Bjorn Van Campenhout
Nassul Kabunga
Tewodaj Mogues
Caroline Miehe

# Community advocacy forums and public service delivery in Uganda

## Impact and the role of information, deliberation and administrative placement

November 2021

Impact Evaluation Report 136

Public Administration

3ie International Initiative for Impact Evaluation

## About 3ie

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed, equitable, inclusive and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesise and quality assure evidence of what works, for whom, how, why and at what cost.

## 3ie impact evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

## About this report

3ie accepted the final version of the report, *Community advocacy forums and public service delivery in Uganda: Impact and the role of information, deliberation and administrative placement*, as partial fulfilment of requirements under grant PW2.15 issued under Policy Window 2. The content has been copyedited and formatted for publication by 3ie.

The 3ie technical quality assurance team for this report comprises Sayak Khatua, Francis Rathinam, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Sebastian Martinez. The 3ie editorial production team for this report comprises Tanvi Lal and Akarsh Gupta.

 Please direct any comments or queries to the corresponding author, Bjorn Van Campenhout, at: b.vancampenhout@cgiar.org.

Funding for this impact evaluation was provided by the Bill & Melinda Gates Foundation. A complete listing of all of 3ie's donors is available on the 3ie website.

Suggested citation: Campenhout, BV, Kabunga, N, Mogues, T and Miehe, C, 2021. *Community advocacy forums and public service delivery in Uganda: Impact and the role of information, deliberation and administrative placement,* 3ie Impact Evaluation Report 136. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: https://doi.org/10.23846/PW2IE136

Badges earned: Open Data Open Materials Preregistered+

Data, replication codes, and readme file are available at: https://doi.org/10.7910/DVN/1J8CIF; preregistration and the pre-analysis plan are available at: https://www.socialscienceregistry.org/trials/5247

Cover photo: Denis Onyodi / URCS-DRK-Climate Centre

# Community advocacy forums and public service delivery in Uganda: Impact and the role of information, deliberation and administrative placement

Bjorn Van Campenhout
International Food Policy Research Institute, Belgium

Nassul Kabunga
Uganda Bureau of Statistics, Uganda

Tewodaj Mogues
International Monetary Fund, USA

Caroline Miehe
KU Leuven, Belgium

**Impact Evaluation Report 136**

**November 2021**

## Acknowledgements

This study received funding from the International Initiative for Impact Evaluation (3ie) under grant number RFQ PW2.18.UG.PG. This research was partly funded by the Consultative Group on International Agricultural Research (CGIAR) Research Program on Policies, Institutions, and Markets (PIM), led by the International Food Policy Research Institute (IFPRI) and carried out with support from CGIAR Fund contributors (https://www.cgiar.org/funders/) under grant number PIM 203002.002.221. We would like to thank Alvina Erman, Jennifer Smart, Marc Charles Wanume, Wilberforce Walukano, Fiona Nattembo, Emmanuel Bizimungu, Leocardia Nabwire, Richard Ariong and people at the Office of the Prime Minister, in particular, Joseph Muserero.

# Executive summary

## Introduction

Ineffective monitoring and weak accountability mechanisms have impaired Uganda's development since its independence in 1962. In 2009, the Government of Uganda initiated community advocacy forums, also known as barazas, to involve the public in holding the government accountable for its performance in relation to the resources spent and eventually improve public service delivery. The baraza programme was initiated by the president of Uganda and implemented by the Office of the Prime Minister.

We proposed a cluster randomised control trial to evaluate the barazas. This study's overall purpose was to establish, in a rigorous way, whether the programme had an impact on public service delivery. A second objective of the study was to inform policymakers about the effectiveness of barazas organised at lower administrative levels (the sub-county) relative to barazas organised at a more aggregate level (the district). The study also set out to explore pathways through which community advocacy forums may affect outcomes. Using a two-by-two factorial design, it differentiated between the impact of: (1) providing citizens with information; and (2) letting citizens engage with public servants and politicians.

## Intervention

To achieve these objectives, we designed four interventions. Sub-county-level barazas were implemented at sub-county level and included information and deliberation. To study the relative importance of both components, we used these barazas and removed either the information component or the deliberation component. District-level barazas were similar to sub-county-level barazas (including information and deliberation), but were organised at district level. We trained local government officials to ensure adherence to our intervention protocols and the Office of the Prime Minister, our main implementing partner, rolled out the interventions.

## Evaluation questions, methods, design, sampling and data collection

These interventions correspond to the following evaluation questions:

- What is the impact of sub-county-level barazas on public service delivery?
- What is the relative importance of the information component and of the deliberation component?
- What is the impact of district-level barazas on public service delivery?

A baseline survey with more than 12,500 households and 400 government officials was conducted in 2015. While it was initially assumed the survey would take two years, the Office of the Prime Minister faced various implementation challenges. Four years after the baseline survey, with about 50 per cent of planned barazas implemented, a trade-off needed to be made between waiting for the remaining barazas to be completed and conducting the endline survey after partial rollout. We decided to collect endline data and employ estimation and data collection strategies to control for potential selection bias. In 2020, endline data were collected on 6,700 households and 260 government officials.

In the first part of this study, we strictly follow a pre-analysis plan that summarises a range of outcomes corresponding to four main sectors (agriculture, infrastructure, health and education) as indices and one overall index of public service delivery. In the second part of the study, we proceed in a more exploratory way: (1) we provide a detailed analysis of individual outcomes; (2) we look at each of the sectors in more detail and for changes in behaviour barazas explicitly target; and (3) we provide results on changes in the perception of citizens on a range of issues. In the third part, we explore heterogeneity in the treatment effect.

## Findings

Judging by the pre-analysis plan's summary indices in the first part, we find little evidence that the baraza intervention had an impact on public service delivery. The only exception is agriculture, where sub-county-level barazas have a positive impact and this impact is superior to the (lack of) impact associated with district-level barazas. The second and third parts add more nuance to this conclusion. For instance, we find that in the agricultural sector, sub-county-level barazas significantly increase access to agricultural extension, a common practice to transmit agricultural information and technologies to farmers. However, this seems to be driven by households that live close to the sub-county headquarters.

Looking at infrastructure outcomes, we find that sub-county-level barazas reduce waiting times at the primary water source. Interestingly, this effect seems to be strongest in more remote areas. For health, we only find effects if we restrict the sample to sub-counties where officials recall that a baraza happened; then, we find that the information and deliberation components affect the use of government health facilities. For education, we see an increase in enrolment rates, but only if enough time has passed between the intervention and endline data collection.

## Cost analysis

The complex picture that emerges from this analysis also means that conclusions in terms of cost-effectiveness are ambiguous. For instance, with regard to public service delivery related to water infrastructure, district-level barazas are far more cost-effective than sub-county-level barazas, because they can reach many more households. However, with respect to agriculture, sub-county-level barazas are the most cost-effective. The same holds for comparisons between the cost-effectiveness of the deliberation and information components. In general, we find that because baraza interventions affect large numbers of households and at relatively low cost, the rate of return is substantial, even if treatment effects are small in size.

## Discussion

These mixed results are puzzling, especially because prior qualitative research suggested an effect of the intervention (Van Campenhout et al. 2018). We suspect that the nature of the intervention can explain the lack of quantifiable impact. Barazas address various issues in heterogeneous settings: different sub-counties face different challenges, so that different issues are discussed and prioritised during the barazas. Consequently, the actual baraza treatments may be far from standardised, and their impact may be highly localised and context specific. As a result, a focus on average

treatment effects may fail to identify a significant impact. The effect is attenuated, because it is averaged over many sub-counties that, in reality, received ‘different’ types of barazas. While it is good that barazas tend to focus on and potentially affect areas that are the most problematic, this complicates the estimation and might be the reason we find only limited effects. Hence, barazas might work but we are unable to detect this. Concerns related to non-standardised treatments are confirmed when looking at heterogeneous treatment effects and a case study of access to water in Bagezza sub-county. That is why we recommend baraza meetings even though they do not have a measurable effect on our pre-registered indices. We suggest a mix of district- and sub-county-level barazas, and recommend the implementation of full barazas that are held several times; for instance, every two years.

# Contents

# List of figures and tables



## Appendix tables

## Abbreviations and acronyms

| | |
|---|---|
| ANCOVA | Analysis of covariance |
| ANOVA | Analysis of variance |
| CAO | Chief administrative officer |
| GoU | Government of Uganda |
| GPS | Global positioning system |
| HC2 | Health centre 2 |
| HC3 | Health centre 3 |
| LC1 | Village local council chairperson |
| LC3 | Sub-county local council chairperson |
| LC5 | District local council chairperson |
| MDE | Minimal detectable effects |
| NAADS | National Agricultural Advisory Services |
| NGO | Non-governmental organisation |
| OPM | Office of the Prime Minister |
| PTA | Parent teacher association |
| RCT | Randomised controlled trial |
| RDC | Resident district commissioner |
| SMC | School management committee |
| UPE | Universal primary education |
| VHT | Village health team |

# 1. Introduction

Since Uganda's independence in 1962, political turmoil and economic mismanagement have thwarted the country's development efforts. In the mid-1980s, after attaining relative stability, the Government of Uganda (GoU), supported by development partners, initiated reforms to address the development challenges of the time. Notable among these initiatives was the liberalisation of the economy and the introduction of a decentralised system of governance (Francis and James 2003; Benin et al. 2007). Decentralisation was particularly viewed as a suitable mechanism for addressing welfare and political challenges by improving the efficiency of public service delivery, formulating more appropriate services and bringing representative governance closer to citizens (Steiner 2007; Francis and James 2003). Major ingredients of decentralisation are enhancing empowerment and building a sense of citizens' ownership to actively participate in planning, implementation and evaluation of development interventions in their locations, and improve local leaders' and service providers' accountability and responsiveness (Burki et al. 1999).

Ineffective monitoring and weak accountability mechanisms have greatly affected the realisation of the benefits of decentralisation in Uganda, especially with respect to beneficiaries holding service providers accountable (Björkman and Svensson 2009; Reinikka and Svensson 2004). In this regard, in 2009, under the stewardship of the Office of the Prime Minister (OPM), the GoU initiated community advocacy forums – or citizen barazas – with the objective of 'enhancing public involvement in holding the government accountable for service delivery in relation to the resources spent' (OPM 2013).

By now, barazas have been implemented in Uganda for more than 10 years. They were first piloted in financial year 2009–2010 in eight communities.[1] Since then, efforts have been underway to roll out barazas in all sub-counties in the country. During the full-scale implementation phase in financial year 2010–2011, 16 more sub-counties in 8 districts held a baraza meeting. And, by the last quarter of 2011–2012, of the country's 1,340 sub-counties spread over 112 districts, 267 had held a baraza meeting. At the beginning of financial year 2012–2013; however, changes in implementation were suggested: subsequent barazas would target district-level reporting to increase participation at a higher level and, at the same time, reduce implementation costs.

As barazas continued to be rolled out beyond the pilot communities, a rigorous evaluation of their effectiveness was still outstanding. This study's overall purpose is to establish, in a rigorous way, whether the programme had an impact on public service delivery. A second objective of the study is to inform policymakers about the effectiveness of barazas organised at lower administrative levels relative to barazas organised at a more aggregate level. The study also sets out to explore pathways through which community advocacy forums may affect outcomes, as we differentiate between the impact of providing citizens with information and the impact of letting

[1] The initial pilot barazas were undertaken in eight lower-level local governments (generically referred to as sub-counties) of the four districts of Masaka, Bushenyi, Kumi and Nebbi, which are respectively located in the four geographical regions of Uganda: Central, Western, Eastern and Northern.

citizens engage with public servants and politicians. At the time of the proposal, the GoU shared the same aspiration to inform policy on programme effects of service delivery to local communities (OPM 2013), because at that time there had been no formal study conducted to test the actual achievements of the baraza initiative against the set objectives. From a policy perspective, it was also important to assess whether the switch from sub-county-level barazas to district-level barazas was cost effective. To date, the OPM has been eager to learn about the results from the different components of the impact evaluations and (preliminary) results have been presented at various high-level meetings between the OPM and the research team.

There have been several studies that look at the impact community involvement has on public service delivery, many of them using Uganda as a case study. A landmark study by Björkman and Svensson (2009) looks at the impact of a community-driven local accountability project on primary healthcare provision in Uganda. They find that the intervention resulted in significant improvements in healthcare delivery, use and health outcomes (most notably child mortality and weight-for-age z-scores) after one year. Björkman Nyqvist and colleagues (2017) confirm that despite minimal follow-up, these effects are still present more than four years after the initial intervention. More recently, however, Raffler and colleagues (2018) have come to more nuanced conclusions when testing an intervention closely modelled on that of Björkman and Svensson (2009). The study, involving a three-wave panel of more than 14,000 households and a factorial design to break down the intervention into its two most important components – similar to what we use – validates the power of information provision to change the behaviour of front-line service providers, but casts doubt on the ability to foster community monitoring or generate improvements in health outcomes, at least in the short run.

Waddington and colleagues (2019) also discuss whether citizens' engagement in the planning, management and oversight of public services affects the quality of and access to those services and citizens' quality of life. In some programmes, citizens participate in setting the priorities for and planning of local services (Touchton and Wampler 2014; Goncalves 2013; Diaz-Cayeros et al. 2014; Beuermann and Amelina 2014; Ananthpur et al. 2014; Giné et al. 2018; Humphreys et al. 2014; Beath et al. 2013). Other programmes evaluate transparency mechanisms that aim to disclose and disseminate information, such as public official or service provider performance information interventions (Humphreys and Weinstein 2012; Grossman and Michelitch 2018; Timmons and Garfias 2015; Capuno and Garcia 2010). Moreover, evaluations of accountability mechanisms included citizens' feedback or monitoring mechanism interventions to hold public service providers and institutions responsible for executing their powers and mandates according to appropriate standards (Berman et al. 2017; Alhassan et al. 2016; Grossman et al. 2017; Björkman et al. 2009; Björkman et al. 2017; Gullo et al. 2017; Bradley and Igras 2005; Molina 2014). This review shows, on the one hand, that interventions promoting citizens' involvement by improving direct engagement between service users and service providers are often effective in stimulating citizen's engagement and improving public service delivery, but complementary interventions that address bottlenecks around service providers' supply chains and service users are needed to improve well-being. On the other hand, interventions that promote citizens' engagement by increasing citizens' pressure on politicians to hold service providers accountable and thus improve governance often do not influence service delivery.

Our study contributes to this literature in various ways. First, this study is one of the few that consider the role of administrative placement in the effectiveness of community monitoring. The level at which the intervention occurs may affect its effectiveness in opposing ways (Donato and Mosqueira 2016). Interventions at a more local level might result in more relevant issues being scrutinised. However, qualitative explorations suggest that, often, issues raised in lower-level barazas fall under the responsibility of higher levels of government or other institutions that are beyond the operational jurisdiction of the participating officials (Van Campenhout et al. 2018). This may be less of a problem when barazas are organised at district level. Most other studies consider interventions placed at fairly local levels. For instance, the intervention in Raffler and colleagues (2018) was implemented in health centres and their associated catchment areas consisted of only a few villages.

Second, we explore pathways through which community advocacy forums may affect public service delivery. A two-by-two factorial design enables us to differentiate between the impact of: (1) providing citizens with information related to budgeting and planning; and (2) letting citizens engage with public servants and politicians in a facilitated question-and-answer session. On the one hand, informational interventions can increase political accountability (Dunning et al. 2019). A citizen who is informed about the performance of politicians and civil servants can monitor the latter and apply pressure (Raffler et al. 2018). There is some evidence that providing citizens with information about public services can increase their ability to hold leaders accountable to improve public service delivery (Pandey et al. 2009; Gilens 2001). On the other hand, deliberation can also increase the quality of public services. Citizens can confront their leaders with urgent and important matters and threaten them if they are not performing. Creating a platform where stakeholders can meet and interact could also increase mutual understanding and result in a better relationship. The impact of deliberation has also been the subject of empirical analysis (Björkman Nyqvist et al. 2017; Goeree and Yariv 2011; Fujiwara and Wantchekon 2013).

Third, our study evaluates the impact of a government initiative, which may instigate an entirely different set of dynamics than interventions organised by local or international non-governmental organisations (NGOs). It has been argued that successful devolution can only happen in the context of a strong state that is able to ensure consistent regulation and a well-informed public backed up by a participatory political culture (Golooba-Mutebi 2005). Many of the actors involved might find that NGOs are not mandated when it comes to public services such as health or education. Furthermore, it is likely to be easier to reallocate resources to problems identified during barazas if they are organised by the government. This is also consistent with suggestive evidence in Raffler and colleagues (2018), which finds that the presence of sub-county officials during their community-based monitoring intervention boosted the impact of the intervention. However, effects may also work in the opposite direction. For example, an intervention to reduce absenteeism in government public health facilities in India was initially very successful, but ceased to have any impact after the local bureaucracy started providing official excuses for most of the nurses' absences (Banerjee et al. 2008). Most of the other studies that are the closest to our study partnered with NGOs for implementation (e.g. Björkman and Svensson 2009; Raffler et al. 2018).

