Online Appendix for: The impact of community-based monitoring on public service delivery: A randomized control trial in Uganda

A.1 Balance in the original design

In Table A.1, we test for balance between the treatment groups at baseline following the original design of the experiment. Baseline data on 12,545 households was collected between June and September 2015. Sample averages are reported in the first column (with standard errors in brackets below). For example, we see that the average household consists of about 6 household members and that only about 11 percent of households report that they had access to agricultural extension. In the second column, we report differences between baseline characteristics of households that were, according to the original design, planned to receive a sub-county level combined information and deliberation baraza, and those that would not be exposed to any baraza. We find that distance to the nearest road is slightly—though statistically significantly—higher in areas that were planned to receive a sub-county level baraza. When comparing households that were planned to be exposed to a sub-county level information baraza to households that would not receive a sub-county level information baraza (column 3), we see that that households are slightly larger in the former group, and the difference is significant. The average household has two to three children attending a public school. We also find a slight pre-treatment imbalance on this outcome for the information treatment. In the fourth column of Table A.1, we report differences between households that would be exposed to a sub-county deliberation baraza and households that would not. For this treatment, we cannot reject balance on any of the variables. In the last column, we report differences in outcomes between households that would be exposed to a district level baraza and households that would not be exposed to any baraza. Also here, we do not find any significant difference. Overall, out of 40 comparisons, we find that three differences are significant, which is what one would expect to find due to chance alone. As such, we conclude that the randomization of the original design was successful.

Table A.1: Balance between treated and control sub-counties (original design)

	mean	sc baraza	information	deliberation	dist baraza
Household size	6.324	0.021	0.304*	-0.003	0.267
	(2.825)	(0.142)	(0.133)	(0.125)	(0.226)
Age of household head (years)	46.501	0.736	0.464	0.725	1.827
	(14.615)	(0.681)	(0.594)	(0.714)	(0.943)
Household head is female (1=yes)	0.191	0.012	-0.014	0.004	-0.006
	(0.393)	(0.014)	(0.013)	(0.015)	(0.012)
Household head finished primary education (1=yes)	0.213	-0.007	-0.020	-0.003	-0.038
	(0.410)	(0.017)	(0.020)	(0.020)	(0.028)
Thatched grass roof (1=yes)	0.298	-0.001	0.009	-0.032	0.020
	(0.457)	(0.026)	(0.025)	(0.023)	(0.034)
Traditional mud wall (1=yes)	0.424	0.021	-0.025	0.038	-0.018
	(0.494)	(0.043)	(0.040)	(0.039)	(0.107)
Distance to nearest all weather road (km)	0.906	0.167*	0.106	0.147	-0.043
	(0.915)	(0.106)	(0.095)	(0.092)	(0.091)
Access to extension (1=yes)	0.108	0.002	0.004	0.007	0.007
	(0.310)	(0.014)	(0.012)	(0.014)	(0.019)
Village health team in village (1=yes)	0.854	0.000	0.006	0.025	0.079
	(0.353)	(0.031)	(0.026)	(0.026)	(0.034)
Number of children in public schools	2.478	0.044	0.165+	0.038	0.165
	(2.074)	(0.095)	(0.091)	(0.089)	(0.146)
Number of observations	12545	5193	10241	10241	4850

Note: Column 1 reports sample means (and standard deviations below); Column 2 reports differences between households that were planned to receive a sub-county information+deliberation baraza and those that would not receive any baraza (and standard errors below); Column 3 reports differences between households that would receive a sub-county information baraza and those that would not receive one (and standard errors below); Column 4 reports differences between households that would receive a sub-county deliberation baraza and those that would not receive one (and standard errors below); Column 5 reports differences between households that were planned to receive a district information+deliberation baraza and those that would not receive any baraza (and standard errors below); **, * and + denote significance at the 1, 5 and 10 percent levels. Reported standard errors are clustered at the level of randomization. Distance to nearest all weather road is trimmed at 5 percent and transformed using the inverse hyperbolic sine transformation.

A.2 Minimum detectable effect sizes

The partial roll-out means a reduction in statistical power. In this section, we use baseline data to simulate an updated set of minimum detectable effects (MDEs) for the outcomes that were pre-registered to be used to judge effectiveness of the intervention. Instead of fixing MDEs and simulating power for different sample sizes as is usually done in power calculations, we fix sample size to what we ended up with after partial roll-out and simulate power for different MDEs. We use a standard significance level of 0.05 (double sided). MDEs are estimated using a simple analysis of covariance (ANCOVA) model that controls for the outcome at baseline and accounts for the clustered nature of the intervention.[1] The simulated MDEs can be a useful guide in deciding whether the intervention really had no effect, or an effect that may have been too small to detect given the (realized) research design.[2]

