040)	(0.132)	(0.031)	(0.041)	(0.036)	(0.042)	(0.011)
Treat+Treat*HC near	0.019	-0.380**	0.063*	0.077*	0.120***	0.003	-0.015
	(0.044)	(0.161)	(0.032)	(0.042)	(0.039)	(0.038)	(0.010)

Notes. This table shows estimated average treatment effects for subgroups of health centers. Each set of subgroup effects is derived from a separate regression, estimated using Equation 2. For continuous variables, *High* indicates that a health center/catchment area's value for the given variable is at or above the median; *Low* indicates that it is below the median. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

H.4 Treatment Effects by Baseline Level of Development

This section presents and summarizes subgroup analyses investigating whether the intervention had stronger treatment effects in health centers with lower baseline levels of development. Table H24 synthesizes the results from Table H22 with regard to treatment quality. Table H25 uses the same specification as described in Section H.3 to investigate heterogeneity of treatment effects by baseline level of mortality. Since this heterogeneous analysis was not pre-specified, we show results for different cutoffs.

Table H24: Estimated treatment effects on treatment quality, by baseline level of development in the health center catchment area

Less developed		More developed		
Subgroup	ATE	Subgroup	ATE	Difference significant
HC2	0.093***	HC3	0.038	no
Low BL TQ	0.090***	High BL TQ	0.050*	no
Few alternative options	0.096**	Many alternative options	0.044	no
HC far	0.060*	HC near	0.077*	no
Rural	0.108***	Urban	-0.03	yes
No health NGO present	0.098***	Health NGO present	0.024	no
High BL U5MR	0.076***	Low BL U5MR	0.053	no

Notes. The table shows estimated average treatment effects on the treatment quality index by subgroup; summarizing results presented in greater detail in column (2) of Table H22.

Table H25: Subgroup effects on child mortality by baseline mortality rate

	(1)	(2)	(3)	(4)
Percentile used to define subgroups:	**58th**	**66th**	**75th**	**90th**
Treat	-0.006	-0.005	-0.008	-0.012
	(0.011)	(0.011)	(0.010)	(0.009)
Treat * U5MR above xxth percentile	-0.014	-0.017	-0.009	0.017
	(0.016)	(0.016)	(0.017)	(0.023)
U5MR above xxth percentile	-0.028*	-0.018	-0.007	-0.006
	(0.017)	(0.017)	(0.018)	(0.018)
Treat + Treat * U5MR above xxth percentile	-0.020*	-0.022*	-0.017	0.005
	(0.011)	(0.012)	(0.013)	(0.021)

H.5 Treatment Effects by Strength of Ruling Party

Table H26: By Presidential vote share above or below the national median

	(1) Citizen monitoring	(2) Citizen pressure	(3) Utilization	(4) Treatment quality	(5) Patient satisfaction	(6) Health outcomes	(7) Ch mort
Treat	-0.014	-0.290*	0.026	0.084**	0.119***	-0.011	-0.0
	(0.039)	(0.168)	(0.028)	(0.037)	(0.033)	(0.040)	(0.0
Treat * Museveni vote share above national median (2011)	0.049	0.371	0.005	-0.034	-0.102*	0.004	0.0
	(0.057)	(0.272)	(0.047)	(0.051)	(0.052)	(0.059)	(0.0
Constant	0.002	0.008	-0.020	0.001	0.000	-0.003	0.06(
	(0.019)	(0.070)	(0.016)	(0.020)	(0.018)	(0.022)	(0.0
N	7,288	187	7,288	7,288	7,288	4,930	18
R^2	0.091	0.096	0.230	0.092	0.039	0.018	0.1
Coeff. Treat + Treat * Museveni vote share above national median (2011)	0.035	0.081	0.031	0.049	0.017	-0.007	-0.0
SE. Treat + Treat * Museveni vote share above national median (2011)	0.041	0.174	0.036	0.036	0.039	0.041	0.0
P-value Treat + Treat * Museveni vote share above national median (2011)	0.400	0.642	0.400	0.174	0.663	0.863	0.8

Notes. Estimates from Equation 2. *Museveni vote share above national median (2011)* is defined at the district level. All models include district fixed effects and demeaned b covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

Table H27: By Presidential vote share above or below the sample median

	(1) Citizen monitoring	(2) Citizen pressure	(3) Utilization	(4) Treatment quality	(5) Patient satisfaction	(6) Health outcomes	(7) Chil morta
Treat	-0.011	-0.304*	0.033	0.081**	0.124***	-0.019	-0.0
	(0.040)	(0.172)	(0.028)	(0.038)	(0.033)	(0.041)	(0.01
Treat * Museveni vote share above sample median (2011)	0.040	0.390	-0.013	-0.028	-0.109**	0.022	0.01
	(0.057)	(0.271)	(0.047)	(0.051)	(0.052)	(0.059)	(0.02
Constant	0.002	0.008	-0.020	0.001	0.000	-0.003	0.060
	(0.019)	(0.070)	(0.016)	(0.020)	(0.018)	(0.022)	(0.00
N	7,288	187	7,288	7,288	7,288	4,930	187
R^2	0.091	0.098	0.230	0.092	0.039	0.018	0.18
Coeff. Treat + Treat * Museveni vote share above sample median (2011)	0.029	0.086	0.021	0.053	0.015	0.003	-0.00
SE. Treat + Treat * Museveni vote share above sample median (2011)	0.040	0.169	0.036	0.035	0.038	0.040	0.01
P-value Treat + Treat * Museveni vote share above sample median (2011)	0.463	0.611	0.563	0.132	0.694	0.932	0.89

Notes. Estimates from Equation 2. *Museveni vote share above national median (2011)* is defined at the district level. All models include district fixed effects and demeaned ba covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

H.6 Robustness Checks for Endline Results

Table H28: Robustness check main outcomes

	(1) Citizen monitoring	(2) Perceived citizen pressure	(3) Utilization	(4) Treatment quality	(5) Patient satisfaction	(6) Health outcomes	(7) Child mortality
A: Without control variables							
Treatment	0.006	-0.131	0.034	0.070***	0.075***	-0.014	-0.010
	(0.028)	(0.105)	(0.023)	(0.027)	(0.025)	(0.028)	(0.008)
Constant	0.003	0.011	-0.022	0.001	-0.001	-0.000	0.061
	(0.019)	(0.070)	(0.017)	(0.021)	(0.018)	(0.022)	0.006
B: Without district fixed effects							
Treatment	0.010	-0.120	0.013	0.064**	0.075***	-0.014	-0.009
	(0.036)	(0.100)	(0.028)	(0.029)	(0.025)	(0.029)	(0.008)
Constant	0.001	0.004	-0.012	0.003	-0.001	0.000	0.063
	(0.026)	(0.062)	(0.020)	(0.023)	(0.018)	(0.023)	0.006
C: With outcome measures aggregated at HC level							
Treatment	-0.001	-0.140	0.035*	0.062**	0.080***	-0.010	-0.011
	(0.029)	(0.102)	(0.021)	(0.027)	(0.025)	(0.029)	(0.008)
Constant	0.004	0.015	-0.014	0.002	-0.001	-0.002	0.061
	(0.020)	(0.064)	(0.014)	(0.021)	(0.019)	(0.022)	(0.006)
D: Difference between post and pre-treatment values							
Treatment	0.013	-0.113	0.004	0.074***	0.067***	-0.023	-0.008
	(0.032)	(0.102)	(0.025)	(0.028)	(0.025)	(0.034)	(0.008)
Constant	0.029	0.042	-0.023*	0.020	-0.011	0.028	0.045
	(0.021)	(0.055)	(0.012)	(0.016)	(0.015)	(0.040)	(0.005)
*E: Without control variable * treatment interaction*							
Program impact	0.006	-0.134	0.027	0.070***	0.077***	-0.009	-0.010
	(0.028)	(0.103)	(0.022)	(0.026)	(0.024)	(0.028)	(0.008)
Constant	0.004	0.013	-0.020	0.001	-0.002	-0.003	0.061
	(0.019)	(0.067)	(0.016)	(0.020)	(0.018)	(0.022)	(0.006)
Observations (A, B & E)	7,288	187	7,288	7,288	7,288	4,930	187
Observations (C)	187	187	187	187	187	187	187
Observations (D)	14,576	374	14,576	14,576	14,576	9,860	374

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control, with the following variations: Panel A shows results without covariates, panel B without district fixed effects, and panel C aggregates outcome measures and covariates at the health center level (the unit of randomization). Panel D shows results from a difference in difference estimation. *** p<0.01; ** p<0.05; * p<0.10

Table H29: Robustness checks – Intermediate outcomes

	(1) **Citizen knowledge**	(2) **HC staff knowledge**	(3) **Efficacy**	(4) **Community responsibility**	(5) **Relationship**	(6) **HC transparency**
A: Without control variables						
Treatment	-0.063*	0.156	-0.024	-0.012	0.042	-0.028
	(0.033)	(0.125)	(0.024)	(0.021)	(0.039)	(0.089)
Constant	-0.005	-0.012	-0.000	-0.002	0.002	0.004
	(0.025)	(0.084)	(0.016)	(0.014)	(0.028)	(0.059)
B: Without district fixed effects						
Treatment	-0.065	0.155	-0.024	-0.018	0.044	0.005
	(0.040)	(0.123)	(0.027)	(0.028)	(0.042)	(0.079)
Constant	-0.005	-0.009	-0.000	0.001	0.002	-0.007
	(0.028)	(0.081)	(0.017)	(0.021)	(0.030)	(0.057)
C: Without ouctome measures aggregated at HC level						
Treatment	-0.058**	0.171	-0.023	-0.013	0.029	0.007
	(0.024)	(0.121)	(0.025)	(0.022)	(0.042)	(0.076)
Constant	-0.008	-0.016	-0.001	-0.001	0.002	-0.006
	(0.017)	(0.080)	(0.016)	(0.015)	(0.030)	(0.053)
D: Difference between post and pre-treatment values						
Treatment	-0.067*	0.188	-0.025	-0.017	0.030	-0.012
	(0.038)	(0.118)	(0.025)	(0.022)	(0.040)	(0.080)
Constant	-0.006	0.046	-0.018	-0.016	-0.042	0.003
	(0.020)	(0.055)	(0.011)	(0.010)	(0.026)	(0.042)
Observations (A & B)	7,288	187	7,288	7,288	7,288	187
Observations (C)	187	187	187	187	187	187
Observations (D)	14,576	374	14,576	14,576	14,576	374

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control, with the following variations: Panel A shows results without covariates, panel B without district fixed effects, and panel C aggregates outcome measures and covariates at the health center level (the unit of randomization). Panel D shows results from a difference in difference estimation. *** p<0.01; ** p<0.05; * p<0.10

We also test the effect on child mortality with a Cox proportional hazards model, leveraging the fact that we have child-month level data on survival over the course of 36 months for over 20,000 children. By estimating the treatment effect on the chance of survival of the individual child, it mimics the data generation process more closely. The Cox model includes the same vector of controls and their interaction with the treatment indicator as Equation 1. Standard errors are clustered by health center catchment area.[34]

Table H30: Child mortality at the child level

	(1) **0-5 years old**	(2) **0-1 year old**	(3) **1-5 years old**
Full treatment	1.059	1.120	0.612
	(0.239)	(0.295)	(0.286)
N	10,118	4,543	8,635
Adjusted p-value (FT)	0.900	0.515	0.450

Notes. Displaying hazard ratios estimated with a Cox proportional hazards model, comparing outcomes between the full treatment and the control group. A hazard ratio below (above) 1 implies that the treatment led to lower (higher) mortality rates. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Standard errors are clustered at the health center level. The dependent variable is the incident of death, observed at the child-month level in the age bracket 0-5 years (1), 0-12 months (2), and 1-5 years (3), respectively. The unit of analysis is the child. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

[34] An important assumption of the Cox model is that the relative effect of a covariate on the hazard function is constant over time (proportional hazard rates). We therefore do not include district fixed effects in our Cox models, since different regions of Uganda experience different seasonal patterns and thus different temporal patterns of child mortality rates. Our results are not affected by the exclusion of district fixed effects.