Fourth, barazas take a comprehensive, multi-sectoral approach that enables cross-sectoral planning and, potentially, allows for reallocations across sectors. Some of the problems users mentioned the most, such as hygiene in health centres or accessibility, involve cooperation between heads of different sectors (e.g. health and infrastructure for water access in health centres or access roads). Bringing sector heads together and confronting them with citizens' priorities may increase information sharing and cooperation between them (Van Campenhout et al. 2018). Most existing studies focus on a single sector; the health sector seems to be especially popular for community monitoring interventions (e.g. Arkedis et al. 2019; Björkman and Svensson 2009; Raffler et al. 2018).

Finally, we evaluate a high-profile policy intervention that receives broad support both within government and among citizens in Uganda. Evaluating policy interventions has its challenges and this one is no exception. As a result, such research has become rare – present-day randomised controlled trials (RCTs) often bypass the political resistance to randomisation among governments, development workers and beneficiaries, as the nature of the partners has changed (NGOs rather than governments) and the interventions have become 'relatively trivial' (de Souza Leão and Eyal 2019).

In this study, we start by providing a brief overview of the government programme we evaluated and explain the theory of change behind the components of the intervention. We then present the four main research questions and provide details on the cluster RCT we used to answer these questions. This section also provides information on the sampling frame and presents detailed power simulations that account for the consequences of the implementation challenges. This is followed by an explanation on how the implementation deviated from what was planned, and the strategies we used to diagnose and remedy the potential bias the deviation introduced. We then present the findings, starting with balance tables and results of a pre-registered analysis. We provide further details and look at outcomes that were not pre-registered, to explore some of the mechanisms behind the intervention. This part also includes an extensive analysis of sub-county-level data that were collected from government officials. We then present heterogeneous treatment effects and reflect on the partial rollout as a threat to the study's validity. Further, we provide a cost-benefit analysis. The penultimate section provides a discussion of the results and the final section concludes the report.

# 2. Intervention

## 2.1 Description

Barazas are platforms for enhancing **information** sharing between policymakers (the clients), public servants (the implementers), and beneficiaries of public goods and services (the users). In addition, they provide an opportunity for citizens to ask questions to policymakers and civil servants, and **deliberate** among themselves. With barazas, citizens, in particular, have the opportunity to participate in the policy process by directly engaging with service providers and to demand accountability for the use of public resources. It is expected that, ultimately, barazas will contribute to effective monitoring, and increase accountability and transparency among all stakeholders.

A typical baraza is initiated from the centre, with the OPM mobilising district and sub-county officials. These include: the chief administrative officer (CAO) as the head of public service delivery at the district level; the resident district commissioner (RDC) as the direct representative of the president; the district local council chairperson (LC5) as the representative of political leadership at the district level; and the various sector heads (agriculture, education, infrastructure and health). For barazas organised at the sub-county level, the sub-county local council chairperson (LC3) – the sub-county-level equivalent of the CAO (the sub-county chief) and the LC5 – also has an important role.

The OPM, in consultation with the district leaders (RDC, CAO and LC5) and other stakeholders, agrees on the date and a neutral venue for holding the baraza event. Again in consultation with the district leaders, a viable moderator and an interpreter into the local language, where applicable, are identified to guide the baraza forum. Village mobilisers and community resource persons are used to publicise the event. A few days before the baraza event, the community mobilisation efforts are further reinforced by adverts in local media, such as radio announcements, printed banners, posters and fliers, and mobile public address systems.

A baraza meeting is chaired by the office of the RDC in each district. In front of the audience, including local citizens, invited opinion leaders, elders and journalists, the RDC seeks accountability and feedback from each head of a major sector. Sector heads are required to present: (1) what services were planned to be delivered in the sub-county (or the entire district in case of a district-level baraza); (2) what was actually delivered and in what quantity and quality; and (3) what issues and challenges emerged and the way forward. The RDC then seeks reactions and feedback from citizens on whether what was presented was what was planned for and actually implemented in different locations. Sector heads are then given another opportunity to clarify or react to any issues citizens raise.

In our study, we do not only want to test whether barazas work. We want to learn which of the main components – the deliberation component or the information component – is responsible for most of the effect. Finally, we also strive to directly compare the effectiveness of district-level barazas to that of sub-county-level barazas. Therefore, we differentiate between four types of barazas: (1) a sub-county-level baraza, (2) an information baraza, (3) a deliberation baraza and (4) a district-level baraza.

The sub-county-level barazas are basically the same barazas the OPM implemented at the district level. They have both an information and a deliberation component. To study the relative importance of these two components, we use this baraza as a starting point and remove either the information or the deliberation component from the generic sub-county-level baraza to test the components' relative importance.

The information component of a baraza involves templates developed to be filled by officials and mounted at a central location in each parish of the district two weeks before the baraza. The templates were designed to inform citizens about planned and actual public expenditures for the previous fiscal year, achievements and challenges encountered during that year, and planned expenditures and targets for the next fiscal year. The sub-county chief needed to fill this for each of the four sectors (agriculture, infrastructure, health and education).

On the day of the baraza event, the CAO provided a brief presentation on the overall budget/finances for the fiscal year, and main achievements and challenges in service delivery, and introduced local officials. After a brief intervention by the OPM, local officials responsible for each sector presented more or less the same information as that required for the templates. An information-focused baraza allowed for the facilitator to collect and ask only 10 clarifying questions.

For the deliberation component of the barazas, posters were also mounted in each parish of the sub-county, but only to announce that a baraza would be held at a particular date and place. At the baraza event, after a brief introduction by the RDC, citizens were guided to break into five groups by sector, discuss problems they faced and draw up a list of priority issues that needed to be addressed. Facilitators in each group were required to anonymously collect these issues and concerns. Facilitators were expected to focus the discussion on what had been done well and what had been the problems during the past year. The discussions were also to result in agreement on what should be done in the next fiscal year. After the break-out sessions, officials were asked to react to the specific comments and requests.

District-level barazas were very similar to sub-county-level barazas (i.e. with both an information and deliberation component), except for the fact that district-level barazas were organised at the district headquarters, and all sub-county chiefs and sub-county chairpersons (LC3s) of each sub-county within the district were expected to attend in case issues arose related to their sub-county.

## 2.2 Theory of change

### *2.2.1 The impact of (sub-county-level) barazas*

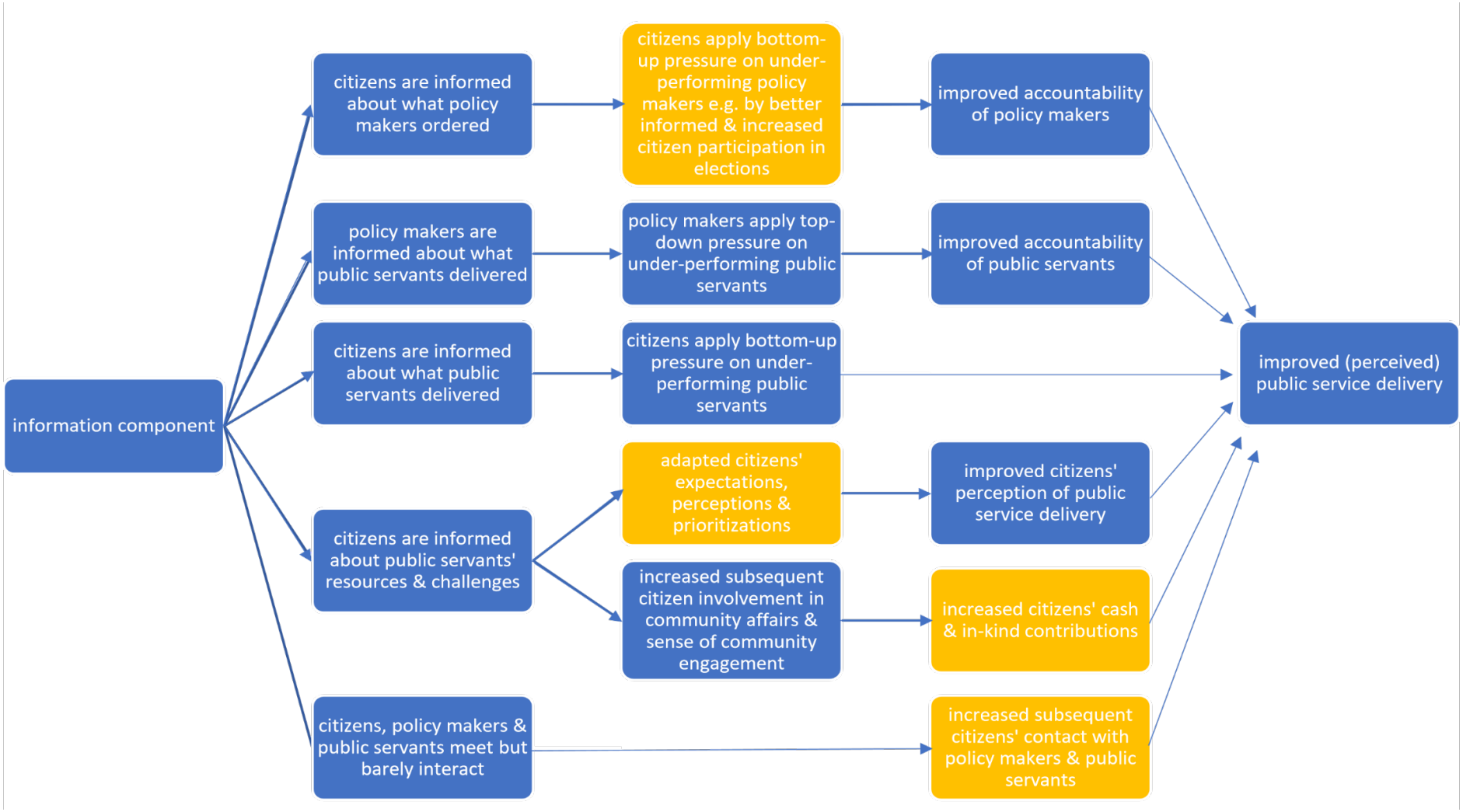
The baraza intervention fundamentally seeks to improve public services by enhancing the accountability of local public decision makers and service providers. The baraza intervention, as conceived by the OPM, is a fairly standard community-based monitoring intervention that combines the provision of information with the possibility for citizens to engage with each other, decision makers and public servants at a fairly local level. Such community-based monitoring has become a popular tool to increase service delivery. However, not all such interventions appear to be successful (Olken 2007). As the sub-county-level baraza combines information and deliberation components, it also works through the components' (combined) theories of change (*Figure 1* and *Figure 2*).

The broad nature of the baraza intervention means that many issues can come up during the meetings. This can make it hard to determine in advance where impact will emerge. If many communities struggle with the same issue (e.g. absences of functioning toilets for girls at the public primary school), it will be easier to pick up an effect of the baraza on that particular issue, even though there may not be an impact on the education sector as a whole. However, different communities might struggle with different issues. In that case, it could be that no effects are found on a particular issue, but all effects within a sector go in the same direction. Furthermore, primary outcomes are mediated through different channels, including enhanced contact with policymakers and service providers, increased citizens' participation in elections, more cash and in-kind contributions to the commons, and changes in perceptions and prioritisation. Some of these mediating channels are less specific and less localised.

### *2.2.2 The information mechanism*

Information treatments are only effective if a lack of information hinders the delivery of public services. We are confident that, here, incomplete and asymmetric information actually blocks effective service delivery for three reasons. First, information provision was one of the main aims of the baraza programme according to the OPM. Second, the hypothesis that information is key in our context derives from previous literature (e.g. Raffler et al. 2018). Third, information frictions were named as a main constraint to public service delivery during our qualitative diagnostic work (Van Campenhout et al. 2018). We identified the barazas' potential to simply reduce information inefficiencies. For these reasons, we confidently decided to isolate and explore the information component in the (existing) intervention.

**Figure 1: Information mechanism**



In situations characterised by incomplete and asymmetric information, targeted efforts to fill knowledge gaps can make a big difference. Indeed, the relationship between citizens and elected officials is reminiscent of the principle-agent problem. In essence, there are three players (elected politicians, civil servants and citizens), with only partly overlapping information sets and potentially competing interests.[2] Bringing stakeholders together in town hall-type meetings is assumed to reduce information asymmetries. Increasing the knowledge of all stakeholders about what the clients (policymakers) ordered and what the implementers (public servants) delivered could be an effective way to improve the quality of public service delivery by: (1) allowing citizens to monitor and apply bottom-up pressure on underperforming civil servants; and (2) increasing top-down pressure on underperforming civil servants by revealing to politicians the discrepancy between what was promised and actual performance, thus enhancing the accountability of service providers. It can also improve the accountability of local public decision makers by allowing citizens to apply bottom-up pressure on underperforming policymakers (e.g. by participating more or having better informed voters in elections).

There is some evidence that channelling information to citizens about the quantity, modality and quality of public services, as well as about the investments and policy decisions made by politicians, bureaucrats and service providers, can increase the ability of users to hold leaders accountable to improve service provision. For example, using a field experiment in India, Pandey et al. (2009) establish that community information campaigns about states' school management obligations had a positive impact on school performance. Gilens (2001) identifies the significant influence of providing policy facts on the public's political judgment. Grossman and Michelitch (2018) disseminate information about job performance for randomly selected Ugandan politicians. While this increases job performance for the politicians on a range of criteria, they find no impact on public service provision. A recent review of 48 empirical studies on the impacts of information on governance and service delivery also suggests that the availability of information alone may not suffice. Information must be deemed relevant to its recipient, and individuals must have both the power and incentives to act on the information (Kosec and Wantchekon 2020).

The information component could also be important in managing client expectations. Citizens might have exaggerated beliefs about the resources at the disposal of decision makers and service providers, or they may not fully appreciate the challenges civil service providers face when doing their job. For instance, during focus group discussions, service providers mentioned that citizens sometime blame officials for things they have no control over. More generally, information could help sensitise citizens to the role of public service provision (e.g. making sure boreholes are present) and that there are limits to what they should expect (e.g. citizens are still required to boil water). When information can rectify inflated expectations and change perceptions, we might not find changes in the quantity or quality of public services, but we could still find modifications in citizens' perceptions of the quality of these services. Informing citizens

[2] Because the public servant must be responsive to the needs of both the client and the community at the same time, the problem can be characterised as a multiple or common agency problem, which adds a collective action component to the standard principle–agent problem (Bernheim and Whinston 1986).

about the resources and challenges of public servants may also increase their involvement in community affairs and sense of community engagement and, consequently, their willingness to contribute to common goods. As citizens, policymakers and public servants meet during the information barazas, this could also raise subsequent contact among these stakeholders.