Figure A.1 plots MDEs against power for the first pre-registered outcome variable that will be used to assess the impact of barazas on public service delivery in the agricultural sector (extension at home, measured as the percentage of households in our sample who report that they were visited by an expert in the previous year). On average, about 11 percent of households in our sample reported that they were visited by an extension officer in the year before baseline data collection. The gray solid line shows the power curve associated with the deliberation treatment, comparing 18 sub-counties that received the deliberation only treatment to 40 pure control sub-counties, which corresponds to a sample of about 2,900 households. The light blue dashed line shows power for different MDEs for the information component of the baraza intervention. Here, we compare 29 sub-counties that received an information focused baraza to 40 control sub-counties, corresponding to about 3,450 households. We also investigate power for the comparison between pure control barazas and sub-county level barazas (black dotted line). Here we compare about 1,000 households in the 20 sub-counties that received the combined information and deliberation treatment to about 2,000 households that did not receive any baraza (in the 40 pure control sub-counties). Finally, the dark blue dashed line plots power for the comparison between about 2,000 households in 40 pure control sub-counties and about 2,000 households that were exposed to a district level baraza (sampled from 40 sub-counties in these districts).

While power curves are generally not too different, we seem to have most power for testing the information treatment. We see that the power curve hits the 80 percent threshold for the first time at an MDE of about 3.7 percentage points. The deliberation experiment is similarly powered, and at 80 percent we can expect to identify effects of 3.75 percentage points or more. Due to the

[1] Note that the level of clustering depends on the comparison. When we look at sub-county level barazas, clustering is at the sub-county level. For the comparison between district level barazas and control, clustering is at the district level.

[2] These new simulated MDEs should not be confused with ex-post power calculations, the dangers of which are well-known to us (Hoenig and Heisey, 2001): The simulations were done before endline data was collected. The only additional information we had at that stage was the realized sample size.

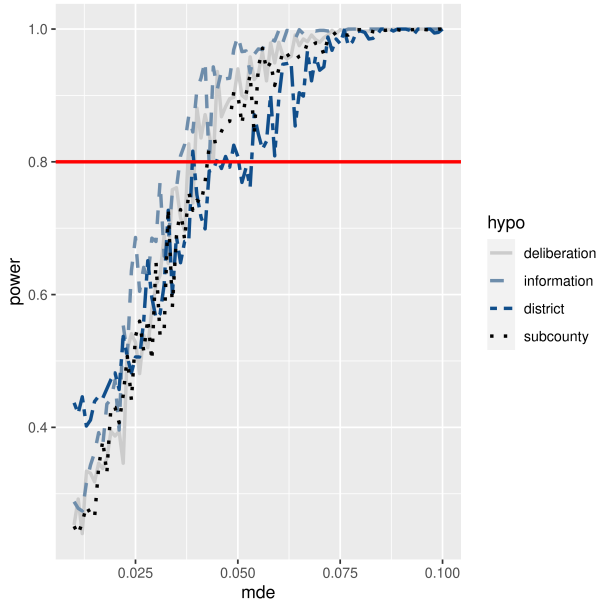

Figure A.1: Power curves for access to extension

smaller sample size, it is harder to compare pure control sub-counties to sub-county barazas that received the interacted treatment. Still, we can comfortably identify an increase of 4 percentage points. Despite a larger sample size available to test the impact of district level barazas, we find power to be comparable to that of the sub-county level baraza. This is due to the clustering at a much more aggregated level. In the online repository, we show results for similar analyses for all pre-registered variables that we will use to judge impact of the baraza intervention.

A.3 Attrition

As more than four years passed between baseline and endline data collection, some level of attrition was to be expected. Overall, attrition amounted to 8.6 percent. Reports from the field indicated that this attrition was due to the fact that some villages that were close to Lake Victoria had been evacuated due to rising levels of the lake. Two villages in areas populated with land insecure settlers were dropped due to security issues. Finally, attrition was generally higher in urban areas, as households that were residing in rented properties returned to their villages. Table A.2 shows that attrition is independent of the various interventions of the study.

Table A.2: Attrition

	mean	sc baraza	information	deliberation	dist baraza
Attrition	0.086	0.010	0.0139	0.003	-0.004
	(0.280)	(0.020)	(0.017)	(0.019)	(0.014)
Number of obs.	7340	2996	5341	5341	3996

Note: Column 1 reports the mean attrition rate (and standard deviations below); Column 2 reports the correlation between attrition and sub-county baraza treatment (and standard errors below); Column 3 reports the correlation between attrition and information baraza treatment; Column 4 reports the correlation between attrition and deliberation baraza treatment; Column 5 reports the correlation between attrition and district baraza treatment. Reported standard errors are clustered at the level of randomization.