Table H31: Main outcomes – Principal component indices

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**
Full treatment	-0.026	-0.232	0.029	0.057	0.094***	-0.011
	(0.032)	(0.150)	(0.031)	(0.037)	(0.033)	(0.031)
Constant	0.013	0.102	-0.053**	-0.030	-0.047*	0.011
	(0.022)	(0.099)	(0.022)	(0.029)	(0.025)	(0.024)
N	7,288	187	7,288	7,288	7,288	4,212
R^2	0.042	0.152	0.284	0.040	0.028	0.042

Notes. Main outcome indices constructed using principal component analysis instead of averaged z-scores. Estimates from Equation 1 comparing outcomes measured at endline between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

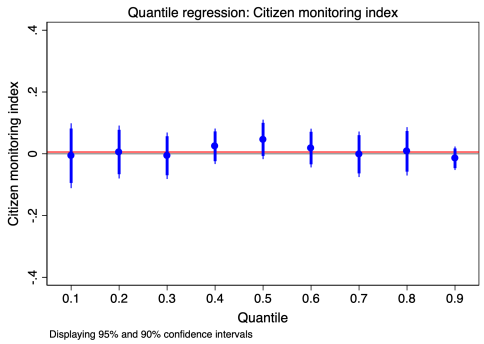
Table H32: Intermediate outcomes – Principal component indices

	(1) **Citizen knowledge**	(2) **HC staff knowledge**	(3) **Efficacy**	(4) **Community responsibility**	(5) **Relationship**	(6) **HC transparency**
Full treatment	-0.106***	0.175	-0.033	-0.017	0.079**	0.018
	(0.032)	(0.140)	(0.036)	(0.027)	(0.035)	(0.115)
Constant	0.052**	-0.087	0.016	0.009	-0.041	0.005
	(0.023)	(0.096)	(0.024)	(0.019)	(0.025)	(0.079)
N	7,288	187	7,288	7,288	7,288	187
R^2	0.132	0.247	0.044	0.053	0.059	0.493

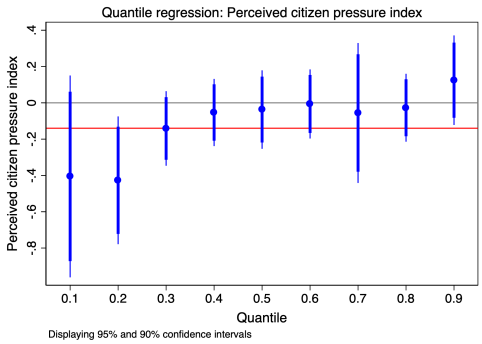
Notes. Intermediate outcome indices constructed using principal component analysis instead of averaged z-scores. Estimates from Equation 1 comparing outcomes measured at endline between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

Figure 7: Quantile regressions of treatment effects

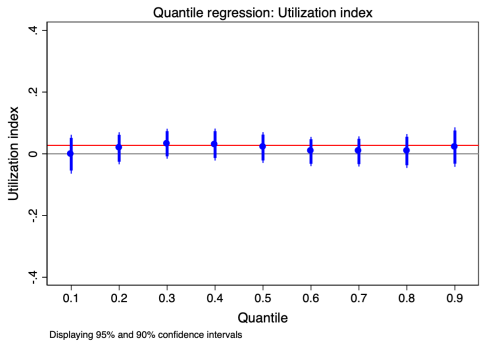
(a) Citizen monitoring



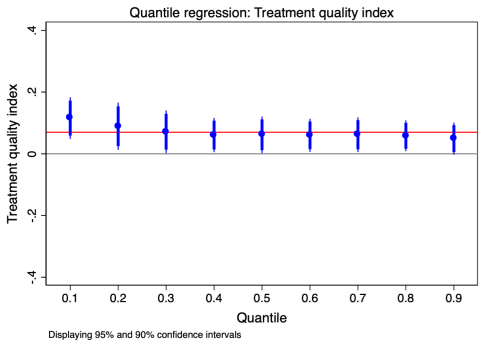
(b) Perceived citizen pressure



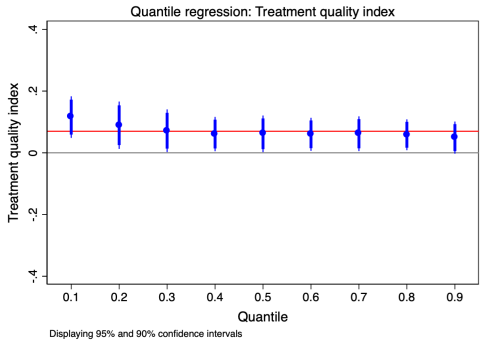
(c) Utilization



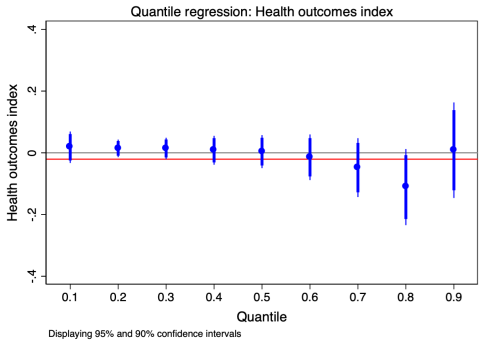
(d) Treatment quality



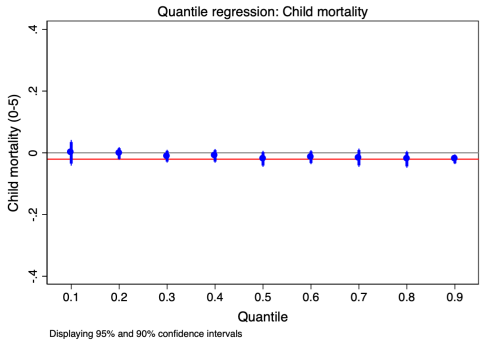
(e) Patient satisfaction



(f) Health outcomes



(g) Child mortality



Additional Robustness Checks on Treatment Quality

Table H33: Robustness check – Excluding subcomponents of treatment quality

	(1) Treatment quality wo/ used equipment	(2) Treatment quality wo/ waiting time	(3) Treatment quality wo/ examined by trained staff	(4) Treatment quality wo/ privacy during exam	(5) Treatment quality wo/ received test when needed	(6) Treatment quality wo/ diagnosis explained clearly	(7) Treatment quality wo/ % staff presence	(8) Treatment quality wo/ facility cleanliness	(9) Treatment quality wo/ drug availability
Full treatment	0.070***	0.021	0.142	0.000	0.015*	0.013	0.023*	0.017	0.016
	(0.026)	(0.013)	(2.829)	(0.002)	(0.008)	(0.014)	(0.013)	(0.041)	(0.019)
Constant	0.000	0.759***	69.736***	0.831***	0.910***	0.742***	0.697***	0.379***	0.691***
	(0.020)	(0.014)	(2.228)	(0.152)	(0.012)	(0.015)	(0.013)	(0.043)	(0.056)
N	7,288	7,288	7,288	7,288	7,288	7,288	7,288	187	187
R^2	0.102	0.023	0.084	0.026	0.039	0.030	0.019	0.299	0.166

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable is the treatment quality index, excluding one subcomponent at a time. *** p<0.01; ** p<0.05; * p<0.10

Table H34: Robustness checks – Main treatment quality outcomes with no HC subcomponents

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**	(7) **Child mortality**
A: Without control variables							
Treatment	0.006	-0.131	0.034	0.038*	0.075***	-0.014	-0.010
	(0.028)	(0.105)	(0.023)	(0.020)	(0.025)	(0.028)	(0.008)
Constant	0.003	0.011	-0.022	-0.002	-0.001	-0.000	0.061***
	(0.019)	(0.070)	(0.017)	(0.015)	(0.018)	(0.022)	0.006
B: Without district fixed effects							
Treatment	0.010	-0.120	0.013	0.036*	0.075***	-0.014	-0.009
	(0.036)	(0.100)	(0.028)	(0.020)	(0.025)	(0.029)	(0.008)
Constant	0.001	0.004	-0.012	-0.002	-0.001	0.000	0.063***
	(0.026)	(0.062)	(0.020)	(0.016)	(0.018)	(0.023)	0.006
C: With outcome measures aggregated at HC level							
Treatment	-0.001	-0.140	0.035*	0.031	0.080***	-0.010	-0.011
	(0.029)	(0.102)	(0.021)	(0.019)	(0.025)	(0.029)	(0.008)
Constant	0.004	0.015	-0.014	-0.002	-0.001	-0.002	0.061***
	(0.020)	(0.064)	(0.014)	(0.015)	(0.019)	(0.022)	(0.006)
D: Difference between post and pre-treatment values							
Treatment	0.013	-0.113	0.004	0.039*	0.067***	-0.023	-0.008
	(0.032)	(0.102)	(0.025)	(0.022)	(0.025)	(0.034)	(0.008)
Constant	0.029	0.042	-0.023*	0.010	-0.011	0.028	0.045***
	(0.021)	(0.055)	(0.012)	(0.012)	(0.015)	(0.040)	(0.005)
Observations (A & B)	7,288	187	7,288	7,288	7,288	4,930	187
Observations (C)	187	187	187	187	187	187	187
Observations (D)	14,576	374	14,576	14,576	14,576	9,860	374

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control, with the following variations: Panel A shows results without covariates, panel B without district fixed effects, and panel C aggregates outcome measures and covariates at the health center level (the unit of randomization). Panel D shows results from a difference in difference estimation. Here, the treatment quality index excludes the components measured at the health center level. *** p<0.01; ** p<0.05; * p<0.10

H.7 Presence of Local Government Officials

Table H35: Main outcomes and presence of an official

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**	(7) **Child mortality**
Full treatment	0.020	-0.148	0.049**	0.054*	0.074***	-0.006	-0.011
	(0.029)	(0.127)	(0.023)	(0.028)	(0.026)	(0.032)	(0.009)
Subcounty official attendance	-0.042	0.024	-0.068*	0.048	0.008	-0.047	0.001
	(0.048)	(0.145)	(0.036)	(0.036)	(0.038)	(0.040)	(0.013)
Constant	0.004	0.015	-0.020	0.000	-0.002	0.004	0.061***
	(0.019)	(0.064)	(0.016)	(0.020)	(0.018)	(0.021)	(0.006)
N	7,288	187	7,288	7,288	7,288	4,212	187
R^2	0.098	0.152	0.231	0.104	0.043	0.026	0.197
Full treatment net effect	-0.023	-0.123	-0.019	0.102***	0.082**	-0.053	-0.010
	(0.046)	(0.112)	(0.034)	(0.035)	(0.036)	(0.036)	(0.012)

Notes. Subcounty official present indicates whether an official from the local government was present at either the community dialogue or the interface meeting. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control. The unit of observation in columns (1) and (3)-(5) is the household, in column (6) it is child, and in columns (1) and (7) the health center catchment area. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at health center level. *** p<0.01; ** p<0.05; * p<0.10

Table H36: Subcomponents of the treatment quality index and presence of an official

	(1) **Treatment quality index**	(2) Used equipment	(3) Waiting time	(4) Examined by trained staff	(5) Privacy during exam	(6) Received test when needed	(7) Diagnosis explained clearly	(8) % st prese
Full treatment	0.054*	0.045	0.005	0.008	0.069	0.023	0.079**	-0.0
	(0.028)	(0.040)	(0.052)	(0.039)	(0.042)	(0.044)	(0.036)	(0.1
Subcounty official present	0.048	0.033	-0.007	0.002	0.013	0.038	-0.066	0.3
	(0.036)	(0.047)	(0.085)	(0.030)	(0.049)	(0.053)	(0.046)	(0.1
Constant	0.000	0.031	-0.030	-0.000	-0.000	0.011	0.005	0.0
	(0.020)	(0.029)	(0.033)	(0.026)	(0.035)	(0.032)	(0.025)	(0.1
N	7,288	7,288	7,288	7,288	7,288	7,288	7,288	18
R^2	0.104	0.023	0.084	0.026	0.039	0.030	0.020	0.3
Full treatment + presence	0.102***	0.079*	-0.003	0.010	0.081	0.061	0.013	0.2
	(0.035)	(0.045)	(0.080)	(0.035)	(0.053)	(0.052)	(0.044)	(0.1

Notes. Subcounty official present indicates whether an official from the local government was present at either the community dialogue or the interfa fixed effects as well as demeaned baseline covariates and their interaction with the treatment indicator. Robust standard errors are clustered at the heal in column (1) is the averaged z-score index of the components presented in columns (2)-(10). The latter are z-scores of (2) whether household me recent visit to the health center, equipment was used during examination, (3) waiting time consisting of the total amount of time spent by the hou consultation and the examination; whether household members declared that, during their most recent visit to the health center, (4) they were examin had privacy during their examination, (6) lab tests were administered, (7) their diagnosis was clearly explained to them; (8) percent of staff in atte the health center, measured at the health center level, (9) condition of the clinic (cleanliness of floors and walls, whether the clinic smelled as obser center), measured at the health center level, (10) share of months in which stock cards indicated availability of six key tracer drugs in the past three level. *** p<0.01; ** p<0.05; * p<0.10.

Table 37: Subcomponents of patient satisfaction and presence of an official

	(1) **Patient satisfaction index**	(2) Satisfied by HC quality	(3) Satisfied with quality of care	(4) Polite staff	(5) Staff interested in health	(6) Free to express clearly	(7) Availability of staff improving
Full treatment	0.074***	0.060*	0.075*	0.111***	0.077**	0.043	0.039
	(0.026)	(0.036)	(0.039)	(0.030)	(0.035)	(0.043)	(0.043)
Subcounty official present	0.008	0.005	-0.002	-0.031	-0.017	-0.008	0.119**
	(0.038)	(0.052)	(0.052)	(0.050)	(0.045)	(0.060)	(0.053)
Control	-0.002	0.017	0.016	0.020	0.013	0.020	-0.024
	(0.018)	(0.023)	(0.027)	(0.024)	(0.025)	(0.029)	(0.032)
N	7,288	7,288	7,288	7,288	7,288	7,288	7,288
R^2	0.043	0.044	0.024	0.018	0.023	0.019	0.055
Full treatment + presence	0.082**	0.065	0.073	0.080	0.060	0.035	0.158***
	(0.036)	(0.048)	(0.049)	(0.051)	(0.042)	(0.058)	(0.054)

Notes. Subcounty official present indicates whether an official from the local government was present at either the community dialogue or the interface meeting. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of the outcomes presented in columns (2)-(8). The latter are z-scores of indicator variables of whether household members declared that (2) the services currently offered at the health center are of "very high quality" or "somewhat high quality", (3) they were "very satisfied" or "satisfied" with the quality of care received during their most recent visits to the health center, (4) during their most recent visit to the health center, the person conducting the examination behaved politely/showed respect, (5) during their most recent visit to the health center, the person conducting the examination appeared to be interested in their health condition, (6) during their most recent visit to the health center, the person conducting the examination listened to what they had to say, (7) during their most recent visit to the health center, they felt free to express themselves to the person conducting the examination, (8) compared to the year before, the availability of medical staff had improved at the health center. *** $p<0.01$; ** $p<0.05$; * $p<0.10$.