### *2.2.3 The deliberation mechanism*

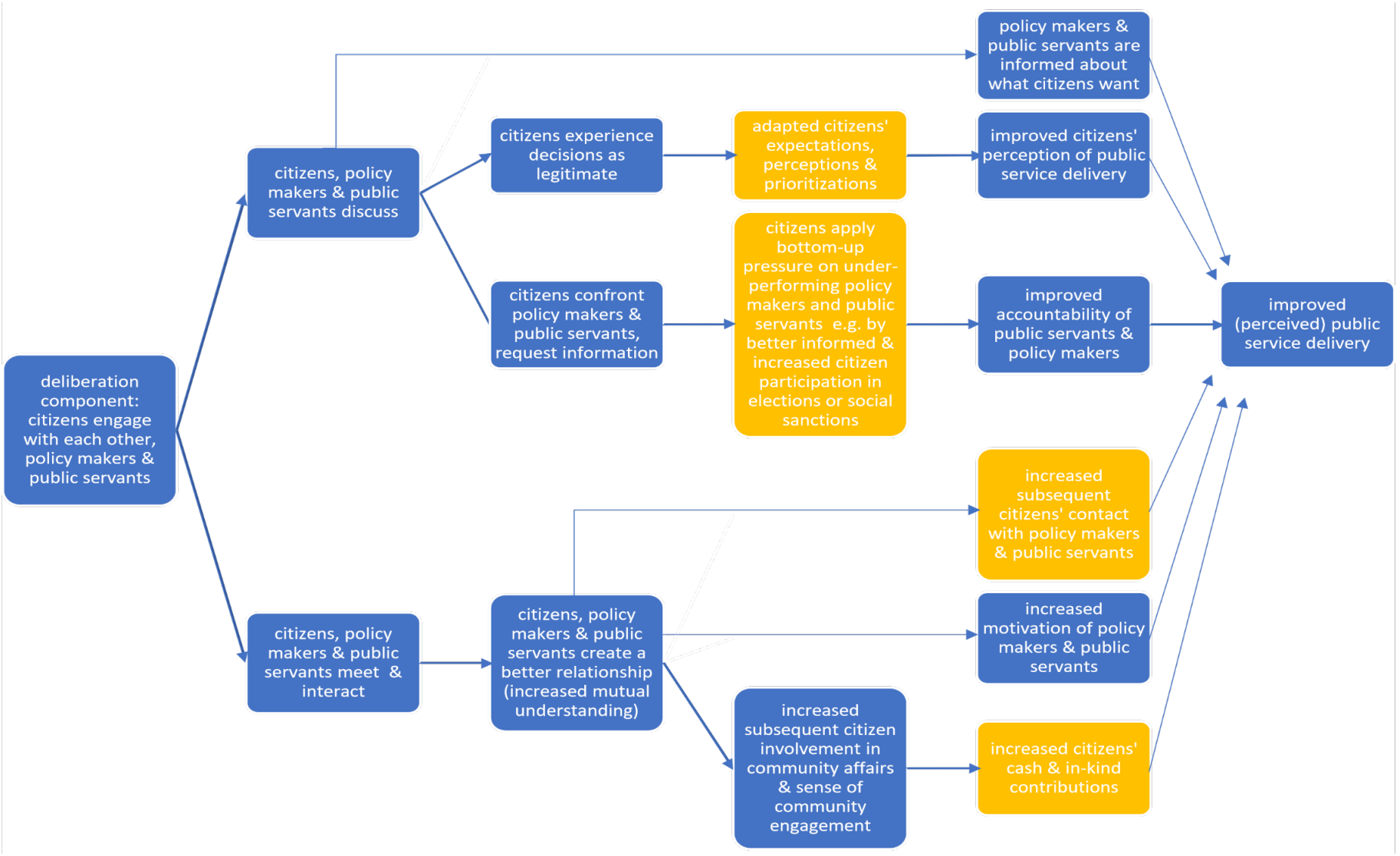
There are various ways deliberation increases the quality of public service delivery. First, it has a legitimating effect on decisions arrived at in this fashion. Effective deliberation assumes an equal voice for the arguments of both marginal and advantaged agents, and the role of evidence that supports the positions articulated. This can change citizens' expectations, perceptions and prioritisations, and improve their comprehension of public service delivery. Second, deliberation can more effectively distil social choice than simple voting and majoritarian rule, in part by building consensus both among citizens and between public servants and citizens. Policymakers and public servants are better informed about what citizens actually want. Third, deliberation has been found to positively affect the vigour and breadth of citizens' subsequent involvement in community affairs (Björkman Nyqvist et al. 2017). Deliberation provides opportunities for citizens to confront their leaders and public servants with issues, and threaten them with social and political sanctions if it is deemed that they are not performing. Citizens apply bottom-up pressure on underperforming policymakers (e.g. through better-informed and increased citizen participation in elections) and underperforming public servants (e.g. through social sanctions), thus improving public servants' and policymakers' accountability.

Creating a platform where stakeholders can meet and interact could also increase mutual understanding and create a better relationship between them. This could lead to more subsequent contact between citizens, policymakers and public servants. Elected officials and service providers could also be more motivated because of this improved relationship. Furthermore, the mutual understanding and better relationship between stakeholders may also increase citizens' involvement in community affairs and sense of community engagement and, therefore, their willingness to contribute to common goods. However, when relationships are poor, public forums that degenerate into name-and-shame sessions may make matters worse. Facilitated, collaborative meetings that jointly engage citizens and service providers in monitoring are often more effective than confrontational meetings (Waddington et al. 2019).

Deliberation also affects information flows. In a baraza, the information component is primarily designed to inform citizens about service providers' activities. To some extent, citizens are passive recipients of this information and officials report what they consider relevant or may even attempt to misrepresent the facts. If citizens can engage with policymakers and civil servants, they may request information relevant to them.

Impacts of deliberative processes have also been the subject of empirical analysis. For example, in addition to the increased community participation mentioned above, experimental evidence shows that deliberative processes make decision outcomes less sensitive to the institution (e.g. voting) rules that bring them about (Goeree and Yariv 2011) or may reduce the prevalence of clientelism (Fujiwara and Wantchekon 2013).

**Figure 2: Deliberation mechanism**



### *2.2.4 Administrative placement*

The baraza intervention can also be distinguished by the administrative level it is implemented at. It was originally planned to implement barazas at the sub-county level, but from 2012 onwards, more and more barazas were implemented at the district level. This administrative placement dimension immediately points to a potential trade-off between attempting to achieve breadth of coverage (through district-level barazas) and attending to depth and quality of coverage (through sub-county-level barazas). While conducting a district-level baraza may be cheaper than conducting sub-county-level barazas in all sub-counties of a district, it is not clear a priori how these cost savings justify the potential reduction in effectiveness of district-level barazas in any given sub-county of the concerned district.

Whether placement at a higher or lower level is more effective will depend on the outcome and the situation. For instance, it has been argued that engaging small groups can be more effective, because they can be coordinated more easily; however, large groups may make more sense if a broader group would enjoy the desired outcome (Donato and Mosqueira 2016). Furthermore, action may be more likely if a large, rather than a small, group of people complains about a highly localised issue (Banerjee et al. 2004). It could also be that issues highlighted at the local level fall under the responsibility of higher-level authorities and vice versa.

## 2.3 Intervention monitoring plan

After completion of the baseline, we trained local government officials and designated facilitators to ensure adherence to the intervention protocols. We agreed with the OPM that for the barazas that were part of the study, facilitators would be selected from among these trained facilitators. We developed detailed scripts RDCs and facilitators were expected to follow. We also produced manuals for RDCs and facilitators. Detailed information can be found in an online appendix.

Two full-time research assistants were assigned to monitor programme implementation. They worked very closely with the OPM staff tasked with the implementation of the barazas. One researcher accompanied the OPM staff to all barazas that were part of the study. They also made sure the information for the information-focused barazas was disseminated in time.

At the end of a baraza, the RDC was required to make a report to the OPM, indicating issues that arose in the baraza meeting. This report pointed out, in particular, policy and programme implementation weaknesses and challenges, which were then expected to further feed into the general government performance management system. These reports were collected to assess implementation and adherence to the intervention protocols ex-post.

# 3. Evaluation questions, methods, design, sampling and data collection

## 3.1 Primary and secondary evaluation questions

Primary evaluation questions look at the impact of the baraza programme and its key components on public service delivery. Four comparisons are made, corresponding to the following evaluation questions:

1. What is the impact of sub-county-level barazas on public service delivery in general, at sector level, and for selected, preregistered individual outcomes?
2. What is the relative importance of the information component of a sub-county-level baraza on public service delivery in general, at sector level and for selected, preregistered individual outcomes?
3. What is the relative importance of the deliberation component of a sub-county-level baraza on public service delivery in general, at sector level and for selected, preregistered individual outcomes?
4. What is the impact of district-level barazas on public service delivery in general, at sector level and for selected, preregistered individual outcomes?

Secondary evaluation questions look into the mechanisms through which the baraza project is assumed to affect public service delivery. These include interfacing with politicians and civil servants, political participation and contributions to common goods. We also investigate how perceptions may have changed as a result of a baraza.

## 3.2 Methods

### *3.2.1 Confirmatory analysis*

In the first part of our analysis, we strictly follow a preregistered analysis plan that takes the form of a 'mock report'. This report was written in December 2019, just before endline data were collected. It contains the results of an analysis on simulated endline data for the four primary research questions. It was preregistered at the American Economic Association's RCT registry with a time stamp. Preregistration and mock reports are effective tools against fishing and false-positive science (Humphreys et al. 2013).

The mock report was prepared using LyX, an open source LaTeX front-end. All Latex and R code to replicate the analysis were placed under revision control using Git. The R scripts are automatically executed when the Lyx document is compiled (using the R package Knitr) and tables are populated. The Git repository can be found at <https://github.com/bjvca/baraza/>. The use of revision control further increases transparency and allows for easy replication (Ram 2013).

Impact is assessed as a simple treatment-control comparison, implemented using an analysis of covariance (ANCOVA) model that also controls for the region (because this was used for stratification) and the baseline outcome. When evaluating the relative importance of the deliberation and information components, we also include all interaction terms of the factorial design (Muralidharan et al. 2019). Standard errors are clustered at the level of randomisation: the sub-county level for the first three hypotheses and the district level for the last hypothesis.

The report describes explicitly which variables are used to assess impact, how they are combined into indices and what transformations are used, referencing the actual names of the variables in the endline data collection application.

This confirmatory analysis focuses on a subset of carefully selected and declared variables that are combined in indices: one overall index and four indices corresponding to the sectors. Our indices combine individual outcome variables covering the use, availability, quality and delivery of services. These outcomes were categorised into four broad sectors: agriculture, health, education and infrastructure (mainly drinking water and roads) following Anderson (2008) to account for multiple hypotheses testing. The four indices are then, in turn, combined into an overall indicator of public service delivery.

For continuous variables, 5 per cent trimmed values were used (2.5% trimming at each side of the distribution). Inverse hyperbolic sine transformations were used if skewness exceeded 1.96. Trimming was always done on end results (e.g. if the outcome was yield at the plot level, then production was first divided by plot area, then the inverse hyperbolic sine transformation was done and the end result was trimmed). Outcomes for which 95 per cent of observations have the same value within the relevant sample were omitted from the analysis to limit noise caused by variables with minimal variation.

#### *3.2.2 Detailed analysis*

A second, more exploratory part of our analysis looks at individual outcomes beyond the indices. Individual outcome variables cover a wide range of outcomes related to the use, availability, delivery and quality of public service. Using ANCOVA models, some outcomes of each of the four key sectors are compared between the different groups. Furthermore, we explore whether barazas affected aspects that are at the core of community-based monitoring, such as participation in elections of local leaders, citizens' contact with policymakers and service providers, perceptions of service quality and prioritisation, and cash and in-kind contributions. Finally, we analyse sub-county-level data collected from government officials because, in addition to household surveys, we conducted surveys with 261 government officials as respondents.

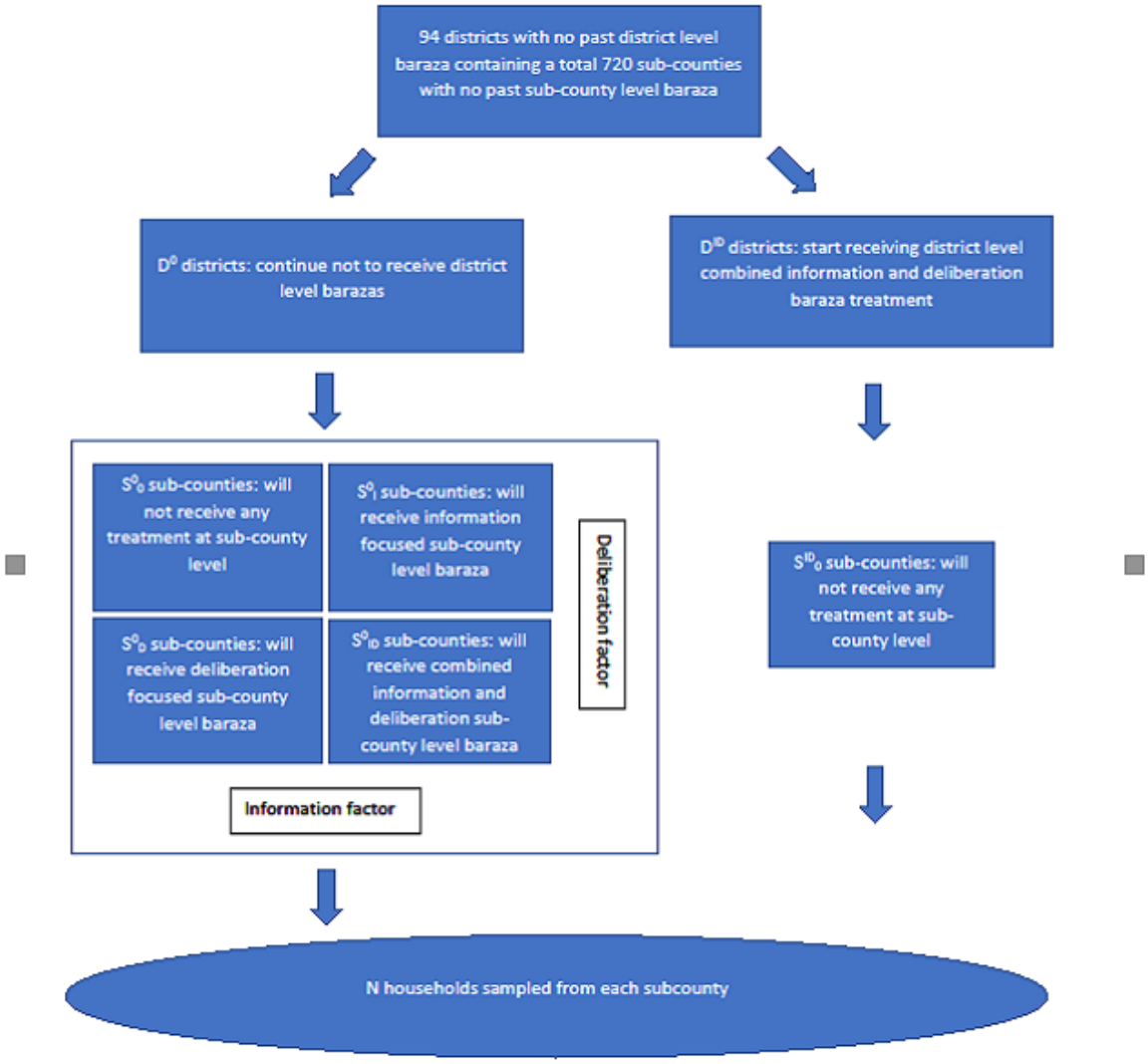
## 3.3 Evaluation design

#### *3.3.1 Study design and identification strategy*

This study proposed a nested, or two-step, randomisation design, illustrated in *Figure 3*. In the first step, we randomly allocate eligible districts to treatment and control conditions. In particular, some of the eligible districts start receiving district-level barazas that contain both the information component and the deliberation component $(D^{ID})$, while other districts do not receive a baraza at this level $(D^0)$. In the second step, we proceed with all eligible sub-counties and randomly allocate each sub-county to one of four conditions in a two-by-two factorial design. In particular, about one quarter of all eligible sub-counties sampled from $D^0$ serve as a pure control and do not receive any baraza at any level $(S_0^0)$. Another quarter receive a sub-county-level baraza that combines both information and deliberation treatment $(S_{ID}^0)$. A third quarter receive a sub-county-level baraza that consists largely of officials providing information and limited opportunity for citizens to engage $(S_I^0)$. The last quarter receives a sub-county-level baraza with a focus on citizens engaging with each other and with officials, without upfront information provision $(S_D^0)$. We also take a random sample of sub-counties from the $D^{ID}$ districts that

received the district-level baraza ($S_0^{ID}$). Within each sub-county, we sample a fixed number of households.

**Figure 3: Study design**




The above design allows us to answer the four research questions. First, to assess the impact of the sub-county-level baraza interventions as implemented by the GoU, one can compare outcomes of households sampled from $S_{ID}^0$ to those sampled from $S_0^0$. Second, to assess the relative importance of the information component of a baraza, one can compare the outcomes of all households exposed to the information component (either as a stand-alone information baraza as implemented in $S_I^0$ or as part of a combined baraza as implemented in $S_{ID}^0$) to outcomes of all households that were not exposed to the information component of the baraza, either because they did not receive a baraza at all ($S_0^0$) or because they only received a deliberation-focused baraza ($S_D^0$). Similarly, to assess the relative importance of the deliberation component of a baraza, one can compare outcomes of all households exposed to the deliberation component (either as a stand-alone deliberation baraza as implemented in $S_D^0$ or as part of a combined baraza as implemented in $S_{ID}^0$) to outcomes of all households that were not exposed to the information component of the baraza, either because they did not receive a baraza at all

$(S_0^0)$ or because they only received an information baraza $(S_I^0)$. Note that, because of the factorial design, much more information can be used to test the last two hypotheses than for the first two. Finally, to investigate administrative placement of the intervention, two comparisons are used. First, and similar to the first hypothesis, we can simply estimate the impact of district-level barazas by comparing outcomes of households sampled from $S_0^{ID}$ to those sampled from $S_0^0$. However, we can also directly compare district-level barazas to sub-county-level barazas by comparing outcomes of households sampled from $S_0^{ID}$ to outcomes of household sampled from $S_{ID}^0$.