A.4 Detailed analysis

A.4.1 Agriculture

We first zoom in on the outcomes that are used to assess the effectiveness of barazas in changing service delivery and associated outcomes in agriculture. Results are reported in Table A.3. We start by looking if the baraza program affected the use of modern inputs in agriculture. A first outcome looks at whether the household used inorganic fertilizers (DAP, Urea, NPK, Foliar, TSP, SSP, MOP) during the last 12 months, an important yield increasing technology. The first column reports sample averages, with standard deviation in brackets below. We find that about 23 percent of households in the sample gave an affirmative answer to this question. In the second column, we report differences in outcomes between households that received a sub-county level baraza and households that did not receive any baraza (β as estimated using the Ordinary Least Squares model in Equation (1) in the paper). We see that the proportion of households that reported using inorganic fertilizer is 1.5 percentage points lower among the sub-group of households that were exposed to a sub-county level baraza that consists of both the information and the deliberation component than among households that did not receive any baraza. However, this difference is not statistically significant. In the third column, we report differences between outcomes of households that live in areas where an information baraza was organized and outcomes of households that live in areas that were not exposed to a baraza. We see that adoption of inorganic fertilizer was 3.4 percentage points higher among households that were exposed to an information focused baraza. However, also here, the difference is not significant. In the fourth column, we report differences between outcomes of households that live in areas where a deliberation baraza was organized and outcomes of households that live in areas that were not exposed to a baraza. We also do not find differences in terms of inorganic fertilizer use. Finally, in the fifth column, we compare households that

were exposed to a district level baraza to pure control households. Again, no impact of the district level baraza is found on this outcome.

The second outcome is related to the use of improved seed, a second important yield improving technology promoted by the government. This input seemed to be used more widely than inorganic fertilizer: 36 percent of households reported that they had been using improved seed during the preceding year. This percentage is 4.3 percentage points higher among households that reside in areas where a sub-county level baraza took place as opposed to in areas where no baraza was conducted, but the difference is not significant. The coefficient estimate is much smaller than the 80 percent MDE reported in Figure 2 in the online repository, so it is unclear if this null result really signifies the absence of an effect, or if it is due to small sample size. We find negative point estimates for the relative effects of both the information and the deliberation component, but effects are imprecisely estimated. Finally, we find adoption of improved seed was lower in areas where a district level baraza was conducted, but the difference with areas that did not receive a baraza is not significant. Adoption of improved seed and inorganic fertilizer was included in the index to judge impact on the agricultural sector in Figure 3 of the paper.

Next, we find that about 12 percent of households reported that they received improved seed from the government extension system (through an extension agent, from the National Agricultural Advisory Services (NAADS) or through Operation Wealth Creation (OWC) that replaced NAADS). We find that this is 5.1 percentage points higher in areas where a sub-county level baraza took place. We further see that this is especially due to the deliberation component of sub-county level barazas. In areas where such a baraza was held, the share is 5.6 percentage points higher than in areas where citizens were not given the opportunity to engage with their leaders. We find no such effects from district level barazas. Direct comparison of district and sub-county level barazas indicate that sub-county level barazas are significantly more effective in increasing the likelihood that households reported to have received these inputs from government.

We then check if household changed with respect to the use of agrochemicals. This includes the use of pesticides, herbicides, fungicides and acaricides during the 12 months prior to endline data collection. Overall, almost half of all households in the sample reported using some form of agrochemicals. We do not find evidence that the baraza intervention affected the use of these inputs. Finally, we consider the use of modern inputs and methods in livestock rearing over the last 12 months. This includes improved animal breeds, the use of modern feeds, drugs, and artificial insemination. 22 percent of households report that they used such inputs, and this proportion is similar across different experimental groups.

We then turn to agricultural advisory services. We first investigate if barazas affected access to extension at home. We estimate the percentage of households in our sample who reported that they were visited by an expert (for example, crop or livestock extension agent, or community-based facilitator or another experienced farmer) at home in the last 12 months. We find that access to

extension, although higher than at baseline, is still low, with only about 18 percent of households reporting that they received such a visit. We find that this percentage is significantly higher among households that were exposed to a sub-county level baraza. The effect is large, amounting to an increase of 5.6 percentage points. The effect seems to come from a combination of the information and deliberation components; the components themselves do not seem to affect the outcome enough to render if significant. We also find that this effect is absent among households that live in sub-counties that received the district level baraza. A direct comparison of extension at home between households that were exposed to a district level baraza and households that were exposed to a sub-county level baraza confirms that sub-county level barazas were significantly more effective.

Home visits by extension officers are not the only way in which households have access to information. Extension offices, demonstration sites and model farmers are also an integral part of the Ugandan agricultural advisory system. Especially after the establishment of NAADS, such a demand-led service component that can be consulted by farmers when the need arises became more important than the more supply-driven component of training and visit. We thus also inquire if anyone in the household visited an extension office, demonstration site, or model farmer in the year preceding endline data collection. We find that about 29 percent of households in our sample reported access to extension in this way. While the results are in line with extension visits at home, differences are not significant. Access to extension, both at home or though extension offices and demonstration sites, was also included as a key outcome to assess impact using in the agricultural sector index.

We find that three quarters of households in our sample mention that there were agricultural enterprises, improved technologies or inputs they would like to adopt, indicating significant demand for advisory services. We also find that, according to citizens, service providers and policy makers were not always aware of the needs of farmers. The table shows that only 26 percent of households believed that officials were aware of which services farmers need. While we do not see that this percentage differs between treatment and control for sub-county level barazas, we do see that a district level baraza reduced this percentage. This may indicate that a district level baraza made the mismatch between what farmers need and what officials think farmers need more salient.