H.8 Main Results, by Treatment Arm

H.8.1 Analysis by Treatment Arm

Although our primary interest is in the impact of the full ACT Health intervention, we use a factorial design to better understand the effects of the program's different elements, as described in Section 4. We combine the information and mobilization components into one treatment arm and cross it with the interface treatment, as depicted in Figure 8. We then randomly assign health centers and their catchment areas to one of the four treatment groups, with treatment assignment blocked by district and health center level. Communities and health centers assigned to the information and mobilization treatment received the CRCs and were invited to attend separate meetings—one for health center staff, another for community members—at which the contents of the CRCs were discussed and action plans were developed in light of the information they contained. Communities and health centers assigned to the interface only treatment did not receive the CRCs but were invited to attend a meeting that brought citizens and health center staff together to discuss how to improve health outcomes in the community. Communities and health centers assigned to the full treatment received both of these components: the CRCs and separate community and health center staff meetings plus the meeting that brought the two parties together. The factorial design enables us to assess the effectiveness of the full ACT Health intervention by comparing units in the bottom right cell to the control group and to learn which aspects of the broader intervention are doing the work in generating the effects we find by making comparisons across all four cells.

		Interface meetings are held between community and health facility staff	
		No	**Yes**
Report card info is reported to community and health facility staff and action plans are developed	**No**	CONTROL 95 HCs (54 HC2 ; 41 HC3)	INTERFACE WITHOUT INFORMATION OR MOBILIZATION 97 HCs (60 HC2 ; 37 HC3)
	Yes	INFORMATION AND MOBILIZATION WITHOUT INTERFACE 92 HCs (55 HC2 ; 37 HC3)	FULL ACT HEALTH INTERVENTION 92 HCs (56 HC2 ; 36 HC3)

Figure 8: Factorial design

To test the effect of each treatment arm, we estimate the model:

$$Y_{ij} = \beta_0 + \beta_1 T_{ij}^{IM} + \beta_2 T_{ij}^{IMI} + \beta_3 T_{ij}^{I} + \beta_4 Y_{ij}^{0} + \beta_5 X_{ij} + \beta_6 X_{ij} * T_{ij}^{IM} + \beta_7 X_{ij} * T_{ij}^{IMI} + \beta_8 X_{ij} * T_{ij}^{I} + \phi_d + u_{ij} \quad (3)$$

where T_{ij}^{IM} is a binary variable indicating whether the health center and catchment area j was assigned to receive only the information and mobilization treatment, T_{ij}^{I} indicates whether the unit was assigned to receive only the interface treatment, T_{ij}^{IMI} indicates whether the unit was assigned to receive the full treatment (i.e., information and mobilization plus interface), and all other terms are defined as in Equation 1. This set-up allows us to compare each cell in the factorial design to the control group.[35]

Main Specification

Table 38: Citizen monitoring index – All treatment arms

	(1) **Citizen monitoring index**	(2) Attended LC1 meetings	(3) HC discussed at LC1 meetings	(4) Community would find out: staff late	(5) Community would find out: staff no effort
Full treatment	0.003	0.051	0.029	-0.012	-0.051*
	(0.029)	(0.054)	(0.072)	(0.033)	(0.028)
Information and mobilization only	0.036	-0.010	0.202***	-0.013	-0.032
	(0.030)	(0.056)	(0.073)	(0.036)	(0.032)
Interface only	0.041	0.041	0.185**	-0.023	-0.037
	(0.029)	(0.054)	(0.074)	(0.032)	(0.029)
Constant	0.005	0.042	0.023	0.012	0.010
	(0.020)	(0.043)	(0.050)	(0.024)	(0.019)
N	14,609	14,609	14,609	14,609	14,609
R^2	0.077	0.082	0.090	0.035	0.026
P-value (Information and mobilization = Interface)	0.874	0.300	0.827	0.774	0.880
P-value (Information and mobilization = Full treatment)	0.284	0.210	0.020	0.967	0.558
P-value (Interface = Full treatment)	0.200	0.829	0.039	0.721	0.647
F-test (joint significance of all 3 treatment groups)	1.063	0.712	3.996	0.174	1.242
P-value (joint significance of all 3 treatment groups)	0.365	0.545	0.008	0.914	0.294

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of columns (2)-(5). The remaining dependent variables are z-scores of: (2) a dummy variable whether household members report having attended at least one LC1 meeting during the last 12 months; (3) a dummy variable whether the local health center was discussed at the most recent LC1 meeting; (4) a Likert-scale variable of whether the community would find out if a staff were regularly late or (5) extended no effort. *** p<0.01; ** p<0.05; * p<0.10

[35] We had initially pre-specified the model $Y_{ij} = \beta_0 + \beta_1 T_{ij}^{IM} + \beta_2 T_{ij}^{IM} T_{ij}^{I} + \beta_3 T_{ij}^{I} + \beta_4 Y_{ij}^{0} + \beta_5 X_{ij} + \phi_d + u_{ij}$, which considers the rows and columns in Figure 8 as well as their interaction. We deem the model described in Equation 3 superior since it relies on fewer assumptions, is easier to interpret, and presents our findings in a way that is consistent with the results in the main specification. Results from the pre-specified model are presented further below.

Table 39: Perceived citizen pressure index – All treatment arms

	(1) **Perceived citizen pressure index**	(2) Community would find out: staff no effort	(3) Community would find out: staff absent	(4) Any report of staff wrongdoing in past 12 months
Full treatment	-0.136	-0.102	-0.305*	0.014
	(0.101)	(0.155)	(0.161)	(0.148)
Information and mobilization only	-0.044	0.015	-0.057	-0.069
	(0.092)	(0.146)	(0.156)	(0.144)
Interface only	-0.015	-0.073	0.045	-0.006
	(0.087)	(0.151)	(0.147)	(0.144)
Constant	0.006	0.032	-0.008	-0.005
	(0.063)	(0.104)	(0.103)	(0.103)
N	376	376	376	376
R^2	0.109	0.118	0.117	0.118
P-value (Information and mobilization = Interface)	0.754	0.564	0.507	0.657
P-value (Information and mobilization = Full treatment)	0.385	0.458	0.148	0.568
P-value (Interface = Full treatment)	0.241	0.856	0.033	0.888
F-test (joint significance of all 3 treatment groups)	0.658	0.262	1.726	0.132
P-value (joint significance of all 3 treatment groups)	0.578	0.853	0.161	0.941

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of columns (2)-(4). *** p<0.01; ** p<0.05; * p<0.10

Table 40: Utilization outcomes – All treatment arms

	(1) **Utilization index**	(2) Vaccination rates, children<36 months	(3) % of visits to HC, vs. other providers	(4) Number of visits to HC
Full treatment	0.027	0.048	0.038	-0.002
	(0.022)	(0.036)	(0.033)	(0.027)
Information and mobilization only	0.013	0.026	0.025	-0.003
	(0.022)	(0.034)	(0.031)	(0.029)
Interface only	0.054**	0.049	0.074**	0.038
	(0.022)	(0.034)	(0.033)	(0.027)
Constant	-0.018	-0.008	0.001	-0.016
	(0.015)	(0.025)	(0.022)	(0.020)
N	14,609	8,548	14,609	14,609
R^2	0.221	0.045	0.173	0.273
P-value (Information and mobilization = Interface)	0.066	0.485	0.151	0.141
P-value (Information and mobilization = Full treatment)	0.518	0.520	0.712	0.962
P-value (Interface = Full treatment)	0.234	0.979	0.296	0.121
F-test (joint significance of all 3 treatment groups)	2.243	0.859	1.702	1.133
P-value (joint significance of all 3 treatment groups)	0.083	0.463	0.166	0.335

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of columns (2)-(4). The remaining dependent variables are z-scores of: (2) vaccination rates of children under 36 months; (3) the share of visits to the designated health center versus other providers; (4) the number of visits to the designated health center by all household members. *** p<0.01; ** p<0.05; * p<0.10

Table 41: Treatment quality outcomes – All treatment arms

	(1) **Treatment quality index**	(2) Used equipment	(3) Waiting time	(4) Examined by trained staff	(5) Privacy during exam	(6) Received test when needed	(7) Diagnosis explained clearly	(8) % staff presence	(9) Facility cleanliness	(10) Drug availability
Full treatment	0.071***	0.062*	0.003	0.006	0.075*	0.041	0.063*	0.030	0.107	0.246**
	(0.026)	(0.037)	(0.049)	(0.032)	(0.040)	(0.041)	(0.033)	(0.136)	(0.137)	(0.109)
Information and mobilization only	0.013	-0.035	-0.011	0.004	0.050	-0.007	-0.019	-0.121	-0.074	0.313***
	(0.029)	(0.040)	(0.043)	(0.039)	(0.043)	(0.046)	(0.036)	(0.140)	(0.135)	(0.111)
Interface only	0.022	0.036	-0.035	0.017	0.015	0.036	0.022	-0.138	-0.099	0.343***
	(0.027)	(0.042)	(0.046)	(0.033)	(0.045)	(0.046)	(0.037)	(0.134)	(0.137)	(0.103)
Constant	-0.002	0.021	-0.035	-0.001	0.001	0.012	-0.003	0.013	0.016	0.010
	(0.021)	(0.028)	(0.034)	(0.026)	(0.035)	(0.032)	(0.025)	(0.106)	(0.109)	(0.088)
N	14,609	14,609	14,609	14,609	14,609	14,609	14,609	376	376	376
R^2	0.102	0.030	0.102	0.015	0.031	0.035	0.021	0.299	0.176	0.367
P-value (Information and mobilization = Interface)	0.740	0.098	0.558	0.677	0.382	0.369	0.293	0.892	0.835	0.748
P-value (Information and mobilization = Full treatment)	0.025	0.011	0.751	0.956	0.475	0.265	0.019	0.239	0.132	0.525
P-value (Interface = Full treatment)	0.032	0.521	0.417	0.616	0.106	0.895	0.247	0.163	0.096	0.292
F-test (joint significance of all 3 treatment groups)	3.222	2.425	0.292	0.137	1.562	0.634	2.130	0.907	1.144	4.065
P-value (joint significance of all 3 treatment groups)	0.023	0.065	0.831	0.938	0.198	0.594	0.096	0.438	0.331	0.007

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of columns (2)-(10). The latter are z-scores of (2) whether household members reported that, during their most recent visit to the health center, equipment was used during examination, (3) waiting time consisting of the total amount of time spent by the household members waiting for the initial consultation and the examination; whether household members declared that, during their most recent visit to the health center, (4) they were examined by trained health care staff, (5) they had privacy during their examination, (6) lab tests were administered, (7) their diagnosis was clearly explained to them; (8) percent of staff in attendance during an unannounced visit to the health center, measured at the health center level, (9) condition of the clinic (cleanliness of floors and walls, whether the clinic smelled as observed during unannounced visit to health center), measured at the health center level, (10) share of months in which stock cards indicated availability of six key tracer drugs in the past three months, measured at the health center level. *** p<0.01; ** p<0.05; * p<0.10

Table 42: Patient satisfaction outcomes – All treatment arms

	(1) **Patient satisfaction index**	(2) Satisfied by HC quality	(3) Satisfied with quality of care	(4) Polite staff	(5) Staff interested in health	(6) Staff listening	(7) Free to express clearly	(8) Availability of staff improving
Full treatment	0.080***	0.105***	0.063*	0.074**	0.104***	0.076**	0.043	0.088**
	(0.024)	(0.035)	(0.034)	(0.036)	(0.030)	(0.031)	(0.040)	(0.041)
Information and mobilization only	0.073***	0.100***	0.056	0.054	0.114***	0.102***	0.077*	0.018
	(0.026)	(0.036)	(0.035)	(0.041)	(0.034)	(0.033)	(0.040)	(0.043)
Interface only	0.064***	0.067**	0.060*	0.048	0.093***	0.062**	0.089**	0.025
	(0.022)	(0.034)	(0.033)	(0.033)	(0.028)	(0.030)	(0.037)	(0.044)
Constant	-0.006	-0.011	0.006	0.005	0.009	0.007	0.008	-0.027
	(0.018)	(0.023)	(0.023)	(0.027)	(0.023)	(0.025)	(0.030)	(0.032)
N	14,609	14,609	14,609	14,609	14,609	14,609	14,609	14,609
R^2	0.040	0.071	0.040	0.019	0.016	0.020	0.016	0.057
P-value (Information and mobilization = Interface)	0.697	0.356	0.896	0.870	0.464	0.182	0.728	0.878
P-value (Information and mobilization = Full treatment)	0.778	0.905	0.839	0.608	0.740	0.403	0.360	0.075
P-value (Interface = Full treatment)	0.441	0.285	0.934	0.397	0.662	0.612	0.174	0.109
F-test (joint significance of all 3 treatment groups)	4.435	3.914	1.662	1.508	5.322	3.403	2.225	1.943
P-value (joint significance of all 3 treatment groups)	0.004	0.009	0.175	0.212	0.001	0.018	0.085	0.122