From a policy perspective, in light of the shift from sub-county-level barazas to district-level barazas from 2012 onwards, the last comparison is the most interesting one and it was preregistered. However, the partial rollout of the intervention means that for this comparison, we are constrained by the number of sub-counties in $S_{ID}^0$ that ended up being treated. Comparisons of outcomes in areas that received a district-level baraza treatment to areas that did not receive a baraza have more statistical power, because more observations can be used.

### *3.3.2 Power and sample size calculations*

To determine the number of districts, sub-counties and households to include in the study, the original research proposal contained an extensive series of power calculations that used data from the Uganda National Household Survey of 2009/2010 and the Demographic and Health Survey of 2011 to estimate standard errors of the outcomes and inter-class correlations. Outcomes used to determine sample size included: (1) weight-for-age z-scores for children; (2) number of days unable to work as a percentage of days sick at household level; (3) number of years the average child within the household goes to school and proportion of children in the household currently attending school; (4) proportion of households visited by an extension worker in the previous year; (5) maize yields; (6) time to get drinking water (including waiting time); and (7) share of households with access to improved drinking water sources. This resulted in the selection of a total sample size of 11,500 households distributed over 230 sub-counties in 40 districts throughout Uganda on which baseline data were collected.[3] More details on the power calculations can be found in the original proposal, which is available as an online appendix.

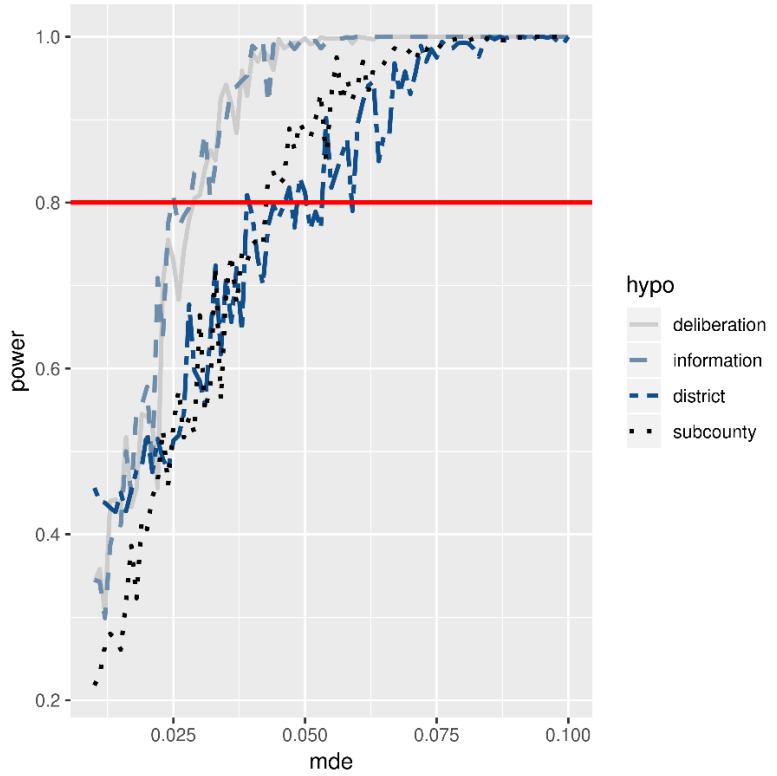
The original power calculations assumed full rollout of the intervention. However, due to implementation challenges that will be explained in detail in section 3.6 on *intervention implementation fidelity*, a series of updated power calculations were performed prior to endline data collection. In particular, we simulated a new set of minimal detectable effects (MDEs) associated with the sample we were about to collect.[4] We applied baseline data to simulate MDEs for a selection of the outcomes we used to judge the effectiveness of the intervention (and which are specified in the preregistered report). We used a standard significance level of 0.05 (double-sided).

*Figure 4* plots MDEs against power for the first outcome variable used to assess the impact of barazas on public service delivery in the agricultural sector (extension at home,

[3] We added three sub-counties in each of the five treatment groups to account for attrition.
[4] Sample size was now largely determined by the extent to which the OPM implemented the interventions.

measured as the percentage of households in our sample that report they were visited by an expert in the previous year). On average, about 11 per cent of households in our sample report during baseline data collection that they were visited by an extension officer in the year preceding the data collection. The grey solid line shows the power curve associated with the deliberation treatment, comparing the 1,900 households that received the deliberation treatment to the 3,450 households that did not receive a deliberation-focused baraza.[5] The light blue dashed line closely tracks the grey line and shows power for different MDEs for the information component of the baraza intervention. Here, we compare the 2,450 households living in sub-counties that received an information baraza to the 2,900 households that did not receive a sub-county information baraza. The dark blue dashed line compares effectiveness of barazas conducted at different levels with the MDE, defined as the difference in outcome between 1,000 households that received the combined information and deliberation sub-county-level baraza and 2,000 households that were exposed to a district-level baraza. Finally, we also investigate power for the comparison between pure control barazas and the sub-county-level baraza (black dotted line). Here, we compare 1,000 households that received the combined information and deliberation sub-county-level baraza to the 2,000 households that did not receive any baraza. MDEs are estimated using a simple ANCOVA model that controls for the outcome at baseline.

**Figure 4: Power curves for access to extension**

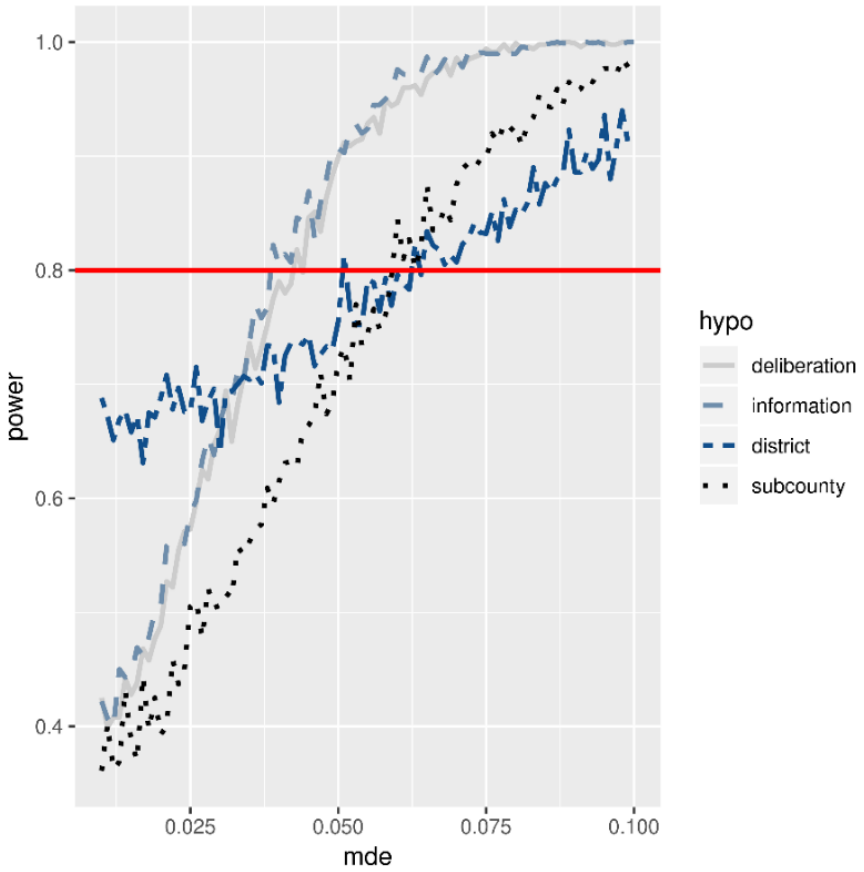



[5] While sample size in treated areas was dictated by what the OPM achieved, we did have some degree of freedom in terms of the sample size in control areas. How the sample size in the control areas was determined is also explained in more detail in section 3.6 on *intervention implementation fidelity*.

Not surprisingly, we have the most power for testing the information treatment. We see that the power curve hits the 80 per cent threshold the first time at an MDE of about 2.5 percentage points. The deliberation experiment is similarly powered and at 80 per cent we can expect to identify effects of 3 percentage points or more. Due to the smaller sample size, comparing sub-county-level barazas to pure control sub-counties seems harder. Here, the difference needs to be at least 4 percentage points. We have the least power when comparing sub-county-level barazas directly to district-level barazas, even though for this comparison, unlike for the previous one, we have the same number of observations in the sub-groups. This is because the unit of randomisation is at a higher level (districts instead of sub-counties).

In *Figure 5*, we plot MDEs for an infrastructure-related outcome: distance in kilometres to the primary water source during the dry season. We find that for the information treatment and the deliberation treatment, we can detect a 4 per cent difference at the standard 80 per cent power level. As the average household lives about 900 metres from the primary water source, this means we can identify effects in excess of 36 metres. Also, the MDE is highest when directly comparing the effect of district-level barazas to sub-county-level barazas. Then, MDEs correspond to about 70–90 metres for the average household in our sample. On GitHub, results are provided for similar power simulations for all the variables used to judge impact of the baraza intervention in section 4.2.2 *below*.

**Figure 5: Power curves for distance to water source**



## 3.4 Ethics

This research was cleared by the Uganda National Council for Science and Technology (UNCST SS 5179) and the Makerere University School of Social Sciences Research Ethics Committee (MAKSS REC 05.19.291), as well as the International Food Policy Research Institute's Institutional Review Board (DSG-19-1053).

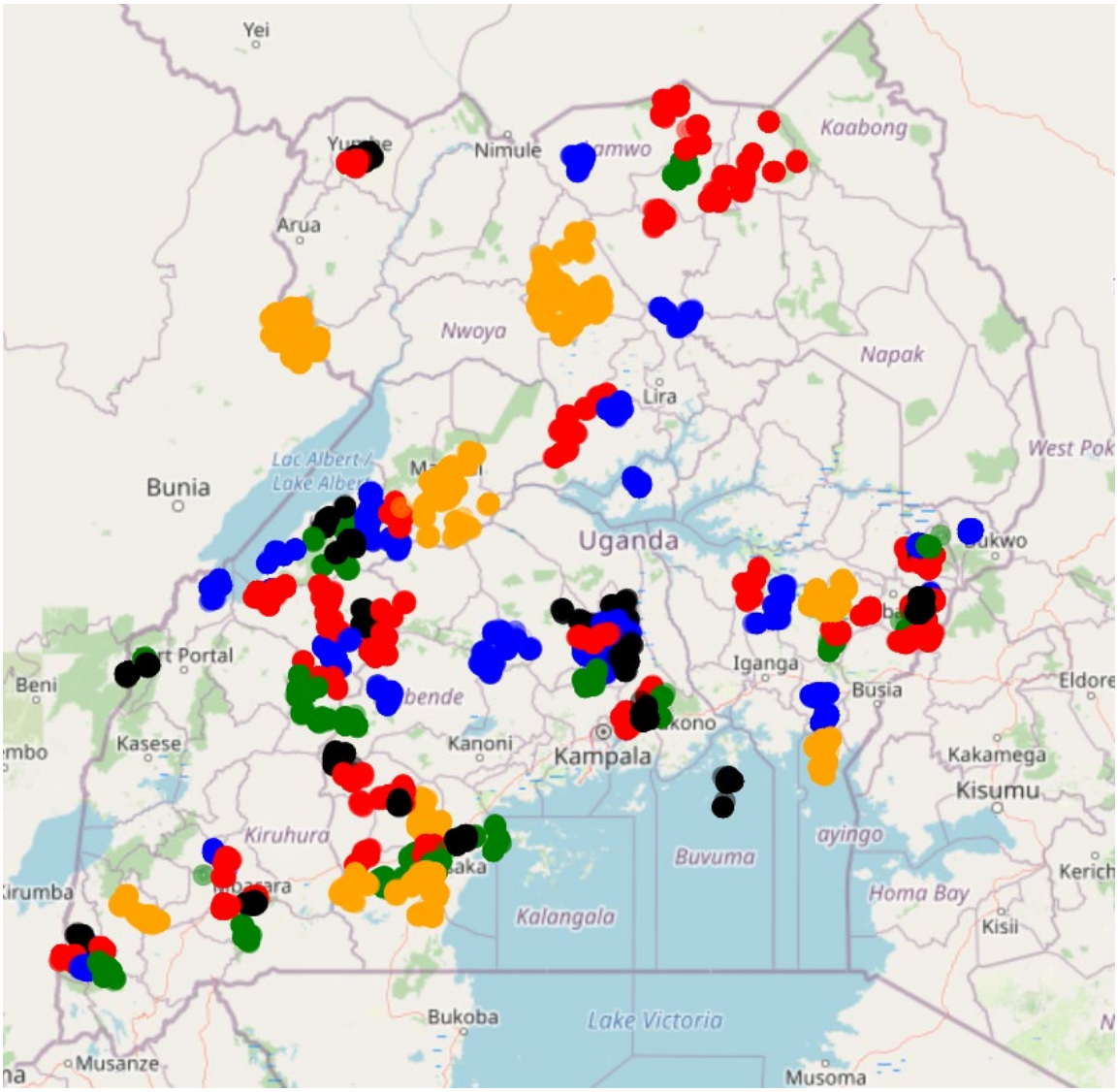
During the implementation, we handled ethical protections carefully and followed protocols and procedures to minimise any potential harm in the field. Throughout baseline and endline interviews, we avoided asking sensitive questions such as those related to religion, political opinions or anything that was considered too personal. Furthermore, we ensured that respondents could refuse to answer questions or abort the interview at any point and this would have no consequences. Interviewees would, for example, still get their token of gratitude if they refused to answer questions or stopped the interview. All field researchers were trained to assure privacy. In addition, we obtained sufficient permission from local authorities in the areas where we worked. Also, during data analysis, we minimised potential harm. While we did collect identifiers to track respondents, data by which respondents could be identified or located were encrypted/anonymised before being analysed or made public.

## 3.5 Sampling and data collection

We designed the experiment to cover districts, sub-counties and households across the four regional blocks (Northern, Western, Central and Eastern) of Uganda. Each regional block has somewhat unique characteristics in terms of ethnicity, and geographical and agro-ecological conditions, as well as cultural history. As noted before, a small share of all sub-counties, though located throughout all of Uganda's 112 districts across the four regions, had already received a sub-county-level baraza intervention. Consequently, we selected our sample of districts from among *eligible districts* and our sample of sub-counties from *eligible sub-counties*. An eligible district was defined as a district where a district-level baraza had not been implemented prior to the start of the study. An eligible sub-county was defined as a sub-county to which two conditions applied: (1) a sub-county-level baraza had not yet taken place; and (2) the sub-county was not located in a district where a district-level baraza had already been implemented. Preliminary analysis of the baraza implementation data at the time of the start of the study indicated that there were 20 or more eligible districts per region, amounting to a total of 94 eligible districts. In each region, there were at least 147 sub-counties that had never been treated and were in eligible districts; the total number of such eligible sub-counties was about 720.

*Figure 6* shows locations of households that were included in the study, clustered in sub-counties (blue, red, black and green) and districts (orange). The colour codes denote the treatment the households were assigned to: (1) blue denotes information-only and green deliberation-only barazas; (2) black are combined deliberation and information sub-county-level barazas; (3) red are control sub-counties; and (4) orange are sub-counties located in districts that received a district-level baraza.

**Figure 6: Study area map**



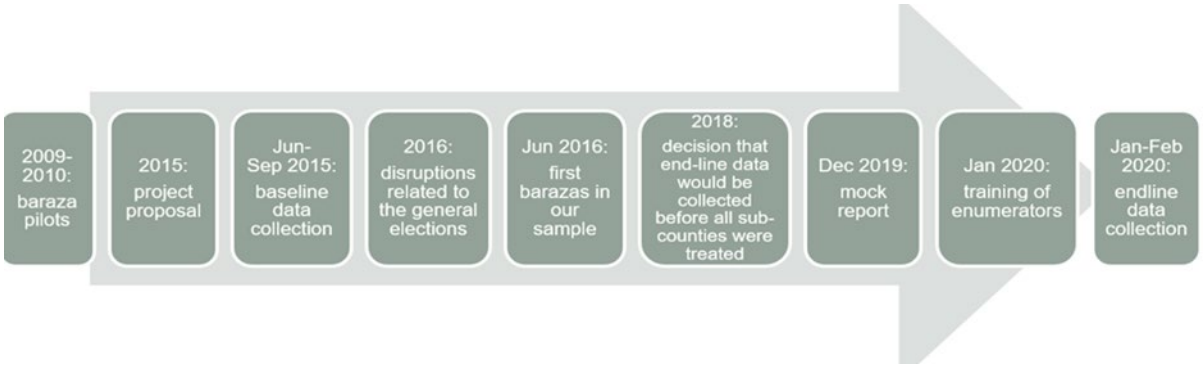

Source: Data collected using Open Data Kit.