Related to the previous outcome, we asked how decisions related to what topics to cover in agricultural extension are made. We define this outcome in a negative way, that is, the indicator is true if decisions are made without consultation. We see that about 32 percent of households indicated that no consultation happens, and the content of extension advisory services is decided upon by experts at the central level. We do not find that the baraza intervention increased participation in extension service planning.

About 40 percent of households reported the presence of farmer groups or cooperatives in their village. In the agricultural sector in Uganda, such groups are very important. They are actively promoted by the government. In fact, to be able to receive inputs from the government extension system, farmers are

strongly encouraged to form such groups. We find that sub-county level barazas increased the likelihood that farmer cooperatives or groups were formed in the villages in Uganda. It seems that the deliberative component is the main driver behind this result. We also find that this effect was specific to interventions at the sub-county level.

We further find that a higher share of farmer groups in areas that received a sub-county level baraza received support from government. NAADS, and more recently OWC, are responsible for the distribution of subsidized agricultural inputs. The Office of the Prime Minister (OPM) reports from the baraza meetings indicate that there are often many complaints related to the absence of NAADS/OWC in sub-counties and the non-transparent way in which decisions are made on who gets to benefit from NAADS/OWC and who does not get inputs. We see that 17 percent of households reported that NAADS/OWC was present in their village. This increased by 7.0 percentage points in areas where a combined information and deliberation sub-county level baraza was held. While the opportunity of households to engage with officials seems to be the main driver behind this result, a deliberation only baraza has no significant impact.

The final two questions focus more on marketing. Connecting farmers to markets is also an important strategy outlined in the Agriculture Sector Strategic Plan (ASSP). The first outcome relates to the likelihood that farmers were supported by government through a village procurement committee. In the sample, about 7 percent of households report that they were assisted by government. A second questions is similar but looked at the role of cooperatives in marketing support. We generally find no effects of the baraza intervention, except perhaps for an increase of almost 3.7 percentage points in the likelihood that cooperatives assisted with marketing in areas that received a sub-county baraza. However, simulated MDEs suggest these results may suffer from a lack of statistical power. Both outcomes were also included in the index to assess agricultural sector impact.

A.4.2 Infrastructure

A second important area in which we expect to see an impact of the baraza program is infrastructure. We primarily focus on drinking water infrastructure. Results are in Table A.4, which is organized similarly to Table A.3.

A first outcome we consider is whether the household used an unprotected water source during the dry season. This is measured as the share of households that reported that the main source of drinking water during the dry season was surface water, an unprotected dug well, or an unprotected spring. We find that about 16 percent of households in the sample reported that they were using an unprotected water source. The baraza intervention did not seem to affect this proportion. This outcome is included in the infrastructure index shown in Figure 3 of the article.

A second outcome we look at (and is also included in the index) is the distance to the primary water source during the dry season. This was measured in

Table A.3: Treatment-control differences (ANCOVA) - Agriculture

	mean	sc baraza	information	deliberation	dist baraza
Used inorganic fertilizer (yes/no)†	0.229	-0.015	0.034	0.001	-0.013
	(0.420)	(0.033)	(0.035)	(0.049)	(0.031)
Used improved seed (yes/no)†	0.364	0.043	-0.030	-0.037	-0.043
	(0.481)	(0.033)	(0.038)	(0.038)	(0.034)
Obtained seed from government (yes/no)	0.121	0.051^{+}	0.004	0.056^{+}	-0.005
	(0.326)	(0.024)	(0.025)	(0.043)	(0.015)
Used agro-chemical inputs (yes/no)	0.469	-0.028	-0.007	-0.005	-0.010
	(0.499)	(0.050)	(0.035)	(0.046)	(0.043)
Used improved livestock methods (yes/no)	0.221	0.029	0.021	0.030	-0.014
	(0.415)	(0.031)	(0.028)	(0.034)	(0.026)
Extension visit at home (yes/no)†	0.178	0.056^{+}	0.037	0.036	-0.027
	(0.383)	(0.018)	(0.030)	(0.048)	(0.014)
Visited extension office, demo site or model farmer (yes/no)†	0.285	0.040	0.036	0.045	-0.013
	(0.452)	(0.028)	(0.035)	(0.044)	(0.023)
Needs extension (yes/no)	0.769	-0.015	-0.043^{+}	-0.029	-0.038
	(0.421)	(0.024)	(0.025)	(0.026)	(0.017)
Extension agents are aware of this need (yes/no)	0.264	-0.006	0.017	-0.001	-0.075*
	(0.441)	(0.024)	(0.027)	(0.035)	(0.023)
Extension agents decided alone (yes/no)	0.316	0.034	-0.041	-0.031	-0.056
	(0.465)	(0.033)	(0.027)	(0.034)	(0.033)
Farmer associations/groups present in village (yes/no)	0.403	0.060^{+}	0.040	0.087*	-0.032
	(0.491)	(0.030)	(0.038)	(0.041)	(0.027)
NAADS/OWC in village (yes/no)†	0.173	0.070*	-0.015	0.053	-0.037
	(0.378)	(0.028)	(0.030)	(0.040)	(0.022)
Support in marketing from marketing committee (yes/no)†	0.069	0.018	-0.013	0.016	-0.014
	(0.254)	(0.022)	(0.017)	(0.017)	(0.012)
Support in marketing from cooperative (yes/no)†	0.062	0.037^{+}	-0.021	-0.001	-0.009
	(0.241)	(0.024)	(0.016)	(0.021)	(0.014)
Number of observations	6704	2738	4858	4858	3687