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of columns (2)-(8). (The latter are z-scores of indicator variables of whether household members declared that (2) the services currently offered at the health center are of "very high quality" or "somewhat high quality", (3) they were "very satisfied" or "satisfied" with the quality of care received during their most recent visits to the health center, (4) during their most recent visit to the health center, the person conducting the examination behaved politely/showed respect, (5) during their most recent visit to the health center, the person conducting the examination appeared to be interested in their health condition, (6) during their most recent visit to the health center, the person conducting the examination listened to what they had to say, (7) during their most recent visit to the health center, they felt free to express themselves to the person conducting the examination, (8) compared to the year before, the availability of medical staff had improved at the health center. *** $p<0.01$; ** $p<0.05$; * $p<0.10$

Table 43: Health outcomes at the child level – All treatment arms

	(1) **Health outcomes index**	(2) **Weight/Age 0-18 months**	(3) **MUAC 0-18 months**	(4) **Weight/Age 18-36 months**	(5) **MUAC 18-36 months**
Full treatment	-0.003	-0.001	-0.013	0.003	0.019
	(0.028)	(0.048)	(0.048)	(0.032)	(0.028)
Information and mobilization only	-0.023	-0.036	-0.026	-0.011	-0.011
	(0.029)	(0.048)	(0.048)	(0.035)	(0.029)
Interface only	-0.011	-0.019	-0.023	0.014	0.026
	(0.028)	(0.046)	(0.046)	(0.036)	(0.032)
Constant	-0.488***	0.005	0.005	-0.461***	-0.630***
	(0.022)	(0.038)	(0.038)	(0.026)	(0.020)
N	10,023	4,379	4,379	5,644	5,644
R^2	0.103	0.011	0.012	0.207	0.328
P-value (Information and mobilization = Interface)	0.653	0.661	0.933	0.500	0.258
P-value (Information and mobilization = Full treatment)	0.472	0.408	0.753	0.687	0.303
P-value (Interface = Full treatment)	0.767	0.651	0.805	0.739	0.840
F-test (joint significance of all 3 treatment groups)	0.251	0.293	0.122	0.154	0.590
P-value (joint significance of all 3 treatment groups)	0.861	0.831	0.947	0.927	0.622

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. Robust standard errors are clustered at the health center level. The dependent variable in column (1) is the averaged z-score index of columns (2)-(7). The latter are z-scores of (2) the average ratio of weight over number of months for children under 18 months, (3) the average ratio of weight over number of months for children 18-36 months old, (4) the average ratio of upper arm circumference over number of months for children under 18 months, (5) the average ratio of upper arm circumference over number of months for children 18-36 months old. *** p<0.01; ** p<0.05; * p<0.10

Table 44: Child mortality at the HC level

	(1) Child mortality 0-5 years old	(2) Child mortality 0-1 year old	(3) Child mortality 1-5 years old
Full treatment	-0.011	-0.006	-0.006
	(0.008)	(0.007)	(0.004)
Information and mobilization only	-0.020**	-0.013**	-0.007
	(0.008)	(0.006)	(0.004)
Interface only	-0.009	-0.004	-0.006
	(0.008)	(0.007)	(0.005)
Constant	0.061***	0.042***	0.020***
	(0.006)	(0.005)	(0.003)
N	376	376	376
R^2	0.151	0.160	0.159
P-value (Information and mobilization = Interface)	0.204	0.212	0.753
P-value (Information and mobilization = Full treatment)	0.281	0.334	0.738
P-value (Interface = Full treatment)	0.820	0.797	0.992
F-test (joint significance of all 3 treatment groups)	2.225	1.509	0.887
P-value (joint significance of all 3 treatment groups)	0.085	0.212	0.448

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. The dependent variable is the child mortality rate in the health center catchment area calculated with the synthetic cohort approach, in the age bracket 0-5 years (1), 0-12 months (2), and 1-5 years (3), respectively. The unit of analysis is the health center catchment area. *** p<0.01; ** p<0.05; * p<0.10

Table 45: Child mortality at the child level

	(1) **0-5 years old**	(2) **0-1 year old**	(3) **1-5 years old**
Full treatment	1.080	1.150	0.614
	(0.241)	(0.301)	(0.292)
Information and mobilization only	0.960	0.900	0.189***
	(0.228)	(0.255)	(0.080)
Interface only	1.166	1.145	0.970
	(0.245)	(0.278)	(0.396)
N	20,371	9,171	17,363
R^2			
P-value (Information and mobilization = Interface)	0.367	0.322	0.000
P-value (Information and mobilization = Full treatment)	0.606	0.352	0.012
P-value (Interface = Full treatment)	0.702	0.986	0.323
F-test (joint significance of all 3 treatment groups)	0.998	1.278	21.373
P-value (joint significance of all 3 treatment groups)	0.802	0.734	0.000

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. The dependent variable is the child mortality rate in the health center catchment area calculated with the synthetic cohort approach, in the age bracket 0-5 years (1), 0-12 months (2), and 1-5 years (3), respectively. The unit of analysis is the health center catchment area. *** p<0.01; ** p<0.05; * p<0.10

Table 46: Robustness check – Excluding control variables interacted with treatment

	(1) **0-5 years old**	(2) **0-1 year old**	(3) **1-5 years old**
Full treatment	1.042	1.131	0.746
	(0.208)	(0.267)	(0.227)
Information and mobilization only	0.894	0.952	0.730
	(0.197)	(0.242)	(0.236)
Interface only	1.013	1.071	0.769
	(0.201)	(0.243)	(0.245)
N	20,371	9,171	17,363
R^2			
P-value (Information and mobilization = Interface)	0.537	0.595	0.876
P-value (Information and mobilization = Full treatment)	0.450	0.452	0.944
P-value (Interface = Full treatment)	0.879	0.787	0.924
F-test (joint significance of all 3 treatment groups)	0.614	0.658	1.288
P-value (joint significance of all 3 treatment groups)	0.893	0.883	0.732

Notes. Estimates comparing outcomes between each treatment arm and the control from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. The dependent variable is the child mortality rate in the health center catchment area calculated with the synthetic cohort approach, in the age bracket 0-5 years (1), 0-12 months (2), and 1-5 years (3), respectively. The unit of analysis is the health center catchment area. *** p<0.01; ** p<0.05; * p<0.10

Table 47: Intermediates outcomes – All treatments

	(1) **Citizen knowledge**	(2) **HC staff knowledge**	(3) **Efficacy**	(4) **Community responsibility**	(5) **Relationship**	(6) **HC transparency**
Full treatment	-0.054**	0.140	-0.019	-0.011	-0.051*	-0.026
	(0.024)	(0.118)	(0.023)	(0.021)	(0.028)	(0.077)
Information and mobilization only	0.001	0.226**	0.006	0.001	-0.032	-0.126
	(0.024)	(0.110)	(0.023)	(0.022)	(0.032)	(0.077)
Interface only	-0.031	0.107	-0.011	-0.025	-0.037	-0.104
	(0.023)	(0.110)	(0.022)	(0.021)	(0.029)	(0.080)
Constant	-0.019	-0.012	-0.003	-0.005	0.010	-0.010
	(0.017)	(0.079)	(0.015)	(0.015)	(0.019)	(0.055)
N	14,609	376	14,609	14,609	14,609	376
R^2	0.193	0.191	0.044	0.047	0.026	0.402
P-value (Information = Interface)	0.179	0.266	0.467	0.240	0.880	0.784
P-value (Information = Full treatment)	0.032	0.457	0.296	0.566	0.558	0.198
P-value (Interface = Full treatment)	0.327	0.772	0.708	0.510	0.647	0.336
F-test (joint significance of all 3 treatment groups)	2.297	1.438	0.447	0.646	1.242	1.194
P-value (joint significance of all 3 treatment groups)	0.077	0.232	0.720	0.586	0.294	0.312

Notes. Estimates comparing outcomes between each treatment arm and the control group for intermediate outcome indices from Equation 3. Each treatment arm enters as an indicator variable. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

Pre-specified Specification

In addition, we show the main results from the pre-specified model, which includes an interaction term rather than estimating average treatment effects for each cell more flexibly, below.

Table 48: Main outcomes – All treatments

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**	(7) **Child mortality**
Information	0.036	-0.043	0.013	0.013	0.073***	-0.023	-0.020**
	(0.030)	(0.092)	(0.022)	(0.029)	(0.026)	(0.029)	(0.008)
Interface	0.041	-0.013	0.054**	0.022	0.064***	-0.011	-0.009
	(0.029)	(0.087)	(0.022)	(0.027)	(0.022)	(0.028)	(0.008)
Information x Interface	-0.074*	-0.074	-0.039	0.036	-0.058*	0.030	0.018
	(0.042)	(0.139)	(0.031)	(0.037)	(0.033)	(0.039)	(0.011)
Constant	0.005	0.002	-0.018	-0.002	-0.006	-0.488***	0.061***
	(0.020)	(0.063)	(0.015)	(0.021)	(0.018)	(0.022)	(0.006)
N	14,609	376	14,609	14,609	14,609	10,023	376
R^2	0.077	0.103	0.221	0.102	0.040	0.103	0.151
Informantion + Information x Interface	-0.038	-0.117	-0.026	0.049**	0.016	0.008	-0.002
	(0.029)	(0.102)	(0.022)	(0.023)	(0.020)	(0.025)	(0.008)
P-value (Information = Interface)	0.874	0.735	0.066	0.740	0.697	0.653	0.204

Notes. Estimates comparing outcomes between each treatment arm and the control. All models include district fixed effects as well as demeaned baseline covariates and their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

Table 49: Intermediates outcomes: subcomponents

	(1) **Citizen knowledge**	(2) **HC staff knowledge**	(3) **Efficacy**	(4) **Community responsibility**	(5) Relationship	(6) HC transparency
Information	0.001	0.226**	0.006	0.001	0.063	-0.126
	0.024	0.110	0.023	0.022	0.039	0.077
Interface	-0.031	0.107	-0.011	-0.025	0.020	-0.104
	0.023	0.110	0.022	0.021	0.042	0.080
Information x Interface	-0.024	-0.193	-0.015	0.012	-0.044	0.203*
	0.034	0.159	0.033	0.030	0.057	0.111
Constant	-0.019	-0.012	-0.003	-0.005	0.004	-0.010
	0.017	0.079	0.015	0.015	0.029	0.055
N	14609	376	14609	14609	14609	376
R^2	0.193	0.191	0.044	0.047	0.050	0.402
Information + Information x Interface	-0.023	0.033	-0.009	0.013	0.019	0.078
	(0.023)	(0.115)	(0.023)	(0.020)	(0.042)	(0.080)
P-value (Information = Interface)	0.179	0.266	0.467	0.240	0.296	0.784

Notes. Estimates comparing outcomes between each treatment arm and the control group for intermediate outcome indices. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

H.9 Midline Results

The tables that follow show the treatment effect on main and intermediate outcome indices at midline, both for the comparison between treatment and the full intervention and for all treatment arms.

Main Results

Table 50: Main outcomes (midline) - Full treatment

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**	(7) **Child mortality**
Full treatment	0.018	-0.132	-0.012	0.060*	0.036	0.024	-0.014
	(0.024)	(0.129)	(0.020)	(0.033)	(0.029)	(0.027)	(0.010)
Constant	-0.001	-0.005	-0.012	-0.002	-0.000	-0.153***	0.054***
	(0.018)	(0.086)	(0.015)	(0.024)	(0.019)	(0.023)	0.008
N	7,204	187	7,204	7,204	7,204	5,337	187
R^2	0.050	0.113	0.255	0.095	0.049	0.102	0.204
P-value (Full treatment = 0)	0.470	0.307	0.560	0.066	0.207	0.382	0.166
Adjusted p-value (FT)	0.548	0.534	0.560	0.463	0.482	0.534	0.482

Notes. Estimates from Equation 1 comparing outcomes measured at midline between the Full treatment arm and the Control. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table 51: Intermediate outcomes (midline) – Averaged z-score indices

	(1) **Citizen knowledge**	(2) **HC staff knowledge**	(3) **Efficacy**	(4) **Community responsibility**	(5) **Relationship**	(6) **HC transparency**
Full treatment	0.009	0.084	0.039**	0.009	-0.041	-0.090
	(0.023)	(0.099)	(0.019)	(0.021)	(0.043)	(0.064)
Constant	-0.009	-0.006	-0.001	-0.005	0.003	-0.016
	(0.017)	(0.072)	(0.014)	(0.015)	(0.026)	(0.050)
N	7,204	187	7,204	7,204	7,204	187
R^2	0.157	0.442	0.044	0.044	0.095	0.352
P-value (Full treatment = 0)	0.693	0.395	0.045	0.684	0.339	0.163
Adjusted p-value (FT)	0.693	0.593	0.273	0.693	0.593	0.488

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control for intermediate outcome indices measured at midline. All models include district fixed effects and demeaned baseline covariates, as well as their interaction with the treatment indicator. Robust standard errors are clustered at the health center level. Adjusted p-values (FT) refer to p-values of the coefficient on Full Treatment which are adjusted using the Benjamini-Hochberg method. *** p<0.01; ** p<0.05; * p<0.10

Table 52: Robustness checks – Main outcomes (midline)