We developed and tested the tool in Bagezza sub-county in August 2019. We trained about 80 enumerators during a three-day training in Kampala early in January 2020 and rolled out the survey in the Northern, Western, Central and Eastern regional blocks simultaneously. Progress was tracked on a daily basis using global positioning system (GPS) mapping to trace out best routes and make sure areas were cleared.

The fact that the Open Data Kit application already had many checks built in meant that little data cleaning was needed. Most of the code to run the analysis was ready and, as a result, the first report was ready by 3 March.[6] *Figure 7* provides a timeline.

[6] This first version of the preregistered report can be found here. However, a coding mistake was later found and corrected. For the comparison between full sub-county barazas to control barazas, we were reporting the interaction effect between information and deliberation instead of the combined information and deliberation effect. This was corrected on 2 April in this commit (9f5afdbfdd6be766).

**Figure 7: Timeline**




## 3.6 Intervention implementation fidelity

One of the main challenges was the slow rollout of the baraza intervention by the implementing partner. At the start of 2018, almost two and a half years after baseline data were collected, only about 25 per cent of the planned interventions had happened. We needed to balance the costs and benefits of waiting until the OPM finished all barazas or collecting baseline information after incomplete rollout. At that time, we developed various scenarios, each with an adapted research design. After an additional six months, with still only 56 out of the 155 barazas implemented, it appeared that the best scenario would be to collect endline data before all sub-counties were treated.

However, endline data collection after partial rollout could introduce selection bias. It may be that, from the randomly assigned sub-counties, particular sub-counties were selected to be treated first and others postponed. For instance, for logistical reasons, the implementing partner may have started with sub-counties that are close to the capital. Furthermore, OPM may have treated particular sub-counties first and other sub-counties later for political reasons. Our implementing partner may have selected politically preferred sub-counties first.

There are various ways in which we diagnose and remedy this potential problem. In the next section, we will present a series of balance tests. In particular, we will compare the balance at baseline between subgroups as originally planned and the final sample. We will also look at the balance between households that were supposed to be treated but ended up not receiving treatment and those that were scheduled to be in the control group. This can be done for characteristics at both baseline and endline. We also propose to check robustness of the findings using a matching estimator.

In addition, because only part of the intervention was implemented, it would not be cost-effective to collect endline data on all sub-counties that did not receive a treatment (either because they were allocated to the control or because they ended up not being treated). This raises the following question: from the potential control sub-counties (either those that were allocated to the control or those that ended up not being treated), which control sub-counties should be included in the data collection? One reasonable suggestion would be to pick them randomly. However, if the rollout was not random, such a strategy could lead to a biased estimate of the causal impact of the intervention. For example, it may be that the implementer prioritised sub-counties that were closer to the capital. Randomly selecting control sub-counties might mean that sub-counties

closer to the capital are relatively underrepresented and sub-counties that are further away may be relatively overrepresented in the control group. A better strategy could be to match, ex ante, each treated sub-county to a control sub-county that is similar in a range of observable pre-treatment characteristics the planner had access to when rolling out the intervention and are likely to affect their decision (Kasy 2016; Bertsimas et al. 2015). For example, based on GPS coordinates of a treated sub-county, a control sub-county that is relatively close to the treated sub-county can be selected from the different candidate control sub-counties. This mitigates the bias a planner who prioritises sub-counties in a particular location (e.g. close to Kampala) could introduce.

We decided to use a range of sub-county characteristics that were likely to be known to OPM staff and may have affected how the intervention was rolled out to match each treated sub-county to a control sub-county similar in terms of these characteristics. Namely, we matched on the following characteristics that were obtained at baseline from a survey of village chairpersons and CAOs of each sub-county: GPS coordinates of the sub-county; road infrastructure within the sub-county (kilometres of tarmac road and kilometres of all-weather [gravel] road); share of households with electricity; share of households with an iron roof or tiles; number of health centres in the sub-county; female primary school dropout rate; number of universal primary education (UPE) schools in the sub-county; percentage of farmers who use improved seed; and political connections of the sub-county (defined by having a minister or a member of parliament from that sub-county). These characteristics are used in a probit regression to predict the likelihood that a sub-county was treated. For each treated sub-county, we then match a potential control sub-county with a likelihood of being treated similar to that of the treated one.[7]

In *Table 1*, we look at baseline balance for the resulting sample. The imbalance that was found in *Table A.1* for the information treatment on household size and the number of children in school has disappeared. Consistent with the indication that the OPM may have prioritised treatment of less remote areas (*Table 1*), we now find that distance to the nearest all-weather road is on average slightly higher in control sub-counties. Two (2) significant results out of 40 comparisons is what can be expected from pure chance alone, so we conclude that with this new sample, we maintain the balance between treatment and control on a range of baseline characteristics for the various comparisons.

[7] We use a greedy matching procedure where we first calculate an adjacency matrix for all treatment and control sub-county populations. We then rank all these elements from the matrix and select those that were the closest (in terms of the predicted likelihood of being treated).

**Table 1: Orthogonality tests for final sample**

| | Mean | Sub-county baraza | Information | Deliberation | District baraza |
|---|---|---|---|---|---|
| Household size | 6.411 | -0.186 | 0.065 | -0.302 | 0.062 |
| | (2.855) | (0.169) | (0.152) | (0.166) | (0.248) |
| Age of the household head (years) | 47.009 | 1.096 | -0.215 | 0.574 | 1.554 |
| | (14.542) | (1.012) | (0.731) | (1.038) | (0.998) |
| Head of household is woman (1=yes) | 0.191 | 0.025 | -0.006 | 0.022 | 0.011 |
| | (0.393) | (0.017) | (0.018) | (0.024) | (0.015) |
| Head finished primary education (1=yes) | 0.208 | 0.005 | -0.016 | 0.014 | -0.018 |
| | (0.406) | (0.029) | (0.025) | (0.035) | (0.031) |
| Thatched grass roof (1=yes) | 0.262 | 0.015 | 0.044 | -0.007 | 0.037 |
| | (0.440) | (0.030) | (0.030) | (0.022) | (0.042) |
| Traditional mud wall (1=yes) | 0.444 | 0.086 | 0.031 | 0.062 | -0.008 |
| | (0.497) | (0.058) | (0.053) | (0.058) | (0.114) |
| Distance to the nearest all-weather road (km) | 0.909 | -0.279+ | 0.027 | -0.104 | -0.229 |
| | (0.912) | (0.136) | (0.140) | (0.135) | (0.112) |
| Access to extension (1=yes) | 0.105 | 0.011 | 0.000 | 0.012 | 0.018 |
| | (0.307) | (0.014) | (0.012) | (0.020) | (0.016) |
| Village health team in village (1=yes) | 0.865 | 0.020 | 0.019 | 0.090* | 0.075 |
| | (0.342) | (0.051) | (0.036) | (0.039) | (0.041) |
| Number of children in public schools | 2.507 | -0.089 | 0.001 | -0.188 | 0.078 |
| | (2.072) | (0.118) | (0.097) | (0.111) | (0.154) |
| Number of observations | 7,340 | 2,949 | 5,298 | 5,298 | 3,999 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels, respectively.

While the ex-ante matching strategy may reduce bias resulting from incomplete rollout, there are also costs involved. First, if sample selection is introduced by the rollout, matching may further reduce external validity of the study, because the control sub-counties are not a random sample of the study population any more. Second, the reduction in potential bias for hypotheses related to the sub-county-level barazas should be traded off against an increase in potential bias when testing differences between control and district-level barazas. Because the sub-county-level analysis weighed higher in terms of research objectives, we decided to prioritise the reduction of bias resulting from incomplete rollout at this level. However, it should be kept in mind that both of these

issues only become relevant if significant selection bias was introduced through the partial rollout. Section 6.4 provides a more detailed discussion of the limitations of matching methods.

*Figure 8* summarises the factorial design that underlies the assessment of the relative effectiveness of the information and deliberation components of sub-county-level barazas. As already noted in section 3.2, one of the main advantages of factorial designs (as opposed to parallel designs) is that to test main effects all observations can be used. For instance, to test the impact of an information baraza, we can compare outcomes of households in sub-counties that received the information treatment (either the information only or the information and deliberation treatments) to outcomes of households that did not receive the information treatment (either because they received no treatment at all or because they only got the deliberation treatment). Had the intervention been implemented as planned, we would have followed the original power calculations and had 104 information sub-counties that could be compared to 102 control sub-counties (and given that 50 households were interviewed in each sub-county, we would have 5,200 treated households and 5,100 control households).

**Figure 8: Factorial design**

| | Control | Information |
|---|---|---|
| Control | Planned: 51<br>Included: 40 | Planned: 51<br>Treated: 29 |
| Deliberation | Planned: 51<br>Treated: 18 | Planned: 53<br>Treated: 20 |

However, the incomplete rollout resulted in the fact that only 67 of a total of 155 sub-counties that would have received any treatment were actually treated. Referring to *Figure 8*, we see that to test the impact of the information baraza, 49 sub-counties that were treated can be used. This means that a total of 157 sub-counties that did not receive the information treatment can be used as control sub-counties. However, optimal power is obtained in designs where the number of treated units is about equal to the number of control units. Thus, with the aim for being cost-effective, we collected information on 49 sub-counties. Because we wanted to formally test whether the partial rollout introduced selection bias by comparing planned control sub-counties to sub-counties that were not treated using endline data (see section 4, *Findings*), we made sure we selected half of these from the first and the other half from the second column in *Figure 8*. To test the impact of the deliberation treatment, we needed 38 control households. Here, too, we made sure half were from the planned controls (the first row in *Figure 8*) and half from sub-counties that were supposed to be treated but were not (the second row). Finally, because we also planned to directly test for the effect of a combined information and deliberation treatment, we needed at least 20 pure control sub-counties. We again made sure half were selected from the upper left cell in *Figure 8* and half from the sub-counties that were assigned to the treatment in the lower right cell of *Figure 8* but did not get the treatment. Note that often the same sub-county could be used to test different hypotheses. For instance, the 10 sub-counties in the upper left cell needed to test whether the deliberation intervention was effective could be taken from the 14 sub-counties that were needed in that cell to test the impact of the information

treatment. Therefore, we simply took the higher number in each cell, which was 14 sub-counties. To allow for attrition, we selected 16 control sub-counties in each treatment cell.

In practice, we started by matching 10 untreated sub-counties from the $S_0^0$ group to the treated sub-counties in the $S_{ID}^0$ group. We then matched a further 10 sub-counties from the $S_{ID}^0$ group that ended up not being treated to the treated sub-counties in the $S_{ID}^0$ group. Next, we looked at the information treatment. In this treatment, 49 sub-counties were treated, either as information-only or as part of the combined information and deliberation treatment. This means we also needed 49 controls. In the previous step, we had already selected 20 pure controls that we could use. Furthermore, 18 pure deliberation treatments could also be used as controls for the information treatment. This means we needed an additional 11 controls. Because we wanted to investigate balance between control and planned but not treated controls, we selected these 11 controls from the sub-counties that were planned to receive the information treatment $S_I^0$, but ended up not receiving it.

Finally, we looked at the deliberation treatment. In this treatment, 38 sub-counties were treated, either as deliberation-only or as part of the combined information and deliberation treatment, so we also needed 38 controls. We already had the 20 pure controls and an additional 11 controls from the previous steps; we needed an additional 7 controls. Because we wanted to investigate the balance between control and planned but not treated controls, we selected these seven controls from the sub-counties that were planned to receive the deliberation treatment $S_D^0$, but ended up not receiving it.

# 4. Findings

## 4.1 Research analyses

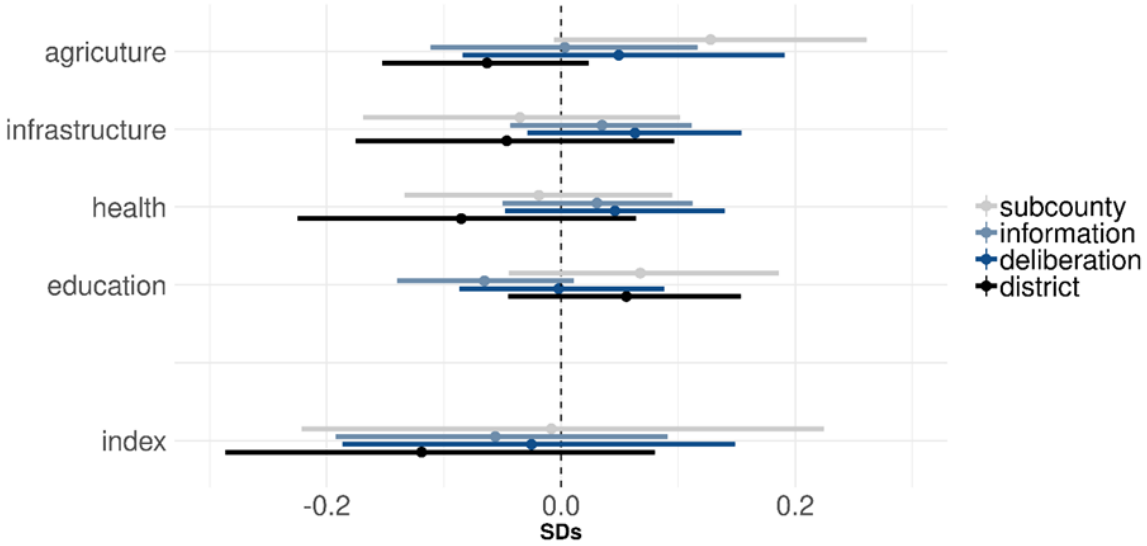
In this section, we provide results for the four main hypotheses outlined in section 3.1. The main results are presented in two parts. In the first part, we strictly follow a preregistered analysis plan and focus on a subset of carefully selected and declared variables that are combined in indices: one overall index and four indices corresponding to the four key sectors (agriculture, infrastructure, health and education). This confirmatory part of the paper will allow us to assess the overall impact of the baraza intervention.

The second part of the analysis is more exploratory in nature and looks at individual outcomes. In this part, we do not follow a preregistered analysis plan. For each of the sectors, a set of outcomes is compared between the different groups using ANCOVA models. We also explore whether the baraza programme affected various aspects that are at the core of community-based monitoring, such as participation in election of local leaders, interfacing with politicians and civil servants, perceptions of service quality and prioritisation, and contributions to public goods (both cash and in-kind). Finally, in this section, we also report results for the analysis of sub-county-level data that were collected from government officials.

### *4.1.1 Confirmatory analysis*

*Figure 9* provides a graphical representation of the overall impact on service delivery of the barazas, as well as on the different sectors. It shows the impact of the four main hypotheses: (1) the impact of the sub-county barazas (subcounty, indicated in grey); (2) the relative effectiveness of the information component (information, light blue); (3) the relative effectiveness of the deliberation component (deliberation, dark blue); and (4) a comparison between sub-county- and district-level barazas (district, black) on the four sectors we consider (agriculture, infrastructure, health and education). The graphs are based on indices that are composed of individual outcomes in each sector as described in our preregistered analysis plan, which are discussed in detail in the next section. We also combine the four indices into one overall index that assesses the impact on public service delivery in general.

**Figure 9: Summary of baraza impact**




The figure provides point estimates for the difference between treatment and control, estimated in an ANCOVA framework with controls for baseline outcome and region dummies. Confidence intervals are obtained following the permutation method explained in Gerber and Green (2012). This method first reconstructs a complete schedule of potential outcomes by adding and subtracting the average treatment effect for control and treated units, respectively. These potential outcomes are then used to simulate all possible random allocations. Average treatment effects are estimated for each allocation and the 2.5th and 97.5th percentiles are then taken as the lower and upper limits of the 95 per cent confidence interval. This provides a conservative estimate of the confidence interval.

We find no significant impact of the baraza programme on overall public service delivery. There are some indications that sub-county-level baraza did make a difference in the agricultural sector, but the difference is only significant at the 10 per cent level. We do find that public service delivery in the agricultural sector was significantly worse in areas that were exposed to district-level barazas than in areas exposed to sub-county-level barazas.

### *4.1.2 Detailed analysis*

The indices combine various outcomes and for some of them the expected direction of the effect is a priori unclear. For instance, an information baraza may increase the quality of services in a hospital or health centre when judged by an objective measure, such as waiting time. However, the information may also result in higher expectations from users. Thus, perceptions of quality may have reduced as a result of an information baraza. It is, therefore, also interesting to look beyond the indices and consider outcomes individually. This part of the analysis is more exploratory in nature and does not follow a preregistered analysis plan.

### *4.1.3 Agriculture*

We first zoom in on the outcomes that are used to assess the effectiveness of barazas in changing service delivery in agriculture. Results are reported in *Table 2*.