Note: Column 1 reports sample means (and standard deviations below); Column 2 reports effects of sub-county barazas (and standard errors below); Column 3 reports effects of information barazas (and standard errors below); Column 4 reports effects of deliberation barazas (and standard errors below); Column 5 reports effects of district barazas (and standard errors below). **, * and + denote significance at the 1, 5 and 10 percent levels using randomization inference. Reported standard errors are clustered at the level of randomization. † indicates that the outcome was included in the index for the confirmatory analysis.

km but trimmed and transformed using the inverse hyperbolic sine transformation. We find that, on average, households had to walk about 0.91 km. While this distance seemed to reduce in all comparisons, and especially for barazas held at the district level where we find a reduction of approximately 9 percent, it is never significantly different from zero. Simulated MDEs in the online repository] indicate that, except for the sub-county level baraza, these null effects are probably not due to limited power.

The third outcome, also part of the index, is the time that one must wait at the water source, measured in minutes. This continuous variable was also trimmed and transformed. We find that households must wait on average about 36 minutes. We find that waiting time in areas that were exposed to the sub-county level baraza intervention reduced by about 25 percent. The deliberation component seems to be mostly responsible for this reduction.

The fourth outcome variable assesses changes in the presence of a water user committee in the village. Overall, about 60 percent of households reported that such a committee was present in their village. We do not find that this share varies between the different experimental groups, but comparison of estimated coefficients to simulated MDEs suggests power is an issue. Consistent with the finding that deliberation has been found to increase subsequent citizen involvement in community affairs (Björkman Nyqvist, de Walque and Svensson, 2017), we see that the deliberation component led to an increase in the likelihood that a household member was a member of a water user committee. We do not find that the baraza intervention increased the likelihood that these committees hold public meetings. Only the first of these three outcomes was pre-registered and included in the index.

Households were also asked if they were satisfied with the quality of the water that was available at the source during the dry season. About 62 percent responded that they were satisfied or very satisfied with the drinking water. We do not find that households that were exposed to the baraza intervention were more or less likely to report that they were (very) satisfied with the quality of drinking water during the dry season. Half of the households reported that they treated drinking water before drinking it, either by boiling it or treating it with chlorine. The likelihood that households treated water reduced somewhat for the information treatment. This may be due to the belief held in many households that water from a safe source does not need additional treatment.

We included one question related to road infrastructure. We asked how far the household is located from the nearest all weather road. We find that in the full sample a household lived on average 26 km from a road. We do not find that the baraza program reduced the distance to the nearest all weather road.

A.4.3 Health

We now look at outcomes in the health sphere (Table A.5). One problem with public health related outcomes is that some outcomes related to quality-of-service delivery in health centers will only be available for households that have visited government health facilities, reducing sample size too much to maintain

Table A.4: Treatment-control differences (ANCOVA) - Infrastructure

	mean	sc baraza	information	deliberation	dist baraza
Household uses unprotected water source during dry season (yes/no)†	0.159	0.031	0.005	0.010	-0.023
	(0.366)	(0.042)	(0.036)	(0.037)	(0.046)
Distance to water source (km)†	0.748	-0.026	-0.040	-0.049	-0.091
	(0.576)	(0.046)	(0.041)	(0.061)	(0.039)
Average waiting time at source (min)†	3.198	-0.286*	-0.006	-0.287+	-0.032
	(1.638)	(0.152)	(0.117)	(0.193)	(0.160)
Water user committee present in village (yes/no)†	0.598	-0.021	0.033	0.032	-0.009
	(0.490)	(0.046)	(0.037)	(0.040)	(0.047)
Satisfied with water qualiy (yes/no)	0.624	0.031	-0.009	-0.062	0.020
	(0.484)	(0.052)	(0.044)	(0.044)	(0.038)
Boil/treat water before drinking (yes/no)	0.500	-0.025	-0.087*	-0.020	-0.050
	(0.500)	(0.045)	(0.037)	(0.046)	(0.056)
Member of commitee (yes/no)	0.163	0.022	0.001	0.040+	0.002
	(0.370)	(0.021)	(0.017)	(0.025)	(0.019)
Committee holds public meetings (yes/no)	0.474	-0.005	0.043	0.060	0.010
	(0.499)	(0.044)	(0.036)	(0.042)	(0.041)
Distance to nearest all weather road (km)†	2.849	0.388	-0.129	-0.286	0.591
	(1.788)	(0.314)	(0.306)	(0.313)	(0.405)
Number of observations	6704	2738	4858	4858	3687