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**	(7) **Child mortality**
A: Without control variables							
Treatment	0.018	-0.143	-0.006	0.057*	0.035	0.021	-0.013
	(0.024)	(0.128)	(0.022)	(0.034)	(0.029)	(0.028)	(0.010)
Constant	-0.001	0.001	-0.015	-0.001	0.001	0.000	0.055***
	(0.018)	(0.083)	(0.017)	(0.025)	(0.020)	(0.020)	0.008
B: Without district fixed effects							
Treatment	0.013	-0.142	-0.018	0.062*	0.038	0.021	-0.012
	(0.027)	(0.130)	(0.026)	(0.035)	(0.030)	(0.029)	(0.010)
Constant	0.001	0.000	-0.009	-0.003	-0.001	-0.000	0.055***
	(0.021)	(0.088)	(0.018)	(0.025)	(0.021)	(0.020)	0.008
C: With outcome measures aggregated at HC level							
Treatment	0.015	-0.132	-0.011	0.050	0.039	0.023	-0.014
	(0.026)	(0.129)	(0.021)	(0.035)	(0.030)	(0.028)	(0.010)
Constant	-0.000	-0.005	-0.003	0.002	0.000	0.003	0.054***
	(0.020)	(0.086)	(0.015)	(0.025)	(0.021)	(0.020)	(0.008)
D: Difference between post and pre-treatment values							
Treatment	0.021	-0.125	-0.035	0.065*	0.029	0.020	-0.011
	(0.027)	(0.126)	(0.024)	(0.035)	(0.030)	(0.028)	(0.010)
Constant	0.027	0.041	-0.023**	0.021	-0.008	0.014	0.045***
	(0.020)	(0.055)	(0.012)	(0.017)	(0.015)	(0.020)	(0.005)
Observations (A & B)	7,204	187	7,204	7,204	7,204	5,337	187
Observations (C)	187	187	187	187	187	187	187
Observations (D)	14,408	374	14,408	14,408	14,408	10,674	374

Notes. Estimates from Equation 1 comparing outcomes between the Full treatment arm and the Control, with the following variations: Panel A shows results without covariates, panel B without district fixed effects, and panel C aggregates outcome measures and covariates at the health center level (the unit of randomization). Panel D shows results from a difference in difference estimation. *** p<0.01; ** p<0.05; * p<0.10

Table 53: Robustness checks – Intermediate outcomes (midline)

	(1) **Citizen knowledge**	(2) **HC staff knowledge**	(3) **Efficacy**	(4) **Community responsibility**	(5) **Relationship**	(6) **HC transparency**
A: Without control variables						
Program impact	-0.001	0.046	0.038*	0.007	-0.041	-0.117*
	(0.031)	(0.103)	(0.020)	(0.022)	(0.045)	(0.070)
Constant	-0.005	0.010	-0.000	-0.004	0.002	-0.001
	(0.022)	(0.076)	(0.014)	(0.016)	(0.027)	(0.058)
B: Without district fixed effects						
Program impact	-0.007	0.088	0.037	-0.004	-0.038	-0.103
	(0.037)	(0.110)	(0.023)	(0.027)	(0.044)	(0.064)
Constant	-0.001	-0.009	0.000	0.001	0.001	-0.011
	(0.025)	(0.083)	(0.017)	(0.019)	(0.027)	(0.051)
C: Without ouctome measures aggregated at HC level						
Program impact	0.008	0.084	0.036*	0.009	-0.051	-0.090
	(0.025)	(0.099)	(0.022)	(0.024)	(0.049)	(0.064)
Constant	-0.011	-0.006	0.003	0.001	0.001	-0.016
	(0.018)	(0.072)	(0.015)	(0.017)	(0.030)	(0.050)
D: Difference between post and pre-treatment values						
Program impact	-0.005	0.117	0.035*	0.002	-0.060	-0.109*
	(0.038)	(0.104)	(0.021)	(0.022)	(0.046)	(0.065)
Constant	-0.003	0.044	-0.016	-0.013	-0.042	0.001
	(0.019)	(0.057)	(0.011)	(0.010)	(0.026)	(0.041)
Observations (A & B)	7204	187	7204	7204	7204	187
Observations (C)	187	187	187	187	187	187
Observations (D)	14408	374	14408	14408	14408	374

Midline Results by Treatment Arm

Table 54: Main outcomes (Midline) - All treatments

	(1) **Citizen monitoring**	(2) **Perceived citizen pressure**	(3) **Utilization**	(4) **Treatment quality**	(5) **Patient satisfaction**	(6) **Health outcomes**	(7) **Child mortality**
Full treatment	0.011	-0.131	-0.015	0.061*	0.039	0.025	-0.014
	(0.025)	(0.131)	(0.021)	(0.033)	(0.029)	(0.028)	(0.010)
Information only	0.032	0.088	-0.031	0.066**	0.044	-0.015	-0.021**
	(0.025)	(0.110)	(0.022)	(0.033)	(0.027)	(0.029)	(0.009)
Interface only	0.029	-0.108	-0.007	0.031	0.018	0.010	-0.011
	(0.025)	(0.128)	(0.023)	(0.033)	(0.028)	(0.029)	(0.010)
Constant	0.003	0.002	-0.009	-0.005	-0.003	-0.147***	0.054***
	(0.019)	(0.089)	(0.016)	(0.024)	(0.020)	(0.022)	(0.007)
N	14,459	376	14,459	14,459	14,459	10,787	376
R^2	0.037	0.113	0.239	0.091	0.048	0.093	0.153
P-value (Information = Interface)	0.894	0.081	0.275	0.263	0.342	0.408	0.269
P-value (Information = Full treatment)	0.384	0.059	0.422	0.871	0.849	0.154	0.392
P-value (Interface = Full treatment)	0.447	0.867	0.708	0.344	0.470	0.586	0.812
F-test (joint significance of all 3 treatment groups)	0.730	1.705	0.736	1.715	1.075	0.714	3.065
P-value (joint significance of all 3 treatment groups)	0.534	0.166	0.531	0.163	0.359	0.544	0.382

Notes. Estimates comparing midline outcomes between each treatment arm and the control. Each treatment arm enters as a separate indicator. All models include district fixed effects and demeaned baseline covariates, as well as their interactions with the treatment indicators. Robust standard errors are clustered at the health center level. *** p<0.01; ** p<0.05; * p<0.10

H.10 Results from T-Tests

This section shows results from two-sided t-tests of difference of means, with the unit of observation being the health center catchment area.

Table 55: Balance Across Treatment Arms (Midline and Endline)

	(1) Full treatment	(2) Information only	(3) Interface only	(4) Control	(5) P-value difference (1) - (4)	(6) P-value difference (1) & (2) - (3) & (4)	(7) P-value difference (1) & (3) - (2) & (4)
A. Midline levels of main outcome indices							
Citizen monitoring	0.77	0.78	0.77	0.76	0.33	0.42	0.51
Perceived citizen pressure	-0.14	0.10	-0.10	0.00	0.25	0.71	0.04
Utilization	5.96	5.88	6.28	6.22	0.37	0.08	0.70
Treatment quality	0.77	0.76	0.76	0.75	0.18	0.14	0.67
Patient satisfaction	0.77	0.77	0.77	0.76	0.26	0.17	0.80
Health outcomes	1.03	1.01	1.05	0.98	0.22	0.97	0.06
Mortality	0.04	0.04	0.05	0.06	0.26	0.12	0.90
B. Midline levels of intermediate outcome indices							
Citizen knowledge	0.40	0.39	0.40	0.40	0.84	0.88	0.89
Health worker knowledge	0.40	0.38	0.37	0.38	0.47	0.45	0.84
Efficacy	0.64	0.64	0.65	0.63	0.16	0.72	0.13
Community responsibility	0.28	0.29	0.28	0.28	0.91	0.97	0.84
Relationship between health workers and community	0.80	0.82	0.81	0.81	0.44	0.70	0.12
Health center transparency	0.19	0.22	0.23	0.22	0.22	0.27	0.58
C. Endline levels of main outcome indices							
Citizen monitoring	0.73	0.74	0.74	0.73	0.72	0.91	0.52
Perceived citizen pressure	-0.11	-0.03	-0.01	0.01	0.23	0.28	0.49
Utilization	6.16	6.19	6.47	6.36	0.55	0.27	0.83
Treatment quality	0.79	0.77	0.78	0.77	0.01	0.27	0.03
Patient satisfaction	0.78	0.78	0.77	0.76	0.01	0.03	0.15
Health outcomes	1.07	1.05	1.06	1.08	0.86	0.80	0.97
Mortality	0.05	0.05	0.06	0.06	0.31	0.11	0.89
D. Endline levels of intermediate outcome indices							
Citizen knowledge	0.43	0.44	0.43	0.45	0.12	0.45	0.13
Health worker knowledge	0.35	0.36	0.34	0.32	0.29	0.12	0.95
Efficacy	0.65	0.66	0.65	0.66	0.29	0.66	0.29
Community responsibility	0.33	0.33	0.33	0.34	0.60	0.92	0.39
Relationship between health workers and community	0.80	0.81	0.80	0.79	0.17	0.17	0.62
Health center transparency	0.28	0.24	0.24	0.30	0.68	0.72	0.78
N	92	92	97	95			

H.11 Multiple Comparison Corrections

Given the number of outcome variables in our study, multiple testing is a concern. Main tables also include corrected p-values for the average effect of the full treatment, calculated using the Benjamini and Hochberg (1995) False Discovery Rate correction, in the bottom panel. This simple step-up procedure is slightly less punitive than a Bonferroni correction since it focuses exclusively on correcting for the false discovery rate (type I errors).

For outcome indices, the family is defined as the set of main outcome indices or the set of intermediate outcome indices, respectively. For components of an index, the family is defined as the set of components of a given index.

H.12 Two One-Sided Tests (TOST) Procedure

Figure 9 shows results from the Two One-Sided Tests (TOST) procedure developed in Schuirmann (1987) and further explained in Lakens (2017). We choose this approach to interrogate our null effects since, for the reasons outlined in Hoenig and Heisey (2001), post-experiment power calculations, often used to determine statistical power post-hoc, are problematic. We conduct equivalence tests (Wald-tests) of the estimated treatment effect with a critical value λ. The two one-sided test procedure means that we test two null hypotheses: (i) that the estimated treatment effect is smaller or equal to $-\lambda$ and (ii) that it is greater or equal to λ. We then report the larger of the two p-values for each λ. When both null hypotheses can be rejected—conventionally, when the larger p-value is smaller than or equal to 0.05—we conclude that the true treatment effect is between $-\lambda$ and λ.

Since the choice of λ is somewhat arbitrary and since we did not pre-specify any particular critical value below which we deem our effect sizes substantively meaningless, we conduct this test for a range of of critical values. The figures, modeled after those in Clayton, De Kadt and Dumas (2019), plot the maximum p-value from the equivalence test against lambda.

As can be seen, we can reject the null hypothesis of even very small treatment effects for citizen monitoring, utilization, child health outcomes, and child mortality. The significant positive treatment effects on treatment quality and patient satisfaction from the main analyses are born out by the TOST graphs as well. For citizen pressure as perceived by health center staff, we are unable to reject even larger treatment effects due to the relatively small sample size (n=187).

I Implementation of ACT Health

I.1 Map of implementation districts

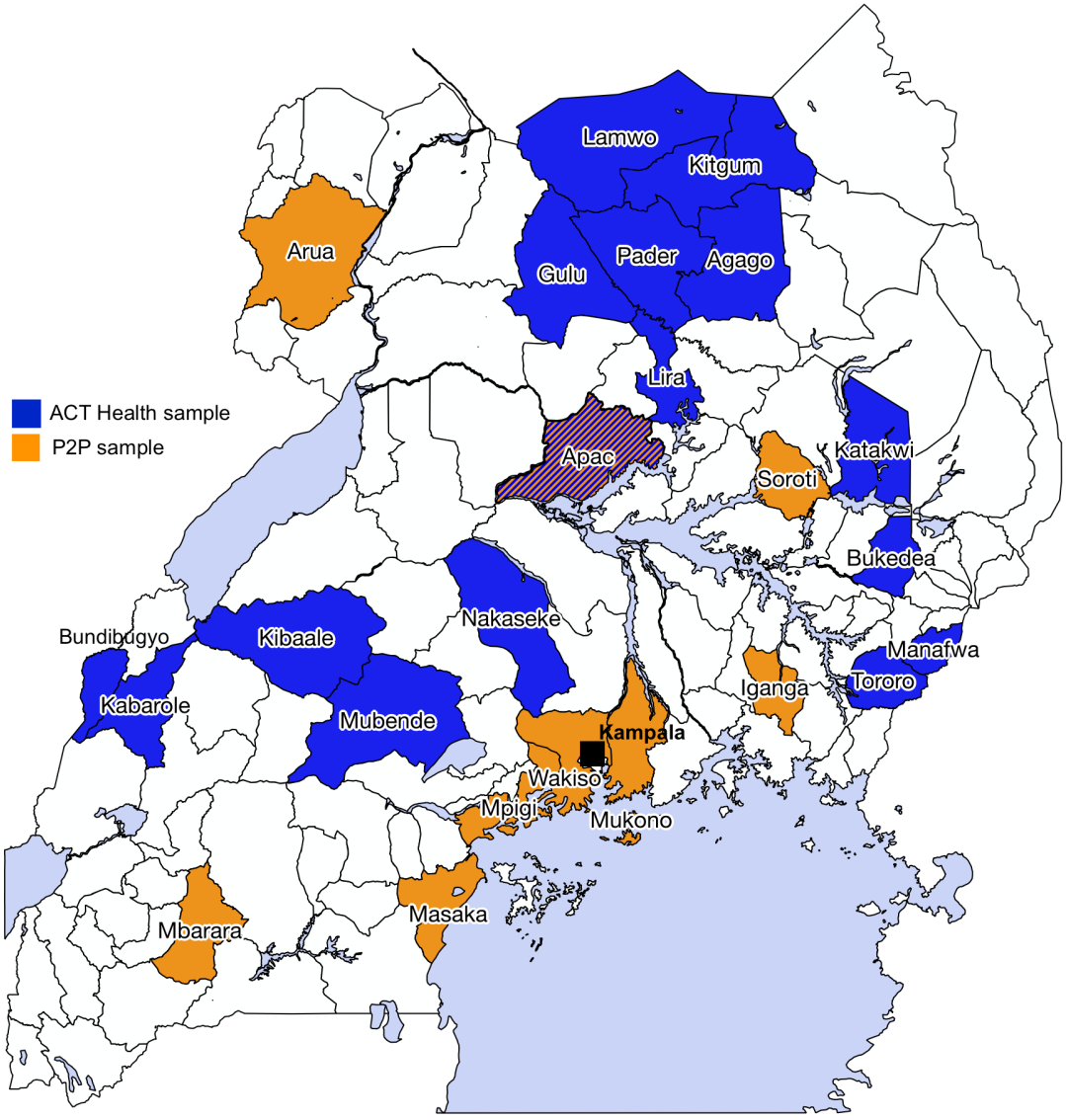


Figure 10: Map indicating the districts in the ACT Health sample, with disctricts that were included in the Björkman and Svensson (2009) study also indicated for reference purposes