**Table 2: Impact of barazas on agricultural outcomes**

| | Mean | Sub-county baraza | Information | Deliberation | District baraza |
|---|---|---|---|---|---|
| Household used inorganic fertilisers?† | 0.229 | -0.015 | 0.034 | 0.001 | -0.013 |
| | (0.42) | (0.033) | (0.035) | (0.049) | (0.031) |
| Household used improved seed?† | 0.364 | 0.043 | -0.03 | -0.037 | -0.043 |
| | (0.481) | (0.033) | (0.038) | (0.038) | (0.034) |
| Received improved seeds from govt? | 0.121 | 0.051* | 0.004 | 0.056 | -0.005 |
| | (0.326) | (0.024) | (0.025) | (0.043) | (0.015) |
| Household used agro-chemicals? | 0.469 | -0.028 | -0.007 | -0.005 | -0.010 |
| | (0.499) | (0.05) | (0.035) | (0.046) | (0.043) |
| Household used improved livestock inputs? | 0.221 | 0.029 | 0.021 | 0.03 | -0.014 |
| | (0.415) | (0.031) | (0.028) | (0.034) | (0.026) |
| Did an agricultural expert visit your home?† | 0.178 | 0.056** | 0.037 | 0.036 | -0.027 |
| | (0.383) | (0.018) | (0.03) | (0.048) | (0.014) |
| Visited extension office/demo site/model farm?† | 0.285 | 0.040 | 0.036 | 0.045 | -0.013 |
| | (0.452) | (0.028) | (0.035) | (0.044) | (0.023) |
| Are officials aware of extension demand? | 0.264 | -0.006 | 0.017 | -0.001 | -0.075* |
| | (0.441) | (0.024) | (0.027) | (0.035) | (0.023) |
| Not consulted for extension content? | 0.316 | 0.034 | -0.041 | -0.031 | -0.056 |
| | (0.465) | (0.033) | (0.027) | (0.034) | (0.033) |
| Are farmer associations/groups in this village? | 0.403 | 0.06+ | 0.04 | 0.087* | -0.032 |
| | (0.491) | (0.03) | (0.038) | (0.041) | (0.027) |
| Farmer groups supported by govt?† | 0.173 | 0.070* | -0.015 | 0.053 | -0.037 |
| | (0.378) | (0.028) | (0.03) | (0.04) | (0.022) |
| Received help in marketing from govt?† | 0.069 | 0.018 | -0.013 | 0.016 | -0.014 |
| | (0.254) | (0.022) | (0.017) | (0.017) | (0.012) |
| Received help in marketing from coop?† | 0.062 | 0.037 | -0.021 | -0.001 | -0.009 |
| | (0.241) | (0.024) | (0.016) | (0.021) | (0.014) |
| Number of observations | 6,704 | 2,738 | 4,858 | 4,858 | 3,687 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; † indicates that the outcome was included in index.

We start by looking at whether the baraza programme affected the use of modern inputs in agriculture. The first outcome considers whether the household used inorganic fertilisers (e.g. diammonium phosphate, urea, nitrogen, phosphorus and potassium, foliar, triple superphosphate, single superphosphate, muriate of potash) during the past 12 months. The first column reports baseline averages, with standard deviation in brackets below. About 23 per cent of households in the sample answer this question affirmatively. In the second column, we report differences in outcomes between households that received a typical sub-county-level baraza (i.e. the crossed treatment of a sub-county information baraza and a sub-county deliberation baraza; the bottom right in *Figure 8*) and households that did not receive any baraza (pure control; the top left in *Figure 8*). We see that the proportion of households that report using inorganic fertilisers 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 those that did not receive any baraza (second column). 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 organised (either an information-only baraza or a combined information and deliberation baraza; top and bottom right of *Figure 8*) and outcomes of households that live in areas not exposed to an information baraza (either pure control or deliberation-only baraza; top and bottom left of *Figure 8*). We see that adoption of inorganic fertilisers is 3.4 percentage points higher among households that were exposed to an information baraza. However, the difference is again not significant. In the fourth column, we report differences between outcomes of households that live in areas where a deliberation baraza was organised (either deliberation-only baraza or a combined information and deliberation baraza; bottom left and right of *Figure 8*) and outcomes of households that live in areas not exposed to a deliberation baraza (either pure control or information-only baraza; top left and right of *Figure 8*). We also do not find differences in terms of inorganic fertiliser 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. This input seems to be used more widely than inorganic fertilisers: 36 per cent of households report that they have used improved seed during the past year. The reported use is 4.3 percentage points higher among households that reside in areas where a sub-county-level baraza took place than in areas where no baraza was conducted, but the difference is not significant. We find negative point estimates for the relative effects of both the information and the deliberation component, but effects are imprecisely estimated. Finally, the adoption of improved seed was lower in areas where a district-level baraza was conducted, but the difference compared to the areas that did not receive a baraza is not significant. Adoption of improved seed and inorganic fertilisers was included in the index that was used for the confirmatory analysis.

Next, about 12 per cent of households report that they received improved seed from the government extension system (i.e. through an extension agent from the National Agricultural Advisory Services [NAADS] or through Operation Wealth Creation, which replace NAADS). This is 5.1 percentage points higher in areas where a sub-county-level baraza took place and this difference is significant at the 5 per cent level. We find no

such effects from district-level barazas. Direct comparison of district- and sub-county-level barazas indicates that sub-county-level barazas are significantly more effective in increasing the likelihood of households reporting they received these inputs from the government.

We then check whether households report any changes with respect to the use of agro-chemicals. This includes the use of pesticides, herbicides, fungicides and acaricides during the past 12 months. Overall, almost half of all households in the sample report using some form of agro-chemical. We do not find evidence that the baraza intervention affected the use of this input. Finally, we consider the use of modern inputs and methods in livestock rearing over the past 12 months, including improved animal breeds and use of modern feeds, drugs and artificial insemination: 22 per cent of households report using such inputs and this proportion is similar across different experimental groups.

We then turn to advisory services. We first investigate whether the barazas have affected access to extension at home. We estimate the percentage of households in our sample that report they were visited by an expert (e.g. crop or livestock extension agent, or community-based facilitator or other experienced farmer) at the home in the past 12 months. Access to extension is low: only about 18 per cent of households report that they received such a visit. Interestingly, we find that this percentage is significantly higher among households that were affected by a sub-county-level baraza. The effect is large, amounting to a 30 per cent increase over the sample mean. Furthermore, the effect seems to come from a combination of the information and deliberation components; the components alone do not seem to affect the outcome enough to render it significant. We also find that this effect is absent among households that live in sub-counties that received a district-level baraza. A direct comparison of access to 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. Comparing realised effects with MDEs indicates that we have sufficient power.

Home visits by extension officers are not the only means for households to 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 farmers can consult when the need arises became more important than the more supply-driven component of training and visits. We thus also inquire whether anyone in the household visited an extension office, demonstration site or model farmer in the past year. About 28 per cent of households in our sample report access to extension in this modality. 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 in the agriculture index.

Three quarters of households in our sample mention that there are agricultural enterprises, improved technologies or inputs they would like to adopt, indicating significant scope for advisory services. We also find that, according to respondents, service providers and policymakers are not always aware of this demand. *Table 2* shows that only 26 per cent of households believe that officials are 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 perceive that a district-level baraza reduced this percentage. Apparently, a district-level baraza makes the mismatch between what farmers need and what officials think farmers need more salient.

With regards to the previous outcome, we ask how decisions related to what topics to cover in agricultural extension are made. We define this outcome in a negative way – the indicator is true if decisions are made without consultation. We see that about 30 per cent of households indicate that no consultation happens and the content of extension advisory services is decided on by experts at the central level. We do not find that the baraza intervention increased participation in extension service planning.

About 40 per cent of households report the presence of farmer groups or cooperatives in their village. In the agricultural sector in Uganda, such groups are very important. The government actively promotes them. In fact, to be able to receive inputs from the government, farmers need to be a member of such a group. Sub-county-level barazas increase the likelihood that farmer cooperatives or groups are formed in the villages in Uganda. Interestingly, it seems that the deliberative component is the main driver behind this result. This effect is specific to interventions at the sub-county level. A higher share of farmer groups in areas that received a sub-county-level baraza received support from government.

The final two questions focus more on marketing. Connecting farmers to markets is also an important strategy outlined in the Agriculture Sector Strategic Plan. The first outcome relates to the likelihood that the government supports farmers through the village procurement committee. In the sample, about 7 per cent of households report that the government assisted them. The second question is similar, but looks at the role of cooperatives. We generally find no effects of the barazas, except perhaps for an increase of almost 4 percentage points in the likelihood that cooperatives assist with marketing in areas that received a sub-county-level baraza. Both of these outcomes were also included in the index to assess the overall impact.

#### *4.1.4 Infrastructure*

Another important area where we expect to see an impact of the baraza programme is in infrastructure, with the primary focus on drinking water infrastructure. Results, similarly formatted as results in the previous section, are in *Table 3*.

The first outcome we consider is whether the household uses an unprotected water source during the dry season. This is measured as the share of households that report that the main source of drinking water during the dry season is surface water, an unprotected dug well or an unprotected spring. About 16 per cent of households in the sample report that they use an unprotected water source. The baraza intervention does not seem to affect this proportion. This outcome is included in the infrastructure index.

The second outcome we look at, which is also included in the index, is the distance to the primary water source during the dry season. This was measured in kilometres, but trimmed and transformed using inverse hyperbolic transformation. On average, households have to walk about 1 kilometre. While this distance seems to reduce in all comparisons – especially for barazas held at the district level, where we find a reduction of approximately 9 per cent – it is never significantly different from zero.

The third outcome, also part of the index, is the time one must wait at the water source, measured in minutes. This continuous variable was also trimmed and transformed. Households must wait on average about 37 minutes. There is a significant reduction in waiting time in areas that were exposed to the sub-county-level baraza intervention and some indication that the deliberation component is mostly responsible for this reduction.

**Table 3: Impact of barazas on infrastructure**

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | District baraza |
|---|---|---|---|---|---|
| Household uses unprotected water source?† | 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 | -0.04 | -0.049 | -0.091 |
| | (0.576) | (0.046) | (0.041) | (0.061) | (0.039) |
| 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) |
| Is there a water user committee in the village?† | 0.598 | -0.021 | 0.033 | 0.032 | -0.009 |
| | (0.49) | (0.046) | (0.037) | (0.04) | (0.047) |
| Is member of water user committee? | 0.163 | 0.022 | 0.001 | 0.04 | 0.020 |
| | (0.37) | (0.021) | (0.017) | (0.025) | (0.038) |
| Water user committee holds public meetings? | 0.474 | -0.005 | 0.043 | 0.060 | -0.050 |
| | (0.499) | (0.044) | (0.036) | (0.042) | (0.056) |
| Satisfied with quality of drinking water? | 0.624 | 0.031 | -0.009 | -0.062 | 0.002 |
| | (0.484) | (0.052) | (0.044) | (0.044) | (0.019) |
| Treat water before drinking (boil or treat)? | 0.5 | -0.025 | -0.087* | -0.02 | 0.010 |
| | (0.5) | (0.045) | (0.037) | (0.046) | (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 | 6,704 | 2,738 | 4,858 | 4858 | 3,687 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; † indicates that outcome was included in index.

The fourth outcome variable assesses changes in the presence of a water user committee in the village. Overall, about 60 per cent of households report that such a committee is present in their village. This share does not vary between the different experimental groups. Similarly, households are not more or less likely to participate in such committees and the committees do not hold more or fewer public meetings.

Households were also asked whether they were satisfied with the quality of the water that was available at the source during the dry season. About 62 per cent responded they were satisfied or very satisfied with the drinking water. Households that were exposed to the baraza intervention were not more or less likely to report they were (very)

satisfied with the quality of drinking water during the dry season. Half of the households reported they treated drinking water before drinking it, either by boiling it or treating it with chlorine. The likelihood that households treat water reduces somewhat for the information treatment. Better access to clean water potentially reduces the necessity to treat it before drinking.

We included one question related to road infrastructure. We asked how far the household was located from the nearest all-weather road. In the full sample, households live on average 26 kilometres from a road. The baraza programme did not reduce the distance to the nearest all-weather road.

#### *4.1.5 Health*

We now look at outcomes in the health sector (*Table 4*). One problem with public health-related outcomes is that some will only be available for households that have visited government health facilities, reducing the sample size too much to maintain acceptable power.

The first two outcomes we consider attempt to assess changes in access to or use of public health facilities. The first indicator measures the use of public health facilities for illness. In particular, we construct an indicator that is true if the household head responds that treatment would be sought in a health centre 2, 3, 4 or in a regional referral hospital if a member of their household had fever. We find that 70 per cent of households respond that they would seek treatment in a government health facility. This proportion is valid for all treatment groups.

A similar indicator attempts to assess the use of the public health system for maternal healthcare and asks whether treatment would be sought in a health centre 2, 3, 4 or in a regional referral hospital if a member of the household was to give birth. This percentage is even higher than for illness: more than 80 per cent would go to a government health facility to give birth. Again, this proportion is not affected by the baraza programme. Both outcomes are included in the health index.

Next, we ask whether a village health team (VHT) is present in the village. VHTs are very important in front-line healthcare in Uganda. They also have prominent roles in government health interventions, such as immunisation campaigns or distribution of bed nets. Overall, nearly 90 per cent of households report a VHT is present in their village. The presence of a VHT is not affected by the baraza intervention.

As barazas try to increase citizens' engagement, we also check whether households that were exposed to a baraza were more likely to participate in VHTs. We asked whether any member of the household was a member of a VHT. In about 10 per cent of our sample, at least one household member is part of a VHT. The baraza intervention does not increase the likelihood that individuals participate 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 organise more public meetings. Overall, 43 per cent of households state that VHTs have organised a public meeting in the past year. This proportion is significantly higher in areas that were exposed to a sub-county-level baraza. This effect seems driven by the deliberative component of a sub-county baraza.

We also consider distance to the nearest government health facility, measured in kilometres. Overall, average distance to the nearest government health facility is almost 50 kilometres. We do not find that barazas reduce this distance.

We then turn to health outcomes. We start by asking whether any member of the household was sick during the past year. This was the case in two thirds of the households in our sample. The intervention did not reduce morbidity in our sample. We then ask for each sick person in the household to record how many days they were ill and use this to calculate the total number of sick days at the household level in the past year. The average household recorded almost 50 sick days according to this definition. We do not find that the intervention affected the (trimmed and transformed) number of sick days. Finally, we look at the number of days household members were unable to go to school or work, which provides an indication of severity of illness. Calculated similarly to the previous outcome, in an average household, about 35 school- or workdays are missed due to illness. Again, there is no significant reduction in this (trimmed and transformed) number. This last health outcome measure was included in the health index.

**Table 4: Impact of barazas on health sector**

| | Mean | Sub-county baraza | Information | Deliberation | District baraza |
|---|---|---|---|---|---|
| Seek treatment for fever in public health facility?† | 0.691 | -0.008 | -0.007 | 0.025 | -0.010 |
| | (0.462) | (0.033) | (0.033) | (0.040) | (0.046) |
| Go to public health facility to give birth?† | 0.813 | -0.029 | -0.033 | -0.016 | -0.070 |
| | (0.390) | (0.034) | (0.029) | (0.035) | (0.043) |
| Is there a VHT in village?† | 0.881 | 0.022 | 0.005 | 0.029 | -0.019 |
| | (0.323) | (0.031) | (0.025) | (0.025) | (0.027) |
| Member of VHT? | 0.113 | 0.022 | 0.003 | -0.001 | -0.024 |
| | (0.317) | (0.017) | (0.014) | (0.015) | (0.012) |
| VHT organises any public meetings? | 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) |
| Any members sick? | 0.658 | 0.003 | 0.024 | 0.037 | -0.015 |
| | (0.475) | (0.023) | (0.028) | (0.033) | (0.024) |
| Number of days ill? | 2.576 | -0.005 | -0.04 | 0.004 | -0.064 |
| | (2.189) | (0.091) | (0.149) | (0.166) | (0.105) |
| Number of days of school/work missed due to illness?† | 2.273 | -0.081 | 0.076 | -0.006 | -0.065 |
| | (2.027) | (0.106) | (0.134) | (0.145) | (0.121) |
| Waiting time before being attended (min.)?† | 4.744 | -0.04 | -0.133 | -0.151 | 0.064 |
| | (1.012) | (0.093) | (0.108) | (0.135) | (0.082) |
| Has visited traditional health | 0.257 | -0.017 | 0.016 | 0.034 | -0.039 |

| | Mean | Sub-county baraza | Information | Deliberation | District baraza |
|---|---|---|---|---|---|
| practitioner?† | | | | | |
| | (0.437) | (0.032) | (0.029) | (0.03) | (0.019) |
| Patient was examined by in-charge/doctor? | 0.411 | 0.044 | -0.049 | -0.070 | -0.041 |
| | (0.492) | (0.041) | (0.032) | (0.042) | (0.025) |
| Time of examination? | 3.403 | 0.048 | -0.099 | 0.015 | -0.002 |
| | (0.761) | (0.066) | (0.070) | (0.091) | (0.083) |
| Paid anything? | 0.179 | 0.01 | -0.008 | -0.013 | -0.005 |
| | (0.384) | (0.023) | (0.025) | (0.042) | (0.024) |
| Received meds in hospital? | 0.709 | 0.000 | -0.003 | 0.000 | -0.024 |
| | (0.454) | (0.036) | (0.023) | (0.033) | (0.034) |
| Satisfied with services at hospital? | 0.682 | 0.048 | -0.026 | -0.038 | -0.011 |
| | (0.466) | (0.033) | (0.031) | (0.038) | (0.026) |
| Number of observations | 6,704 | 2,738 | 4,858 | 4,858 | 3,687 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; Column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; † indicates that outcome was included in index.