Note: Column 1 reports sample means (and standard deviations below); Column 2 reports effects of sub-county barazas (and standard errors below); Column 3 reports effects of information barazas (and standard errors below); Column 4 reports effects of deliberation barazas (and standard errors below); Column 5 reports effects of district barazas (and standard errors below). **, * and + denote significance at the 1, 5 and 10 percent levels using randomization inference. Reported standard errors are clustered at the level of randomization. † indicates that the outcome was included in the index for the confirmatory analysis. Distance to water source, average waiting time at source, and distance to nearest all weather road are trimmed at 5 percent and transformed using the inverse hyperbolic sine transformation.

acceptable power. Furthermore, the intervention may also affect the likelihood that households visit a government health facility, potentially introducing selection bias.

The first two outcomes we consider attempt to assess changes in access or use of public health facilities. A first indicator measures demand for public health services for illness. In particular, we construct an indicator that is true if the household head responded that treatment would be sought in a health center 2, 3, 4 or in a regional referral hospital if a member of his/her household had fever. We find that 69 percent of households responded that they would seek treatment in a government health facility. This proportion is independent of the treatment groups.

A similar indicator attempts to assess demand for public health services for maternal health care, and asks if treatment would be sought in a health center 2, 3, 4 or in a regional referral hospital if a member of the household was to give birth. This percentage was even higher: about 81 percent would go to a government health facility to give birth. Again, this proportion was not affected by the baraza program. Both outcomes are included in the health index.

Next, we asked if a Village Health Team (VHT) was present in the village. VHTs are very important in front-line health care provision in Uganda. They also have prominent roles in government health interventions, such as immunization campaigns or the distribution of bed nets. We find that overall, about 88 percent of households reported that a VHT was present in their village. The presence of a VHT was not impacted by the baraza intervention.

As the baraza tries to increase citizen engagement, we also checked if households that were exposed to a baraza are more likely to participate in VHTs. We thus asked if any member of the household was a member of a VHT. We see that in about 11 percent of our sample, at least one household member was part of a VHT. The baraza intervention did not increase the likelihood that individuals participated as VHT members. Furthermore, the baraza intervention attempts to encourage sharing of information. As such, we expect that being exposed to a baraza may encourage VHTs to organize more public meetings. We find that overall, 43 percent of households stated that VHTs organized a public meeting in the preceding year. We find that this proportion was significantly higher in areas that were exposed to a sub-county level baraza.

We also consider distance to the nearest government health facility, measured in km (trimmed and transformed using the inverse hyperbolic sine transformation). Overall, average distance to the nearest government health facility was almost 50 km. We do not find that barazas reduced this distance.

We then turn to health outcomes. We started by asking if any member of the household fell sick during the year prior to endline data collection. This was the case in two thirds of the households in our sample. The intervention did not reduce morbidity in our sample. We then asked for each sick person in the household to record how many days he or she was ill and use this to calculate the total number of sick days at the household level in the last year. The average household recorded almost 50 sick days according to this definition. We also do not find that the intervention affected the (trimmed and transformed)

number of sick days. We also look at the number of days household members were unable to go to school or to work, which provides an indication of the severity of illness. Calculated similarly to the previous outcome, we find that in the average household about 35 school- or workdays were missed due to illness. Again, there was no significant reduction in this (trimmed and transformed) number. This last health outcome measure is included in the health index.

Next, we asked how long one had to wait before being attended to (in minutes, trimmed and transformed). We find that the sample mean for this outcome is about 90 minutes. While we see that waiting time reduced for most comparisons, the differences are never significant. However, we only have sufficient power to confidently claim null effects for the components of the sub-county level baraza, not for the combined sub-county level baraza. This outcome was also included in the health index.

The next question that was included in the index was again asked to all households. In particular, we inquired if a traditional health practitioner was consulted in the last year. In one in four households in our sample, this was the case. The baraza intervention did not affect this percentage.

One problem that often crops up in the health sector is absenteeism (Banerjee and Duflo, 2006). To assess this, we asked who examined the patient in the health center. Ideally this should be the doctor or in-charge. If this person was absent, patients were generally examined by nurses or lab technicians. We thus construct an indicator that is one if the household responds that the patient was investigated by the doctor or the in-charge, and zero otherwise. Only in 41 percent of the cases, a qualified person appeared to do the examination. Surprisingly, the deliberation component of the baraza seems to reduce this likelihood somewhat. We also look at the time that the examination takes. The average examination in our sample took about 22 minutes. There is no change related to the intervention.

Health care in Uganda is supposed to be free. However, corruption is widespread and often patients are required to make payments to receive care. We find that about 18 percent of households reported that payment was required the last time they visited a government health facility. There was no impact of the intervention. Related, users often complained about a lack of drugs in government health facilities. We asked if, during the last visit to a government health center, drugs were received (indicating that drugs were available). We also asked if drugs had to be purchased from outside of the hospital (indicating that at least some drugs were missing). While 71 percent of households reported that they received medicines in the health center, almost all of them also mentioned that they had to also buy drugs outside of the hospital. For neither indicator, the intervention seems to make a difference.