I.2 Implementing Organizations

The Irish NGO, GOAL, launched the ACT Health program in 2014 with funding from DFID. The program was implemented by three Ugandan regional partners across 15 districts and by GOAL

Uganda in one district. All three regional partners had prior experience working on community mobilization and/or public health, and already had a strong footprint in the regions (if not the districts) in which they implemented ACT Health. In particular, the Coalition for Health Promotion and Social Development (HEPS), which was founded in 2000, focused on access to health care and essential medicines, maternal health rights, community-based empowerment work, and health advocacy prior to becoming involved in the implementation of ACT Health. The Multi-Community Based Development Initiative (MUCOBADI) was started by HIV positive teachers in 2000 and focused on HIV prevention, access to primary health care, community mobilization, and livelihood development. Finally, the Kabarole Research and Resource Centre (KRC), which was founded in 1996, focused on leadership mobilization, good governance, and research and advocacy. Two of the three regional partners, HEPS and MUCOBADI, were actively involved in the design and implementation of a pilot of the ACT Health program in Bugiri district. Organizational leadership aside, implementing staff for the intervention were specifically recruited for the program and had to have prior experience in community mobilization and/or public health. All implementing staff underwent extensive training and were continuously monitored and supervised by GOAL Uganda.

I.3 Implementation Monitoring

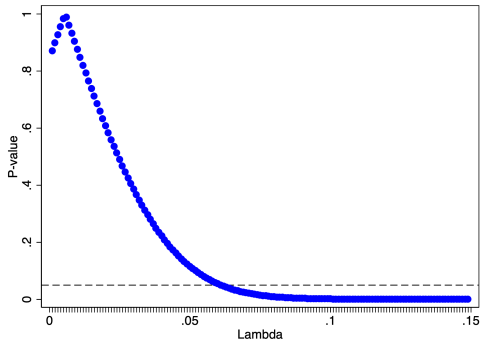
GOAL ensured fidelity to both the intervention and the randomized impact evaluation protocol through several quality control measures.

- Detailed procedure manuals for each of the three variations of the program (the full program, the information and mobilization program—called separate dialogues by GOAL—and the interface-only program). These manuals were used for training and as a reference to ensure that all partners had clear and precise instructions regarding every detail of the intervention. These manuals were used for training and as a reference so that all implementing partners had clear and precise instructions of every detail of the interventions.
- Extensive monitoring data captured in an online monitoring database, which tracked the dates and numbers of people participating in each dialogue. The database also includes the actions agreed upon in the action plans and social contracts developed during the dialogues, and tracked their progress at each follow-up meeting. The reports include data on the dates and number of people participating for each program activity, including HC and community meetings, interface meetings, and each of the follow-up meetings. They also record all the actions agreed to in the action plans and social contracts and track their progress at the follow-up meetings.
- Direct observation by GOAL's monitoring team. To assure quality across the life of the intervention, GOAL had "mentor" managers and monitoring, evaluation, and learning (MEL) officers embedded within the teams of each partner organization to provide direct support and programmatic guidance, as needed. Mentor manager and MEL officers observed a portion of each partner's dialogues and follow-ups for quality assurance purposes. In particular, 97% of health centers were monitored at least once by either a mentor manager or MEL officer (see Section E of Table 1 for more details). During these direct observations, officers recorded information about facilitator behavior, the presentation of the citizen report cards, the nature of participation during the meeting, and whether the action plans and social contracts met certain quality criteria. The feedback tool for these observations is reproduced below.

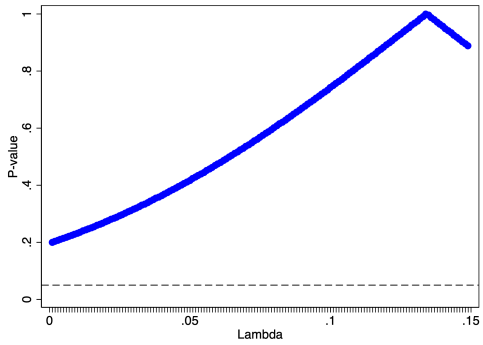
- Issues tracking. GOAL tracked issues as they came up during implementation. They gave field teams a detailed protocol of issues to watch out for and flag. Issues were shared with the evaluation team and solutions were jointly decided to resolve the issues.

Figure 9: Plotting p-values from two one-sided tests against lambda for main outcome indices at endline

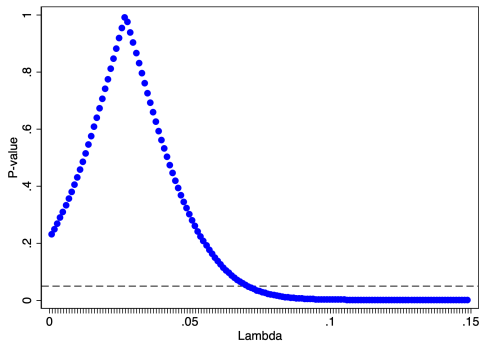
(a) Citizen monitoring



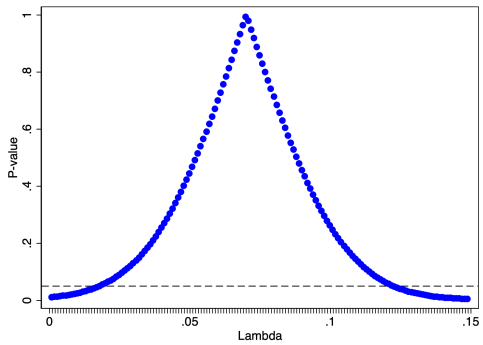
(b) Perceived citizen pressure



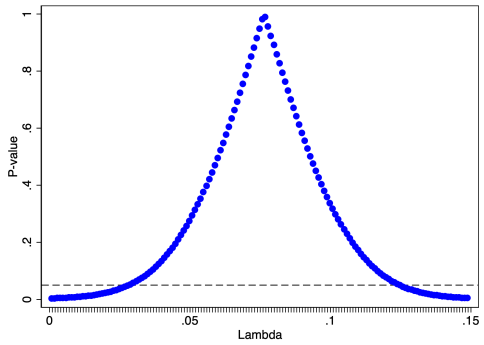
(c) Utilization



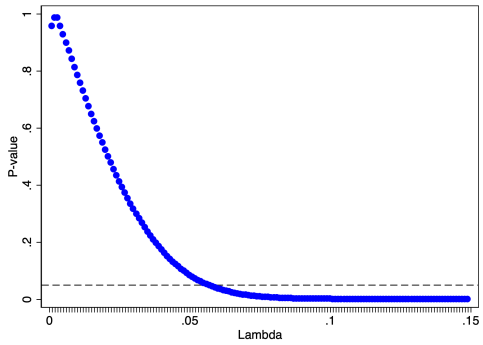
(d) Treatment quality



(e) Patient satisfaction



(f) Health outcomes



(g) Child mortality

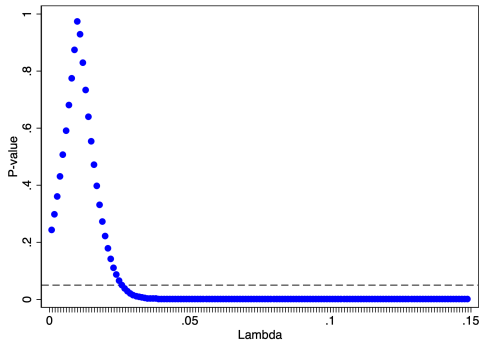


Table 1: GOAL attendance and monitoring data

	Full treatment			Information only			Interface only		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**
A. HC information separate dialogue									
# of HC staff present	90	6.14	3.26	92	6.24	3.42	n/a	n/a	n/a
% of HC staff present	90	67.25	33.89	92	71.04	27.56	n/a	n/a	n/a
B. Community information separate dialogue									
# of community members present	90	102.36	20.99	91	99.01	25.17	n/a	n/a	n/a
# of female community members present	90	35.67	8.79	91	34.26	9.91	n/a	n/a	n/a
% of HCs where at least one SC official is present	90	0.10	0.30	91	0.36	0.48	n/a	n/a	n/a
C. Interface meeting									
# of community members present	91	33.10	14.94	n/a	n/a	n/a	97	67.38	13.67
# of female community members present	90	16.87	8.83	n/a	n/a	n/a	97	34.79	10.49
# of HC staff present	91	3.97	2.31	n/a	n/a	n/a	97	3.80	2.45
% of HCs where at least one SC official is present	91	0.30	0.46	n/a	n/a	n/a	97	0.35	0.48
D. Follow-up meetings (average across three meetings)									
# of community members present	91	41.01	14.77	92	54.51	20.14	97	55.67	17.54
# of female community members present	91	20.54	8.13	92	27.37	11.01	97	29.21	11.41
# of HC staff present	91	3.71	2.22	90	3.14	3.30	97	3.46	2.10
% of HCs where at least one SC official is present	91	0.39	0.54	.	.	.	97	0.47	0.60
E. Monitoring & oversight									
% supervision during the initial activity	92	0.80	0.40	92	0.62	0.49	97	0.58	0.50
% supervision during at least one follow-up meeting	92	0.76	0.43	92	0.92	0.27	97	0.73	0.45
% supervision at least one time (initial activity or follow-up)	92	0.98	0.15	92	0.99	0.10	97	0.93	0.26

Notes. Data is drawn from implementers' monitoring tools, and verified by GOAL's monitoring team through direct observation in the share of meetings indicated in Panel E.

Version 10-Mar-2015

Observation-Feedback Tool for ACT Health Community Activities

*This form is for use by Line Managers and Mentorship Managers when observing **dialogues** and **interfaces.** Compete for each observation and share your feedback with the Officers after the visit. Copy should be kept in file of Officer and Observer.*

Observerer	Name			Title		Organisation	
Date of observation	Day	Month	Year	**Procedure # (Tick One ONLY)**	#2	#3	#4
Location	Health Centre			Sub-county		District	
Activity observed	h of WDWW. Linking the HC Dialogue		Community Dialogue	Interface			
Facilitators Observed	Name		Organisation	Name		Organsiation	

	Was the Facilitator . . .	YES	NO	Partly
Facilitator behaviour	Well-prepared on the process and purpose of meeting?			
	Effective in managing expectations of participants?			
	Clear and audible so participants understood?			
	Managing time? (Covered all agenda items, not exceed 5 hrs)			
	Listening actively? (Eye contact, summarising, repeating)			
	Enabling diverse participants to speak and participate?			
	Ensuring mutual respect among participants?			
	Bringing focus to Responsibility of community members?			
	Bringing focus to Responsiveness of HC staff?			
	Impartial? (non-biased, non-judgemental, not taking sides)			
	Managing challenges? (conflicts, disturbances)			
CRC Presentation **Applicable to** Separate Dialogues Full Programme Observer must also study the CRC before meeting	**Did the Facilitator(s)…**	**YES**	**NO**	**Partly**
	Clearly understand the CRC content before the meeting?			
	Know and clearly communicate the MoH standards?			
	Bring the appropriate posters for the HC level?			
	Present clearly? (word choice, local language, explain boxes)			
	Check to ensure that participants understood the CRC?			
	Use CRC information to challenge perceptions and excuses?			
	Use probing questions to improve Responsibility?			
	Use probing questions to improve Responsiveness?			
	Use probing questions to improve Relationships?			
	Ensure mis-conceptions raised by participants are corrected			
Participation	**How participatory was the meeting?**	**YES**	**NO**	**Partly**
	Were *all* HC staff present?			
	Were all social groups represented as per mobilisation list?			
	Were all small groups working effectively?			
	Was the participation in the large group balanced? Women talked?			
	Was the posture and positioning of facilitator conducive?			
Action Plan/Social Contract	**Action plan/social contract meet basic quality criteria**	**YES**	**NO**	**Partly**
	Issues are clearly stated			
	Issues are related to the CRC/Information in CRC (Procedure #2 & #4)			
	Actions are related to the issue			
	Actions are achievable with local resources (low/no cost)			
	Inclusive of actions for community AND health centre staff			
	How many priority issues from women's groups are included?			
	Inclusive of issues/actions from all social groups?			
	Mix of actions for now (6months), soon (12 months) and later			

Observation-Feedback Tool for ACT Health Community Activities

*This form is for use by Line Managers and Mentorship Managers when observing **dialogues** and **interfaces.** Compete for each observation and share your feedback with the Officers after the visit. Copy should be kept in file of Officer and Observer.*

Observerer								
	Name			Title			Organisation	
Date of observation				**Procedure # (Tick One ONLY)**				
	Day	Month	Year		#2	#3	#4	
Location								
	Health Centre			Sub-county			District	
Activity observed	h of WDWW. Linking the							
	HC Dialogue		Community Dialogue		Interface			
Facilitators Observed								
	Name		Organisation		Name		Organsiation	

Feedback for Facilitator/Officer

The ACT Health programme has a component of on-the-job training and support. This support (mentorship) is very important for continuous learning and implementation of a high quality programme. These feedback tips should be shared with the faciltiators(s) after the meeting or in a visit to the Officer shortly (within three (3) working days) of the observation.