We then ask how long one had to wait before being attended to (in minutes). The sample mean for this outcome is about 90 minutes. While waiting time reduces for most comparisons, the differences are never significant. Potentially, the reduced sample size resulted in too little power to detect a difference. This outcome was also included in the health index.

A final question that was included in the index was again asked to all households. In particular, we inquired whether a traditional health practitioner was consulted in the past 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. To assess this, we ask who examined the patient in the health centre. Ideally this should be a doctor or in-charge. If such a person is absent, patients are 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 40 per cent of cases does a qualified person appear to have done the examination. The baraza intervention does not seem to lead to less absenteeism. We also look at the time the examination took. The average examination in our sample took about 22 minutes. There is no change related to the intervention.

Healthcare in Uganda is supposed to be free. However, corruption is widespread and often patients are required to make payments to receive care. Almost 20 per cent of households report that payment was required the last time they visited a government health facility. The intervention had no impact. Relatedly, users often complained about a lack of drugs in government health facilities. We asked whether, during the last visit to a government health centre, drugs were received (indicating that drugs were available). We also asked whether drugs had to be purchased from outside of the hospital (indicating that at least some drugs were missing). While 70 per cent of households report they received medicines in the health centre, almost all of them also mentioned they had to buy drugs outside of the hospital. The intervention did not seem to make a difference for either indicator.

We further probe for a subjective assessment of the overall quality of care at the health facility. Most households report that they are satisfied or very satisfied with services received at the government health facility. This seems 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 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 immunisation. However, we already find close to 100 per cent immunisation rates in our baseline data. Another outcome is child mortality. Child mortality rates at baseline were estimated at 38 per 1,000 live births, which was deemed too low to include in the endline analysis. Raffler and colleagues (2018) find similar child mortality rates at baseline and speculate that the fact 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 1,000 live births.

#### *4.1.6 Education*

Education outcomes to assess the impact of the intervention suffer from a similar problem as health outcomes: not all households in the sample have children in school, so for many of the outcomes related to education the sample size is small. This also affects the indices. Results are presented in *Table 5*.

**Table 5: Impact of barazas on education**

| | Mean | Sub-county baraza | Information | Deliberation | District baraza |
|---|---|---|---|---|---|
| Number of children in UPE 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)?† | 1.42 | 0.025 | -0.047 | -0.044 | -0.002 |
| | (0.763) | (0.057) | (0.067) | (0.071) | (0.042) |
| Has complete boundary fence?† | 0.347 | 0.064 | -0.061 | -0.057 | -0.008 |
| | (0.476) | (0.048) | (0.046) | (0.049) | (0.045) |
| Has electricity? | 0.338 | 0.165** | -0.04 | -0.017 | 0.035 |
| | (0.473) | (0.049) | (0.042) | (0.049) | (0.038) |
| Has water facility?† | 0.703 | 0.106* | -0.023 | 0.026 | 0.073 |
| | (0.457) | (0.041) | (0.048) | (0.05) | (0.050) |
| Has PTA? | 0.945 | -0.007 | -0.029 | 0.000 | 0.000 |
| | (0.227) | (0.014) | (0.019) | (0.028) | (0.012) |
| Has SMC?† | 0.915 | 0.008 | -0.034 | 0.002 | 0.037 |
| | (0.279) | (0.024) | (0.023) | (0.033) | (0.020) |
| Informed about SMC?† | 0.882 | 0.021 | -0.036 | -0.042 | 0.009 |
| | (0.323) | (0.023) | (0.024) | (0.032) | (0.019) |
| Inspectors visited schools?† | 0.639 | -0.004 | -0.075+ | -0.035 | 0.015 |
| | (0.48) | (0.051) | (0.043) | (0.048) | (0.036) |
| Number of observations | 6,704 | 2,738 | 4,858 | 4,858 | 3,687 |

**Note:** UPE = universal primary education; USE = universal secondary education; PTA = parent-teacher association; SMC = school management committee; column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; † indicates that outcome was included in index.

If the quality of public education is poor, households are less likely to send their children to public schools. Thus, the first obvious outcome is to simply compare the number of children within the households that attend public school (either universal primary education or universal secondary education). The average household in our sample had almost two children in government schools, but enrolment rates were not affected by the baraza intervention.

Access to public education is also influenced by the distance to a public school, so we recorded distance to primary or secondary school (or the average if both are reported). On average, households live about 3 kilometres from a government-operated school. The baraza programme did not have an impact on this outcome either.

We also look at school infrastructure. First, we ask households whether 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 per cent of schools had such a

fence.[8] We also ask whether the school had electricity and whether there was a water source available in the school. Overall, about 34 per cent of schools had electricity and about 70 per cent had a water source. Sub-county-level barazas seem to have improved 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 suggest there were generally no issues related to these outcomes.

We also look at how the school was managed and how parents were involved. For instance, we consider whether the school had a parent teacher association (PTA) and a school management committee (SMC). Almost all schools had a PTA and 91 per cent of households stated the primary or secondary school attended by any of their children had an SMC. However, not all households were informed about SMC meetings. The baraza intervention does not seem to affect how schools are managed, parents can participate or information is shared. Finally, we ask households whether an inspector had visited the school in the year before the survey. About 64 per cent of households indicate that schools had been inspected. Surprisingly, this proportion reduces as a result of the information component of a baraza.

#### *4.1.7 Contact with policymakers and service providers*

As mentioned, one of the main aims of the community forums is to increase communication between politicians, civil servants and citizens. Therefore, we try to assess whether citizens interact more with politicians and service providers as a result of the meetings. In particular, we asked how long it had been since the respondent spoke personally to various officials for reasons related to service provision in agriculture, health, education, water or roads. Based on the answer, we constructed an indicator variable that denotes whether the household had a meeting or not. The time frame changes depending on the official. For instance, for the village local council chairperson (LC1), the indicator takes the value of one if the respondent spoke to them within the past month. For the head teacher, the reference period is six months. For other officials (sub-county chief, health management unit member and water committee member), the indicator is true if contact was sought in the past year. Results are presented in *Table 6*.

About 43 per cent of households in our sample had met with the (village) LC1 chair in the month before the endline data were collected. The baraza intervention did not affect the likelihood that citizens met with the LC1 using this definition. About 20 per cent of respondents reported they had met with the sub-county chief in the past year. We do not find that the baraza intervention changed this likelihood.

Furthermore, we see that the information component of the baraza increases the likelihood that citizens interface with the head teacher or members of the SMC. Also, the information component of the sub-county-level baraza increases the likelihood of meetings with water committee members. Finally, and similarly to meetings with sub-county chiefs, few citizens reported meeting with health unit management committee members, but all coefficients on sub-county-level interventions are positive. The index also shows that the information component had the largest effect on meetings, but the effect is not significant.

[8] The lack of a fence was a frequent complaint from parents during qualitative work.

**Table 6: Impact of barazas on meetings**

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | District baraza |
|---|---|---|---|---|---|
| LC1 chairperson | 0.426 | 0.001 | 0.030 | 0.035 | -0.034 |
| | (0.495) | (0.025) | (0.035) | (0.048) | (0.024) |
| Sub-county chief | 0.196 | 0.031 | 0.035 | 0.053 | -0.040 |
| | (0.397) | (0.02) | (0.035) | (0.052) | (0.020) |
| Head teacher/SMC member | 0.486 | 0.038 | 0.058* | 0.048 | 0.019 |
| | (0.5) | (0.028) | (0.028) | (0.037) | (0.024) |
| Health unit management committee member | 0.155 | 0.040 | 0.020 | 0.061 | -0.021 |
| | (0.362) | (0.024) | (0.036) | (0.051) | (0.017) |
| Water committee member | 0.382 | -0.016 | 0.060+ | 0.044 | 0.012 |
| | (0.486) | (0.040) | (0.034) | (0.051) | (0.037) |
| Contact index | 0.000 | 0.037 | 0.089 | 0.107 | -0.036 |
| | (0.649) | (0.036) | (0.059) | (0.094) | (0.033) |
| Number of observations | 6,704 | 2,738 | 4,858 | 4,858 | 3,687 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels, respectively.

### *4.1.8 Participation in elections*

A second key aim of the baraza programme is to increase citizens' empowerment. One way citizens can influence policy is through political participation. Thus, we expect that the baraza intervention would affect the likelihood that citizens participate in elections at various levels. We also ask whether any of the household members hold any political or traditional position. Results are in *Table A.3* in the Appendix. Results show fairly high overall participation in elections at various levels and no impact of barazas. About 30 per cent of households report that at least one member holds a political or traditional position.

### *4.1.9 Cash and in-kind contributions*

The baraza programme also attempts to increase a sense of community engagement. One way citizens can participate is by contributing to common goods, such as public infrastructure, education or health services. We differentiate between cash and in-kind contributions.

*Table 7* shows that about 32 per cent of households indicate they made in-kind contributions to public schools in their community in the past two years. Overall, most in-kind contributions were targeted towards drinking water facilities and the least in-kind contributions were going to a dam or irrigation facility, which is consistent with the difference in public nature of these two facilities. Cash contributions are distributed similarly, except for the fact that contributions to bridges and roads generally take the form of labour contributions.

**Table 7: Impact of barazas on contributions**

| | Mean | Sub-county baraza | Information | Delibe-ration | District baraza |
|---|---|---|---|---|---|
| In-kind contributions to the school? | 0.321 | 0.006 | -0.085** | -0.019 | -0.059 |
| | (0.467) | (0.032) | (0.032) | (0.035) | (0.031) |
| In-kind contributions to the health centre? | 0.126 | 0.011 | -0.03 | -0.031 | -0.061+ |
| | (0.332) | (0.023) | (0.025) | (0.021) | (0.023) |
| In-kind contributions to the road/ bridge? | 0.384 | 0.025 | -0.039 | -0.011 | -0.052 |
| | (0.486) | (0.043) | (0.037) | (0.037) | (0.035) |
| In-kind contributions to the drinking water facility? | 0.452 | 0.047 | -0.01 | 0.059 | -0.004 |
| | (0.498) | (0.046) | (0.042) | (0.038) | (0.043) |
| In-kind contributions to the dam/irrigation facility? | 0.093 | 0.022 | -0.024 | -0.028 | -0.019 |
| | (0.291) | (0.031) | (0.020) | (0.029) | (0.020) |
| In-kind contributions to any government structure? | 0.233 | 0.04 | -0.073* | 0.012 | -0.025 |
| | (0.423) | (0.034) | (0.029) | (0.034) | (0.029) |
| In-kind Contribution Index | 0.000 | 0.063 | -0.107+ | -0.016 | -0.093 |
| | (0.609) | (0.068) | (0.057) | (0.058) | (0.057) |
| Cash contributions to the school? | 0.382 | -0.005 | 0.053 | 0.021 | 0.101* |
| | (0.486) | (0.026) | (0.035) | (0.039) | (0.028) |
| Cash contributions to the health centre? | 0.121 | -0.023 | 0.053 | 0.051 | -0.014 |
| | (0.326) | (0.024) | (0.035) | (0.040) | (0.017) |
| Cash contributions to the road/ bridge? | 0.097 | -0.017 | 0.001 | 0.021 | 0.001 |
| | (0.296) | (0.022) | (0.015) | (0.031) | (0.013) |
| Cash contributions to the drinking water facility? | 0.37 | -0.044 | 0.107* | 0.057 | 0.094 |
| | (0.483) | (0.034) | (0.043) | (0.048) | (0.043) |
| Cash contributions to the dam/irrigation facility? | 0.04 | 0.001 | 0.008 | 0.001 | -0.003 |
| | (0.197) | (0.015) | (0.012) | (0.014) | (0.011) |
| Cash contributions to any government structure? | 0.26 | 0.008 | -0.027 | 0.007 | 0.029 |
| | (0.439) | (0.030) | (0.026) | (0.039) | (0.030) |
| Cash Contribution Index | 0.000 | -0.033 | 0.076* | 0.063 | 0.067 |
| | (0.536) | (0.041) | (0.037) | (0.056) | (0.032) |
| Number of observations | 6,704 | 2,738 | 4,858 | 4,858 | 3,687 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels, respectively.

The information component of the sub-county baraza reduces in-kind contributions, but increases cash contributions. In-kind contributions are particularly reduced for schools and government or community buildings. The increase in cash contributions as a result of the baraza intervention is especially for drinking water infrastructure.

#### *4.1.10 Perceptions and prioritisation*

In this section, we provide results on changes in citizens' perceptions of a range of problems. Respondents were given a statement and, using a 10-point Likert scale, had to indicate how much they disagreed (1) or agreed (10) with the statement. The statements were based on extensive qualitative work, where various stakeholders were interviewed and asked about key problems surrounding public service provision in different sectors.

*Table 8* shows that, as a result of the information component of a sub-county baraza, households tend to agree more that access to drinking water sources is a serious problem. Households that received a deliberation-focused sub-county-level baraza are also more likely to agree that drinking water is usually dirty.

In the area of public health provision, households that were exposed to a sub-county-level information baraza are more likely to agree that access to a health centre or hospital is a serious problem. We further find that households that were exposed to a sub-county-level baraza indicate that lack of medicines at health centres or hospitals is less of a problem than in control areas. A direct comparison for this outcome between sub-county-level barazas and district-level barazas also yields a significant difference. We also asked about perceptions related to friendliness of staff and absenteeism. We find that households that live in areas that received the district-level treatment are more inclined to say that absenteeism is a problem, but the difference is not significant.

We then look at perceptions in the area of education. We see that households are generally most concerned about poor-quality learning outcomes, but think absenteeism is less of a problem. There is no significant difference between groups for any of the school-related perceptions. Also the perception of access to roads as a serious problem does not change as a result of the barazas.

Respondents seem to perceive agricultural service delivery as the most problematic area. Averages on the Likert scales are fairly high when asked whether farmers agree extension officers visit rarely and there is a lack of transparency in how farmers are selected to receive inputs from government. We see that the issue of transparency reduces somewhat after a sub-county-level baraza, but the effect is not significant. A perception index that combines all outcomes indicates only a significant difference between sub-county-level barazas and district-level barazas.

#### *4.1.11 Sub-county-level analysis*

In addition to household surveys, we conducted surveys with 261 government officials as respondents. As in the previous subsection, we analysed data on agriculture, infrastructure, health and education. Obviously, sample sizes are much smaller here, so results should be interpreted with this caveat in mind. The full analysis of sub-county data can be found in *Table A.4*.