We further probe for a subjective assessment of the overall quality of care at the health facility. Most households reported that they were satisfied or very satisfied with services received at the government health facility. This seemed to increase in areas where a sub-county level baraza took place and there is also a sizable difference in outcomes when comparing sub-county level baraza outcomes to district level baraza outcomes. However, none of the differences

are significant.

We considered several other health related outcomes that feature prominently in other studies. One key outcome in Björkman and Svensson (2009) is immunization. However, we already find close to 100 percent immunization rates in our baseline data. Another outcome is child mortality. Child mortality rates at baseline were estimated at 38 per 1000 live births, which was deemed too low to include in the analysis. Raffler, Posner and Parkerson (2020) find similar child mortality rates at baseline and suggest that the fact that they do not find an effect while Björkman and Svensson (2009) do is due to differences in baseline conditions: child mortality at baseline in Björkman and Svensson (2009) was 117 per 1000 live births.

A.4.4 Education

Education outcomes to assess impact of the intervention suffers from a similar problem as the one encountered with health outcomes: not all households in the sample had children in school, and so for many of the outcomes related to education, sample size becomes small. This also affects the indices. Results are presented in Table A.6.

If the quality of public education is poor, households will be less likely to send their children to public schools. A first obvious outcome is thus to simply compare the number of children within the household that attended public schools (either Universal Primary Education or Universal Secondary Education). We find that the average household in our sample had almost two children in government schools, but that enrollment rates were not affected by the baraza intervention, although for all but the information component, we cannot differentiate between null effects and insignificant results due to low statistical power.

Access to public education is also influenced by the distance to a public school. We thus recorded distance to primary or secondary school (or the average if both are reported; trimmed and transformed). We find that on average, households lived about 3 km from a government operated school. Also for this outcome, the baraza program did not make an impact, but estimated coefficients are below MDEs and so insignificance may be due to small sample size.

We also look at school infrastructure. First, we asked households if the primary or secondary school attended by any of their children had a complete boundary fence. In the complete sample, it was reported that only about 35 percent of schools have such a fence. The lack of a fence was a frequent complaint from parents during qualitative work. We also asked if the school had electricity and if there was a water source available in the school. We find that overall, about 34 percent of schools had electricity and about 70 percent had a water source. We find that sub-county level baraza seem to improve school infrastructure. We considered many other infrastructure related outcomes, such as the number of classrooms and availability of functioning toilets for both girls and boys, but baseline data suggested there were generally no issues related to these outcomes.

Table A.5: Treatment-control differences (ANCOVA) - Health

	mean	sc baraza	information	deliberation	dist baraza
Sought treatment for fever in public health facility (yes/no)†	0.691	-0.008	-0.007	0.025	-0.010
	(0.462)	(0.033)	(0.033)	(0.040)	(0.046)
Went to public health facility to give birth (yes/no)†	0.813	-0.029	-0.033	-0.016	-0.070
	(0.390)	(0.034)	(0.029)	(0.035)	(0.043)
VHT present in village (yes/no)†	0.881	0.022	0.005	0.029	-0.019
	(0.323)	(0.031)	(0.025)	(0.025)	(0.027)
Member of the VHT (yes/no)	0.113	0.022	0.003	-0.001	-0.024
	(0.317)	(0.017)	(0.014)	(0.015)	(0.012)
VHT organised public meetings (yes/no)	0.429	0.076*	-0.018	0.058	-0.046
	(0.495)	(0.041)	(0.033)	(0.040)	(0.036)
Distance to nearest govt health facility (km)†	3.875	0.256	-0.162	-0.252	0.445
	(1.377)	(0.219)	(0.233)	(0.263)	(0.342)
HH members were sick (yes/no)	0.658	0.003	0.024	0.037	-0.015
	(0.475)	(0.023)	(0.028)	(0.033)	(0.024)
Number of days sick	2.576	-0.005	-0.040	0.004	-0.064
	(2.189)	(0.091)	(0.149)	(0.166)	(0.105)
Number of days work/school missed due to illness†	2.273	-0.016	0.037	0.023	-0.065
	(2.027)	(0.093)	(0.134)	(0.153)	(0.121)
Waiting time before being attended (min)†	4.744	-0.040	-0.133	-0.151	0.064
	(1.012)	(0.093)	(0.108)	(0.135)	(0.082)
Visited traditional health practitioner (yes/no)†	0.257	-0.017	0.016	0.034	-0.039
	(0.437)	(0.032)	(0.029)	(0.030)	(0.019)
Was examined by in-charge/doctor (yes/no)	0.411	0.044	-0.049	-0.070+	-0.041
	(0.492)	(0.041)	(0.032)	(0.042)	(0.025)
Time of examination (min)	3.403	0.048	-0.099	0.015	-0.002
	(0.761)	(0.066)	(0.070)	(0.091)	(0.083)
Paid for services (yes/no)	0.179	0.010	-0.008	-0.013	-0.005
	(0.384)	(0.023)	(0.025)	(0.042)	(0.024)
Received medication (yes/no)	0.709	0.000	-0.003	0.000	-0.024
	(0.454)	(0.036)	(0.023)	(0.033)	(0.034)
Satisfied with services at health center (yes/no)	0.682	0.048	-0.026	-0.038	-0.011
	(0.466)	(0.033)	(0.031)	(0.038)	(0.026)
Number of observations	6704	2738	4858	4858	3687