WELL DONE! These are the areas where you excelled. Thanks for your work!	
1)	
2)	
3)	
There are a few things you can work on for next time. Let me know how I can support you best.	
1)	
2)	
3)	

As an observer, what do you think participants in this community feel about the ACT Health programme?

Observer	
Name	
Signature	
Date of sharing with facilitator	

Facilitator	
Name	
Signature	
Date of discussing with facilitator	

I.4 Intervention Materials

I.4.1 Steps of the intervention

Procedures Table

Procedure		Citizen Report Card (CRC)	Health Centre Dialogue	Community Dialogue	Interface	Final Output Action Plan or Social Contract	Follow-up every six (6) months
1	No intervention (control)	Baseline data will be collected but no activities.	None will be held.	None will be held.	None will be held.	None will be developed.	*Survey at 12 months and 36 months after baseline.
2	Information provided (CRCs) and two separate action plans are developed in community and health centre dialogues. There is no interface between HC staff and community.	The Citizens Report Card will be shared in community dialogue and health centre dialogue.	Health centre staff have a dialogue and develop an action plan.	Community members have a dialogue and develop an action plan.	This will not be held.	Two separate action plans will be developed – one by the health centre staff and one by the community members.	Every six months separate follow-up dialogues for community members and health centre staff. *Survey at 12 months and 36 months after baseline.
3	No information (CRC) provided and no health centre or community dialogues are held. Interface between health centre staff and communities yields one social contract.	The Citizens Report Card will not be shared.	This will not be held.	This will not be held.	The interface will bring together community members and health centre staff.	One social contract developed at the interface combining community and HC actions.	Every six months follow-up interface with community members and HC staff. *Survey at 12 months and 36 months after baseline.
4	**Information (CRC) provided in separate dialogues at health centre and community levels. During these dialogues, each group develops an action plan which is discussed at the interface. The interface yields one social contract.**	**The Citizens Report Card will be shared at community dialogue and health centre dialogue.**	**Health centre staff dialogue and develop an action plan.**	**Community dialogue and develop an action plan.**	**The interface will bring together health centre staff and representatives of the communities.**	**One social contract developed combining community and HC actions.**	**Every six months follow-up interface with community members and HC staff jointly. *Survey at 12 months and 36 months after baseline.**

ACT Health Procedure #4 Full Programme

6

I.4.2 Example of a mobilization protocol and citizen report card

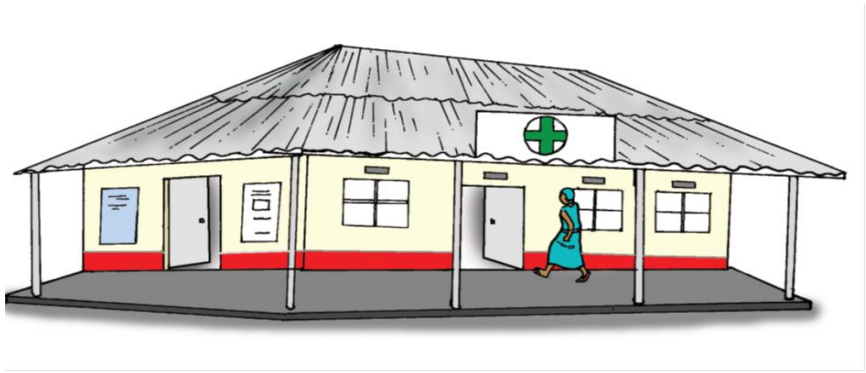
Version: 27-August-2014

Community Mobilisation List

Mobilising diverse social groups (women and men of different ages, income levels and social standing) is very important! We want to hear voices of all social groups in the community. Please think of someone in the village who meets the social group description below and would be willing to participate. Thank you for your time!

Health Centre	
Village	
Name of VHT carrying out mobilisation	

Dialogue Meeting Participants *at least 50% of the participants from each village should be women		
Social Group to Target	**Individual Name**	**Will attend interface meeting?**
1. LC1 Chairperson		
2. LC Women Representative		
3. LC Youth Representative (15-20 years old)		
4. LC Representative with disability		
5. Mother		
6. Mother		
7. Mother		
8. Male Lowest Income Group		
9. Female Lowest Income Group		
10. Female youth (15-20) in Lowest Income Group		
11. Male youth (15-20) in Lowest Income Group		
12. Male Highest Income Group		
13. Female Highest Income Group		
14. Male youth (15-20)		
15. Female youth (15-20)		
16. Male adult (21-49)		
17. Female adult (21-49)		
18. Male elder (50+)		
19. Female elder (50+)		
20. VHT Member		
21. VHT Member		
22. VHT Member		
23. VHT Member		
24. VHT Member		
25. HUMC Member		



Kagote Health Center III

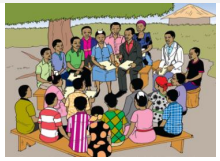
Kabarole District

Citizen Report Card

Survey dates: 10th October 2014 to 26th October 2014

Report Card Prepared: 27th October 2014

Responsibility	Responsiveness	Relationships
Individuals have good health-seeking behaviour. They seek preventive care (ANC, immunisations, testing, etc.) and go early for treatment of illness to avoid complications.	Health Center staff use resources effectively and provide care as per Ministry of Health standards in the Uganda National Minimum Health Care Package (UNMHCP).	Mutual understanding and trust between community members and health Center staff. Includes better understanding of each other's constraints.



Note: This Citizens Report Card has been compiled from responses to household surveys and HC staff interviews.

Rights and Responsibilities

Issue	Households say	Health Center says
Who could name at least 5 health rights and entitlements	0%	Could name 3

Health Rights

Some Major Health Rights in Uganda Patient's Charter
Right to choose
Right to complaint and redress
Right to access essential medicine
Right to access information
Right to privacy and confidentiality

Health Responsibilities

Health Responsibilities include
Responsibility to be healthy
Responsibility to participate

What services does our HC III provide?

GOVERNMENT STANDARD* Services that should be provided by HCIII	Health Center says
Antenatal care	Yes
Delivery	No
Outpatient care	Yes
HIV counselling and testing (HCT)	Yes
Immunisation	Yes
Lab services	Yes
Family planning methods (simple)	Yes
Family planning methods (advanced)	Yes
Health education (at HC)	Yes
Family planning education	Yes
Health Outreach (villages)	Yes
Prevention of Mother to Child Transmission (PMTCT)	Yes
Anti-retroviral therapy (ART)	Yes

*Uganda National Minimum Health Care Package

How many people use this HC? (Responsibility)

The community member visits to Kagote health Center III in the past 12 months.

Use patterns (adults and children)
17% of all health care visits in this community were to Kagote health Center

Reasons why community do not use health Center	
Reasons why the households in the community DO NOT visit Kagote health Center	60% Lack of drugs 20% Long waiting time 20% Long Distance 0% Cannot afford payment 20% Attitude of Staff 0% Unclean facility 10% Poor quality services 10% Have not been sick 10% Don't provide treatment I need 10% Others

Community member visits to other health providers in the past 12 months.

Other providers	Average utilisation (adults and children)
Private not for Profit (PNFP) e.g. NGO, missionary health Center	5% of all health care visits
Private for profit	19% of all health care visits
Traditional healer	1% of all health care visits
Community health worker e.g. VHT	6% of all health care visits
Self-treatment (pharmacy, drug shop)	32% of all health care visits
Other government health facilities e.g. HC III, IV, hospital	20% of all health care visits

How does our community compare?		
Health care provider	Kagote health Center	District use patterns of nearest government health centers
Use patterns	17%	31%

How many of us use ANC and deliver at our HC III? (Responsibility)

GOVERNMENT STANDARD = pregnant mothers should have four (4) ANC visits

Community's utilisation of antenatal care, family planning	
Percentage of households with pregnant women who have visited **Kagote health Center** for **antenatal care** since **September 2013**	63%
Percentage of those pregnant in the last year who **delivered** at **Kagote** health Center since **September 2013**	0%
Percentage of women who received an **HIV test** during ANC visit **(PMTCT)**	100%

Reasons why we (community members) do not deliver at this HC

Why do pregnant women in the community choose **NOT** to deliver at **Kagote** health center	**0%** Cannot afford **20%** Health Center was not open **0%** Use traditional birth attendant **0%** Attitude of staff **0%** Was not treated well at the HC **10%** Delivered quickly **20%** Referred to another health center **20%** Other provide better services **20%** Did not have the requirements **40%** Other

How do we compare? Antenatal care and maternity care		
Use pattern of antenatal care and maternity care	Among pregnant women in this community	Among pregnant women in Kabarole District
Percentage of households with pregnant women who have visited **their closest government** health Center for antenatal care	**63%**	**46%**
Percentage of pregnant women who made four (4) ANC visits to the nearest health center.	**10%**	**23%**

How do we compare? Immunisation		
Immunisation	In this community	*Among children in District*
% of children <5 immunised in **Kagote** catchment area	**98%**	**98%**

How many of us use family planning services at our HC III? (Responsibility)

Community's utilisation of family planning	
Percentage of households who have visited **Kagote health Center** for **family planning** since **September 2013**	29%

Why do households in the community choose NOT to use family planning services at Kagote health Center?	**0%** Attitude of staff **9%** Use natural methods **N/A** Not interested **Data not collected** **6%** Fear side effects **3%** Do not provide family planning education **19%** Do not need (young/want children/too old) **0%** Partner does not want **19%** Go elsewhere **3%** Health center lacks family planning drugs **3%** Did not know about the service **3%** Refused to answer **41%** Other

What community says about staff attendance at our HC III (Responsiveness)

GOVERNMENT STANDARD = absenteeism is any *unexcused* absence

Percentage of household saying medical staff attend work at Kagote health Center	
Always at work	71%
Sometimes at work	22%
Rarely at work	7%

GOVERNMENT STANDARD = HC III should have eleven (11) medical staff + eight (8) other staff for a total of nineteen (19) staff

Type of Staff	Government Standard	Staff actually allocated	Staff present on survey day
Medical	11	12	8
All staff	19	18	11

Medical staff attendance at Kagote health Center on survey day	
Total number of medical staff out on leave and/ or training on the survey day	2
Total number of medical staff out for outreach on the survey day	0
Percentage of households who said the health Center was open when they last visited	93%

What community says about drug availability

Household rating of drug availability

Household rating of drug availability at Kagote health Center	
Patients who received drugs at their last visit	88%
Drugs are always available	16%
Drugs are sometimes available	76%
Drugs are rarely available	9%

Do community members know when drugs are received?

Health issue	Households say	Health Center says
Do you know when drugs are delivered to Kagote health Center?	31% yes	Yes, we do distribute information on drug deliveries

Households reporting about the drugs they have	
Average number of type/brands of drugs received per visit per person	2
Percentage of patients who say it was clearly explained how to take the drugs	79%

GOVERNMENT STANDARD = All six (6) items should be available at all times

Health Center reporting stock outs of the following tracer items in the last 3 months	
1. Cotrimoxazole (CTX)	No
2. Artemether/Lumefantrine	No
3. Oral Rehydration salts (ORS)	Yes
4. Depo Provera	No
5. Measles Vaccine	No
6. Sulfadoxine and Pyrimethamine (SP)	No

Minimum standard drug storage conditions	
Method in place to control temperature	Yes
Windows that can be opened or there are air vents	Yes
Direct sunlight cannot enter the area	Yes
Area is free from moisture	Yes
Cold storage in the health Center	Yes
Medicines are stored directly on the floor	No
There is evidence of pests in the area	No

Fees at our HC (Responsiveness)

Government Standard	Health Center says	Community says	
0.00 UGX for government health facilities	No	Average amount paid	
		Cash	Value: In kind
		300 UGX	0 UGX

What did we bring / buy most?	
Top 3 things that have to be bought or brought to Kagote health Center	1.Exrecise book for prescription 2.N/A 3.N/A

Fees – HC III services	Households say	Health Center says	District Averages (Households say)
User fees (Cash)	2%	No	2%
Average amount paid for user fees (cash)	300 UGX	N/A	3,610 UGX
User fees (In-kind)	0%	N/A	0.1%
Average amount paid for user fees (in-kind)	N/A	N/A	1,200 UGX
Center charges for antenatal care (answered by pregnant women)	10%	No	1%
Average amount paid for antenatal care	300 UGX	0 UGX	1,650 UGX
Center charges for delivery (answered by women who delivered there)	0%	No	2%
Average amount paid for delivery	N/A	0 UGX	5,000 UGX
Center charges for drugs (including injections)	0%	No	0.1%
Average amount paid for drugs	N/A	0 UGX	1,800 UGX
Patients have to pay for immunization	0%	No	0.4%
Average amount paid for immunization	N/A	0 UGX	2,890 UGX

Satisfaction

Waiting times

GOVERNMENT STANDARD = waiting time should be less than one hour

Waiting time until first attended to

Government Standard	Community says	Health Center says
Less than 1 hour	00 Hour 39 Minutes	30 Minutes

Health Unit Management Committees (HUMCs)

Percentage of households who DO know at least two (2) roles of the HUMC	4%

Satisfaction with Relationships between HC Staff and Community

Overall satisfaction with relationship between community members and HC staff	Households say	HC says
Very satisfied	18%	Satisfied
Satisfied	60%	
Not satisfied	22%	

Health issue	Households say	Health Center says
Were patients treated politely	86% yes they were polite/ extremely polite	" Yes, we sometimes treat patients politely"
Average exam time for patients at their last visit	14 minutes	45 minutes

Health issue	Percentage
Percentage of patients who said the health worker listened to what they said at their last visit	89% yes/ very interested and asked questions
Percentage of patients who said the staff clearly explained their medical condition	70%
Percentage of patients who were examined at their last visit	65%
Percentage of patients who said health worker wore uniform at their last visit	59%
Percentage of patients who said they had privacy during the examination at their last visit	89%

How do we compare?		
Household says	Kagote health Center	District
Average waiting time for patients	39 minutes	46 minutes

I.4.3 Example of a community action plan, health center action plan, and social contract

Instructions
Please record below the action plan that was developed. Please do not edit. Type it exactly as members developed. You will need a verison in the local language and you will work with the secretary to translate to English for analysis/tracking.
For **Procedure #2 (Separate Dialogues)** PLEASE USE THE SOCIAL CONTRACT TEMPLATE Because the actions in the social contract template are final, not "suggested."
For **Procedure #3 (Interface Only)** and **Procedure #4 (Full Programme)**, the action plan should be placed in the file. Only the social contract is submitted with the report to your manager.

Document Type	**ACTION PLAN**			For Document type enter "Community Action Plan" or "HC Action Plan"
District	**Kabarole**			
Sub-County	**West division**			
Health Centre	**Kagote health center III**			
Procedure #	**4**			
Facilitator name(s)	**Makasi and Hilary**			
Facilitator Organisation	**KRC**			
Action Plan By (tick one)				
	COMMUNITY			
Date developed	1	12	2014	
	Day	Month	Year	

#	Issue	Reasons for Issue	***Suggested*** **Action**	***Suggested*** **Person Responsible**	***Suggested*** **Completion Date**
1	Information on drug availability	Community is not informed whenever drugs are available at the health center	Writing to the in charge to always inform the community through the notice board whenever drugs are available at the health center	VHT Nkayezu	30/12/2015
2	Staff putting on uniform	Not following up staff who come on duty without putting on uniform by the incharge	Writing to the in charge to speak to the staff to always put on Uniform while on duty	VHT Coordinator Richard mwagushia	30/1/2015
3	Lab equipments	Some of the Lab equipment's are not at the health center like the one for Typhoid	Writing to the in charge to ensure that all the Lab equipment's are available at the health center	VHT Nkayezu	30/1/2015
4	The behavior for the staff	The staff do not mind about the patients at the health center	Writing to the in charge to speak to the staff about there behaviors in handling the patients at the health center and come up with the solution	Kairu christopher (Elder)	30/1/2015
5	UHMC roles	Community does not know the role of HUMC at the health center	Writing to the incharge to inform the community on the role of HUMC through the village notice boards	incharge Mugisa Brian	30/1/2015
6	Coming late by patients for treatment at health center	Community members have a thinking that there are always no drugs at the health center	HVTs should sensitize the community members to always come early for treatment at the health center before the sickness worsens	VHT Coordinator Richard mwagushia	30/12/2015

Figure 11: Sample Community Action Plan from Kabarole-Kagote HC3

Instructions
Please record below the action plan that was developed. Please do not edit. Type it exactly as members developed. You will need a verison in the local language and you will work with the secretary to translate to English for analysis/tracking.
For **Procedure #2 (Separate Dialogues)** PLEASE USE THE SOCIAL CONTRACT TEMPLATE Because the actions in the social contract template are final, not "suggested."
For **Procedure #3 (Interface Only)** and **Procedure #4 (Full Programme)**, the action plan should be placed in the file. Only the social contract is submitted with the report to your manager.

Document Type	**ACTION PLAN**			For Document type enter "Community Action Plan" or "HC Action Plan"
District	**KABAROLE**			
Sub-County	**WEST DIVISION**			
Health Centre	**KAGOTE HEALTH CENTER 3**			
Procedure #	**4**			
Facilitator name(s)	**MAKASI & HILARY**			
Facilitator Organisation	**KRC**			
Action Plan By (tick one)				
		HEALTH CENTRE		
Date developed	24	11	2014	
	Day	Month	Year	

#	Issue	Reasons for Issue	***Suggested*** Action	***Suggested*** Person Responsible	***Suggested*** Completion Date
1	Bringing mothers to deliver at the health center	Service was not being offered at the health center	Using VHTs to give information that the center now conducts deliveries, Carry out health education at the HC	Health assistant Muhumuza Michael	end of Feb 2015
2	Information on drug availabiltiy	Information gap	Displaying dlivery of drug on public notice boards	Medical records officer Henry & Beatrace	end of Jan 2015
3	Information on health rights & responsibility	Information gap	Display alist of health rights and responsibility on the public notice boards	In cahrge Mugisa	end of Feb 2015
4	Involvement of HUMC in HC activity	Not community members & have over stayed in office	Write to the office of the town clack about formation of HUMC at the HC	In charge	end of march 2015
5	Community dialogue	CRC not disemineted to the community	Disemineting the CRC	Hilary and Makasi	end of Dec 2014

Figure 12: Sample HC Action Plan from Kabarole-Kagote HC3

Instructions

Please record below the social contract that was developed in the interface. Please do not edit - type it exactly as members have developed. You will need a verison in the local language and you will work with the secretary to translate to English for analysis/tracking.

For **procedure #2 (Separate Dialogues)** please use this format to develop action plans.
For **Procedure #3 (Interface Only)** and **Procedure #4 (Full Programme)**, attach a copy of the social contract to the Interface report and submit to Manager within five (5) working days of the Interface.

Document Type	**SOCIAL CONTRACT**		
District	KABAROLE		
Sub-county	WEST DIVISION		
Health Centre	KAGOTE HEALTH CENTER 3		
Procedure #	4		
Facilitator name(s)	MAKASI K. EDWARD & RUYOOKA HILARY		
Facilitator Organisation	KRC		
Date developed	8	12	2014
	Day	Month	Year

For "Document" you will enter "social contract" if this is used in Interface. Enter "Commuity Action Plan" or "HC Action Plan" for Procedure #2 (Separate Dialogues)

#	Issue	Action	Person Responsible	Expected Completion Date	Evidence of Progress on Action	Person Responsible for Monitoring Progress
1	Putting on Unifrom by staff	Incharge to inform the staff in a meeting to always put on uniform while on duty at the health center	In-charge Mugusa Bra	End of March 2015	Finding staff in uniform while on duty and the minuts for the meeting	O/c Kagote station Muhindo
2	Staff conduct	Incharge to hold a meeting with staff to discuss about there conduct towards the patients while on duty at the health center and come up with a solution	In-charge Mugusa Bra	End of Feb 2015	Minutes for the meeting and the change in the conduct for the staff towards the patients	Kabasiguzi Beatrace (elder)
3	HUMC Functionality	Electing the new HUMC	Town clark and the incharge	End of June 2015	Council Minutes and the new HUMC to be in place	Tuhaise Aisha (elder)
4	Information on Health rights and responsibility	Providing information on health rights and responsibilites by VHTs in the villages and health center staff at OPD and hang it on the notice board	Health Assistant Muhumuza michael	End of Feb 2015	VHTs Reports	VHT kayezu
5		Health center staff provide information on health right at and responsibility at OPD and hang it on the notice board	Health Assistant Muhumuza michael	End of Feb 2015	Information on health rights and responsibility hanged on the notie board	VHT kayezu
6	Coming late for teatment at the H/C by the patients	Providing information on the need to always come early for treatment by VHTs in the villages and	VHTs	End of Dec 2015	VHTs Reports	Health Assistant

Figure 13: Sample Social Contract from Kabarole-Kagote HC3

References

Andrabi, Tahir, Jishnu Das and Asim Ijaz Khwaja. 2017. "Report cards: The impact of providing school and child test scores on educational markets." *American Economic Review* 107(6):1535–1563.

Arkedis, Jean, Jessica Creighton, Dixit Akshay, Archon Fung, Stephen Kosack, Dan Levy and Courtney Tolmie. 2021. "Can transparency and accountability programs improve health? Experimental evidence from Indonesia and Tanzania." *World Development* 142:1 –14.

Banerjee, Abhijit, Rema Hanna, Jordan Kyle, Benjamin A. Olken and Sudarno Sumarto. 2018. "Tangible information and citizen empowerment: Identification cards and food subsidy programs in Indonesia." *Journal of Political Economy* 126(2):451–491.

Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, Rachel Glennerster and Stuti Khemani. 2010. "Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India." *American Economic Journal: Economic Policy* 2(1):1–30.

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

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the false discovery rate: A practical and powerful approach to multiple testing." *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 289–300.

Björkman, Martina and Jakob Svensson. 2009. "Power to the people: Evidence from a randomized field experiment of a community-based monitoring project in Uganda." *Quarterly Journal of Economics* 124(2):735–769.

Björkman Nyqvist, Martina, Andrea Guariso, Jakob Svensson and David Yanagizawa-Drott. 2019. "Reducing child mortality in the last mile: Experimental evidence on community health promoters in Uganda." *American Economic Journal: Applied Economics* 11(3):155–192.

Christensen, Darin, Oeindrila Dube, Johannes Haushofer, Bilal Siddiqi and Maarten Voors. 2021. "Building resilient health systems: Experimental evidence from Sierra Leone and the 2014 Ebola outbreak." *Quarterly Journal of Economics* 136:1145–1198.

Clayton, Amanda, Daniel De Kadt and Nicolas K Dumas. 2019. "Daughters do not affect political beliefs in a new democracy." Working paper.

Cox, D. R. 1972. "Regression Models and Life-Tables." *Journal of the Royal Statistical Society* 34(2):187–220.

Croft, Trevor N., Aileen M. J. Marshall and Courtney K. Allen. 2018. *Guide to DHS Statistics*. Maryland: ICF.

Fiala, Nathan and Patrick Premand. 2018. "Social accountability and service delivery: Evidence from a large-scale experiment in Uganda." Working Paper.

Freire, Danilo, Manoel Galdino and Umberto Miznozzetti. 2020. "Bottom-up accountability and public service provision: Evidence from a field experiment in Braziil." *Research and Politics* pp. 1–8.

Hoenig, John M and Dennis M Heisey. 2001. "The abuse of power: the pervasive fallacy of power calculations for data analysis." *The American Statistician* 55(1):19–24.

Keefer, Philip and Stuti Khemani. 2014. "Mass media and public education: The effects of access to community radio in Benin." *Journal of Development Economics* 109:57–72.

Kling, Jeffrey R., Jeffrey B. Liebman and Lawrence F. Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica* 75(1):83–119.

Lakens, Daniël. 2017. "Equivalence Tests: A Practical Primer for t-Tests, Correlations, and Meta-Analyses." *Social Psychological and Personality Science* 8(4):355–362.

Lieberman, Evan S., Daniel N. Posner and Lily L. Tsai. 2014. "Does information lead to more active citizenship? Evidence from an education intervention in rural Kenya." *World Development* 60:69–83.

Lin, Winston, Donald Green and Alexander Coppock. 2016. Standard Operating Procedures for Don Green's Lab at Columbia. Technical report.
URL: *https://github.com/acoppock/Green-Lab-SOP*

Mohanan, Manoj, Vikram S. Rajan, Kendal M. Swanson and Harsha Thirumurthy. 2020. "Information and facilitation interventions for accountability in health and nutrition: Evidence from a randomized trial in India." ERID working paper no. 295.

Olken, Benjamin. 2007. "Monitoring corruption: Evidence from a field experiment in Indonesia." *Journal of Political Economy* 115(2):200–249.

Pandey, Priyanka, Sangeeta Goyal and Venkatesh Sundararaman. 2009. "Community participation in public schools: Impact of information campaigns in three Indian states." *Education Economics* 17(3):355–75.

Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, Arminda Alisjahbana and Rima Prama Artha. 2014. "Improving educational quality through enhancing community participation: Results from a randomized field experiment in Indonesia." *American Economic Journal: Applied Economics* 6(2):105–126.

Ravallion, Martin, Dominique van de Walle, Puja Dutta and Rinku Murgai. 2015. "Empowering poor people through public information? Lessons from a movie in rural India." *Journal of Public Economics* 132:13–22.

Reinikka, Ritva and Jakob Svensson. 2005. "Fighting corruptoin to improve schooling: Evidence from a newspaper campaign in Uganda." *Journal of the European Economic Association* 3(2-3):259–267.

Rosner, Bernard. 2015. *Fundamentals of Biostatistics*. Nelson Education.

Rowland, Donald T. 2003. *Demographic methods and concepts*. New York: Oxford University Press.

Schuirmann, Donald J. 1987. "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability." *Journal of Pharmacokinetics and Biopharmaceutics* 15(6):657–680.

UNDG. 2003. *Indicators for monitoring the Millennium Development Goals: definitions, rationale, concepts and sources*. Vol. 95 United Nations Publications.