Note: Column 1 reports sample means (and standard deviations below); Column 2 reports effects of sub-county barazas (and standard errors below); Column 3 reports effects of information barazas (and standard errors below); Column 4 reports effects of deliberation barazas (and standard errors below); Column 5 reports effects of district barazas (and standard errors below). **, * and + denote significance at the 1, 5 and 10 percent levels using randomization inference. Reported standard errors are clustered at the level of randomization. † indicates that the outcome was included in the index for the confirmatory analysis. Distance to nearest govt health facility, number of days sick, number of days work/school missed due to illness, waiting time before being attended, and time of examination are trimmed at 5 percent and transformed using the inverse hyperbolic sine transformation.

We also look at how the schools were managed, and how parents were involved. For instance, we look at whether the school had a Parent Teacher Association (PTA) and a School Management Committee (SMC). Almost all schools had a PTA. We further find that 92 percent of households state the primary or secondary school attended by any of their children had a SMC. However, not all households were informed about SMC meetings. The baraza intervention did not seem to affect how schools were managed, how parents could participate, or how information was shared. Again, statistical power may be an issue in many of these comparisons, as illustrated by the large MDEs in the online repository. Finally, we asked households if an inspector had visited the school in the year before the survey. We find that about 64 percent of households indicated that schools were inspected. Surprisingly, this proportion reduced as a result of the information component of a baraza.

Table A.6: Treatment-control differences (ANCOVA) - Education

	mean	sc baraza	information	deliberation	dist baraza
Number of children in UPS or USE†	1.797	0.149	-0.168	-0.078	0.021
	(1.914)	(0.139)	(0.101)	(0.109)	(0.136)
Distance to public school (km)†	1.420	0.025	-0.047	-0.044	-0.002
	(0.763)	(0.057)	(0.067)	(0.071)	(0.042)
School has complete boundary fence (yes/no)†	0.347	0.064	-0.061	-0.057	-0.008
	(0.476)	(0.048)	(0.046)	(0.049)	(0.045)
School has electricity (yes/no)	0.338	0.165**	-0.040	-0.017	0.035
	(0.473)	(0.049)	(0.042)	(0.049)	(0.038)
School has water facility (yes/no)†	0.703	0.106*	-0.023	0.026	0.073
	(0.457)	(0.041)	(0.048)	(0.050)	(0.050)
School has PTA (yes/no)	0.945	-0.007	-0.029	0.000	0.000
	(0.227)	(0.014)	(0.019)	(0.028)	(0.012)
School has SMC (yes/no)†	0.915	0.008	-0.034	0.002	0.037
	(0.279)	(0.024)	(0.023)	(0.033)	(0.020)
Is informed about SMC (yes/no)†	0.882	0.021	-0.036	-0.042	0.009
	(0.323)	(0.023)	(0.024)	(0.032)	(0.019)
Inspectors visited schools (yes/no)†	0.639	-0.004	-0.075+	-0.035	0.015
	(0.480)	(0.051)	(0.043)	(0.048)	(0.036)
Number of observations	6704	2738	4858	4858	3687

Note: Column 1 reports sample means (and standard deviations below); Column 2 reports effects of sub-county barazas (and standard errors below); Column 3 reports effects of information barazas (and standard errors below); Column 4 reports effects of deliberation barazas (and standard errors below); Column 5 reports effects of district barazas (and standard errors below). **, * and + denote significance at the 1, 5 and 10 percent levels using randomization inference. Reported standard errors are clustered at the level of randomization. † indicates that the outcome was included in the index for the confirmatory analysis. Distance to public school was trimmed at 5 percent and transformed using the inverse hyperbolic sine transformation.

References

Banerjee, Abhijit and Esther Duflo. 2006. "Addressing Absence." *Journal of Economic Perspectives* 20(1):117–132.
URL: *https://www.aeaweb.org/articles?id=10.1257/089533006776526139*

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

Björkman Nyqvist, Martina, Damien de Walque and Jakob Svensson. 2017. "Experimental evidence on the long-run impact of community-based monitoring." *American Economic Journal: Applied Economics* 9(1):33–69.

Hoenig, John M and Dennis M Heisey. 2001. "The Abuse of Power." *The American Statistician* 55(1):19–24.
URL: *https://doi.org/10.1198/000313001300339897*

Raffler, Pia, Daniel N. Posner and Doug Parkerson. 2020. "Can citizen pressure be induced to improve public service provision." *Unpublished manuscript* .