**Table 8: Impact of barazas on perceptions**

| | Mean | Sub-county baraza | Information | Delibe-ration | District baraza |
|---|---|---|---|---|---|
| Access to a drinking water source is a serious problem | 5.151 | 0.048 | 0.606** | 0.410 + | 0.143 |
| | (3.264) | (0.265) | (0.223) | (0.227) | (0.160) |
| Drinking water is usually dirty | 4.428 | 0.072 | 0.057 | 0.442+ | 0.049 |
| | (3.129) | (0.232) | (0.199) | (0.229) | (0.254) |
| Access to a government health centre or hospital is a problem | 5.819 | -0.193 | 0.365+ | 0.016 | 0.290 |
| | (3.092) | (0.273) | (0.218) | (0.261) | (0.150) |
| Government health centres or hospitals do not have relevant medicines | 6.495 | -0.412+ | -0.027 | 0.018 | 0.169 |
| | (3.024) | (0.204) | (0.182) | (0.206) | (0.168) |
| Staff at government health centres or hospitals are rude to patients | 5.040 | -0.048 | 0.015 | 0.096 | 0.053 |
| | (2.913) | (0.224) | (0.155) | (0.205) | (0.165) |
| Medical staff at government health centres or hospitals are often absent | 4.776 | 0.032 | 0.081 | 0.127 | 0.301 |
| | (2.757) | (0.173) | (0.142) | (0.202) | (0.152) |
| Access to a government primary school is a serious problem | 4.930 | 0.032 | 0.046 | 0.021 | 0.037 |
| | (2.905) | (0.246) | (0.205) | (0.21) | (0.227) |
| Teachers in government schools are often absent | 4.847 | -0.074 | 0.011 | -0.061 | 0.124 |
| | (2.72) | (0.182) | (0.17) | (0.211) | (0.189) |
| Children's learning outcomes in government schools are poor | 6.360 | -0.194 | 0.166 | 0.14 | -0.246 |
| | (2.918) | (0.18) | (0.155) | (0.187) | (0.154) |
| Availability/access to all-weather roads is a serious problem | 5.157 | -0.348 | -0.023 | -0.18 | 0.118 |
| | (3.14) | (0.289) | (0.225) | (0.229) | (0.113) |
| Agricultural inputs supplied by the government are of poor quality | 5.845 | 0.227 | -0.027 | -0.105 | 0.130 |
| | (2.788) | (0.16) | (0.129) | (0.16) | (0.176) |
| There is lack of transparency in how farmers are selected to receive agricultural inputs from govt. | 6.352 | -0.351 | 0.22 | 0.042 | -0.024 |
| | (3.165) | (0.229) | (0.25) | (0.259) | (0.198) |
| Agricultural extension agents rarely visit | 6.372 | -0.189 | -0.001 | 0.103 | 0.007 |
| | (3.218) | (0.268) | (0.301) | (0.344) | (0.233) |
| Agricultural extension agents are not aware farmers' needs | 6.098 | -0.01 | 0.082 | 0.13 | 0.458+ |
| | (3.074) | (0.224) | (0.254) | (0.321) | (0.162) |
| Perception Index | 0.000 | -0.033 | 0.035 | 0.026 | 0.031 |
| | (0.514) | (0.039) | (0.033) | (0.037) | (0.030) |
| Number of observations | 6,704 | 2,738 | 4,854 | 4,854 | 3,685 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels.

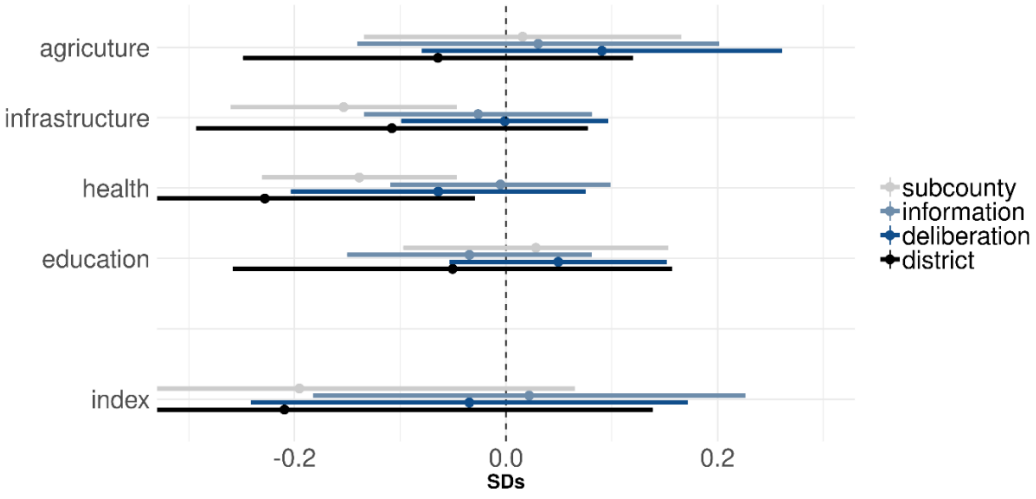
## 4.2 Heterogeneity of impacts

### *4.2.1 Heterogeneity in the timing of the intervention*

The slow rollout of the intervention over an extended period also introduces variation in the time that passed between treatment administration and endline data collection. For instance, the first barazas were held around June 2016 (about one year after the baseline), so more than three years will have passed between treatment administration and endline data collection. For the most recent barazas, there will only be a few months between treatment administration and endline data collection. One may argue that sub-counties or districts that were treated early on have been exposed to the programme much longer; hence, one may expect larger effects on a range of outcomes for these sub-counties or districts than areas that only recently received treatment. Places might have also been treated with barazas first because they were politically favoured. It is furthermore possible that our implementing partner made a greater effort at the beginning than at the end of the programme. Therefore, the effects of these early interventions may be systematically different. At the same time, for some outcomes, effects of the baraza intervention may dissipate – or even reverse – over time as enthusiasm fades, plans are abandoned and promises forgotten.

We find that the OPM organised quite a few barazas in May 2019. We thus reran the analysis and added an interaction term between the treatment indicator and an indicator variable that takes the value of one if the baraza the household was exposed to happened more than one and a half years before the endline data collection (the indicator is coded as zero for the control group). Results are summarised in *Figure 10*. It displays average treatment effects for the four hypotheses on the four families of outcomes and one overall index, similar to the summary in *Figure 9*.

**Figure 10: Heterogeneity at sub-county level – effects more than one and a half years after implementation**



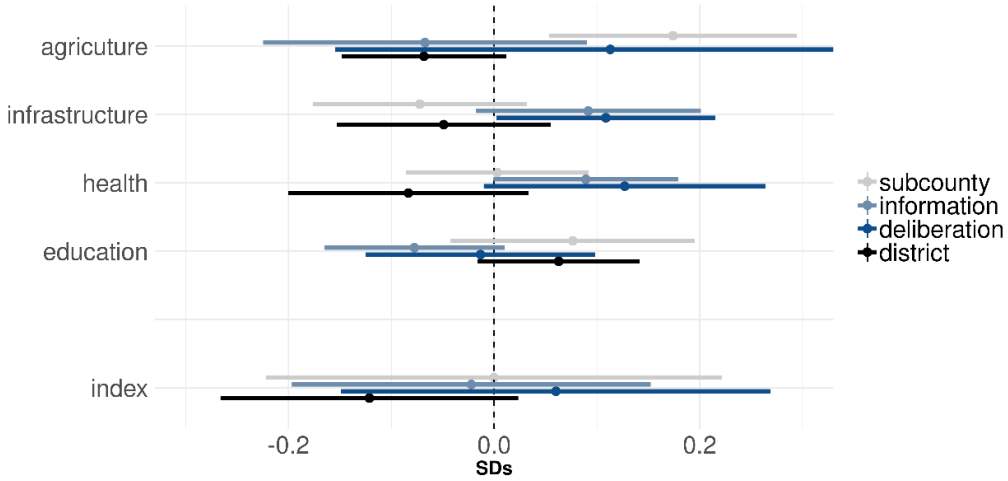

For the agricultural sector, we do not find that the time elapsed between the intervention and endline data collection affects the impact of sub-county-level barazas as we found in *Figure 9*. There is a significant negative interaction effect from sub-county-level barazas for the infrastructure index. For outcomes in the health sector, there are negative interaction effects for sub-county-level and district-level barazas. The negative

interaction effect for the sub-county-level barazas seems to be driven by the reduced likelihood that households turn to government health facilities for maternal healthcare in the long term. There are indications that these households are switching back to traditional healers, which may point to disappointment in the lack of progress in public health facilities. For education, we do not find significant interaction effects for the index. We detect a positive effect of sub-county-level barazas on enrolment in the long run. Overall, results suggest that the effectiveness of barazas seems to dissipate over time.

#### *4.2.2 Heterogeneity related to officials recalling barazas*

Our treatment indicator is based on information from the implementing partner. However, we also asked officials at sub-county headquarters whether they recalled if a baraza took place in the past five years. We also use this variable to check for heterogeneous treatment effects. Results are in *Figure 11*.

**Figure 11: Heterogeneity at sub-county level – officials recall baraza**




Even though restricting the sample to households that live in a sub-county where officials recall a baraza took place in treatment areas reduces sample size by only about 25 per cent, we find substantial changes in the results. For the agricultural sector, differences with the full sample are minor. If anything, the positive effect of sub-county barazas on public service delivery in agriculture stands out even more. For infrastructure, the deliberation component seems to increase service provision, driven by a significant reduction in distance to water source.

Interestingly, we now also perceive significant results for the health sector. The deliberation component of a sub-county baraza increases the likelihood that households seek treatment in government health facilities when ill. The information component is associated with greater use of government health facilities for maternal health and also increases the likelihood that a VHT is present. Both components also reduce waiting time before being attended to. There are no effects on education service delivery.

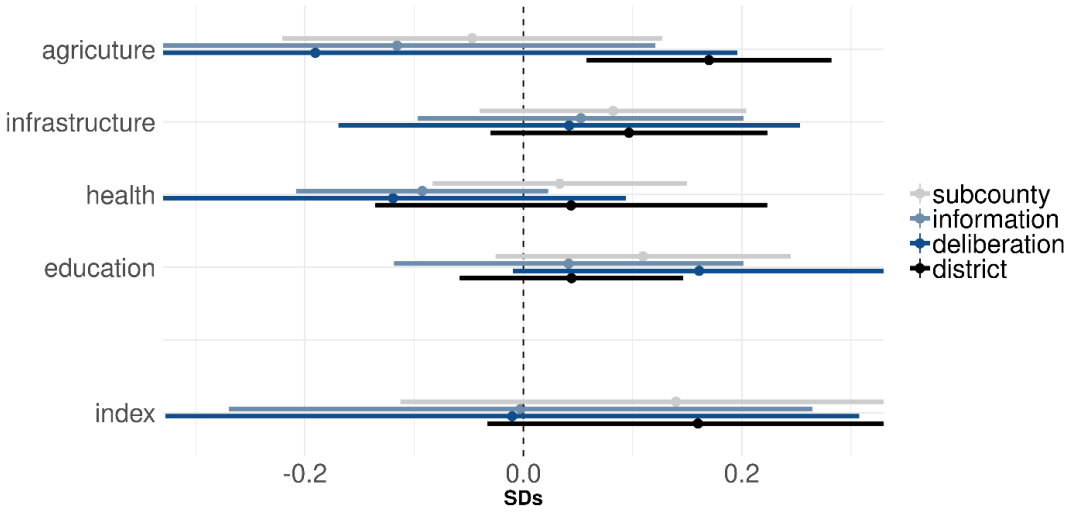
The generally larger impacts we find in this sub-sample, particularly for the health sector, are intriguing. It is possible that officials who recall a baraza are intrinsically more motivated and, thus, more receptive to community-based monitoring. Alternatively, it may be that the information we received from the OPM is inaccurate and some sub-counties indicated as being treated were, in fact, not.

### *4.2.3 Heterogeneity related to remoteness*

Differences in the timing of the treatment and the fact that an official recalls a baraza introduce heterogeneity at treatment level. However, heterogeneity may also depend on household characteristics. Because outcomes are likely to be correlated within sub-counties, we will have more statistical power to assess heterogeneity related to household characteristics than heterogeneity that originates at the treatment level.

One potential source of treatment heterogeneity at the household level is related to remoteness. Indeed, during discussions with stakeholders, it was often argued that barazas may have different effects on households that live close to the sub-county headquarters than those that live in remote areas. At baseline, we collected data on the distance between the homestead of the household and the sub-county headquarters. This median distance was 5 kilometres, so we reran the regressions, but only for households that live 5 kilometres or more away from the district headquarters. Results are summarised in *Figure 12*.

**Figure 12: Heterogeneity at individual level – living >5 km from sub-county headquarters**



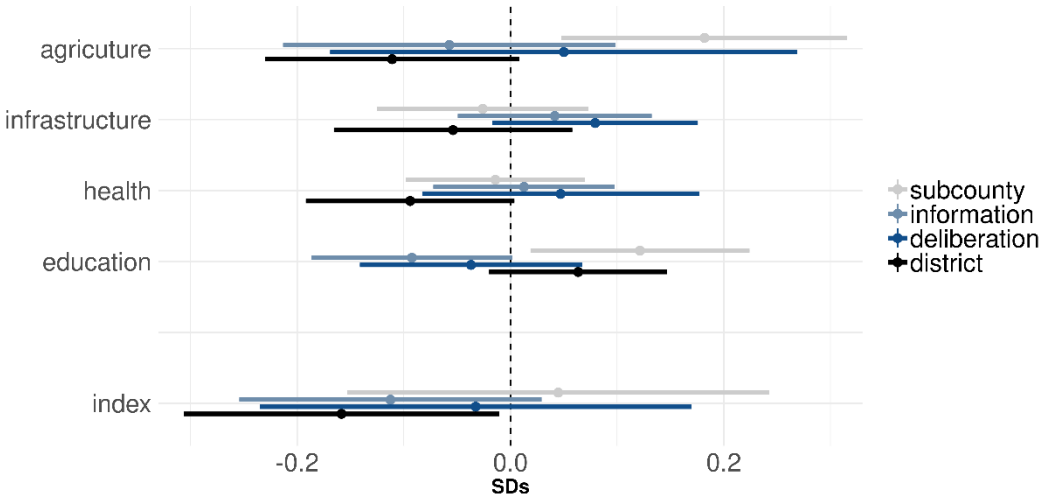

There are some indications that district-level barazas are effective for households that live further away from the sub-county. This is particularly the case for outcomes in the agricultural sector. For households in remote areas, we find that a district-level baraza leads to a significant increase in the likelihood that: (1) households are visited by extension officers at home; (2) they visit an extension office, demonstration site or model farmers; (3) they are assisted by NAADS; and (4) they are supported by a marketing cooperative. Remote households also benefit from sub-county-level barazas in terms of access to a protected water source, while district-level barazas reduce waiting time at the source for this sub-group. We also find more positive treatment effects in the education sector if we focus on households that live further from the district headquarters. Overall, there are some indications that sub-county- and district-level barazas are particularly effective for households living in more remote locations.

### *4.2.4 Heterogeneity related to households being aware of barazas*

Of the total sample of 6,700 households, about 3,160 households responded they were aware of the concept of a baraza (and about 1,750 reported they remember that in the

past five years, such a meeting was held in their sub-county). Being aware of the concept of a baraza may indicate that one is better informed or more interested in governance and public service delivery, which might also be an important source of heterogeneity at the individual level. Consequently, we reran the analysis, but only for the subset of households that indicated they were aware of the concept of a baraza. Results are in *Figure 13*.

**Figure 13: Heterogeneity at individual level – knows baraza**




The positive impact of sub-county-level barazas on agriculture becomes stronger. The effects are driven by more extension officers visiting and an increase in the support of NAADS/Operation Wealth Creation in the village. This suggests that households that are well informed and interested in public service provision are, particularly, able to cash in on the baraza. We also detect a clear effect of sub-county-level barazas on the education sector. While this effect is caused by increased enrolment in public schools, there are also significant positive effects on school infrastructure, such as fencing and access to water on the school premises. The information component of the baraza significantly reduces outcomes. One reason might be that providing only information (but no voice) to people who are receptive to participatory governance may lead to frustration, causing them to view some of the outcomes in a more negative light. This explanation is consistent with the results of the underlying individual outcomes. While outcomes such as enrolment rates and school infrastructure are no different between households exposed to information and those that are not, the former group complains significantly more about not being informed about the SMC.

## 4.3 Threats to validity/robustness

In this study, the primary threat to validity is the possibility that the partial rollout introduced selection bias. We already showed in *Table 1* that the updated balance table that compares a range of baseline characteristics of actual treated households to control areas displays similar balance to the original balance table comparing planned treatment areas to planned control areas. While this is reassuring, in this section, we present additional balance checks to further explore whether the rollout of the intervention was not random.

First, we can investigate whether selection bias was introduced by comparing outcomes in control sub-counties to outcomes in sub-counties that were allocated to receive treatment but ended up not receiving it.[9] The idea is that if the rollout was random, sub-counties that were allocated randomly to a particular treatment at the design stage but did not end up receiving treatment can be interchanged with sub-counties that were randomly selected at design stage to function as control sub-counties. Finding no significant differences in outcomes between these two groups would support the hypothesis that the partial rollout did not introduce selection bias. If the incomplete rollout introduced selection bias, comparing these two groups may also be informative to assess the direction and magnitude of bias.

*Table 9* presents the original balance table (*Table A.1*), but after dropping sub-counties that were treated. Thus, instead of comparing pre-treatment characteristics between treatment sub-counties and control sub-counties, the table compares sub-counties that were allocated to a particular treatment (but did not end up receiving it) to the (planned) control sub-counties for that particular treatment. The table seems to suggest that the rollout did not introduce imbalance, at least as judged by the pre-treatment characteristics that were in the original balance table. We find that, out of 30 comparisons, we reject the null hypothesis of no difference at the 1% significance level once, at the 5% level once and at 10% level once. Also, this would be expected by pure chance alone, so we conclude that the partial rollout did not seem to introduce selection bias.

We also revisit the risk that the partial rollout may have introduced selection bias. While results in *Table 1* and *Table 9* are reassuring, it should be noted that pre-treatment characteristics were collected some time ago and results could be different if more recent data were used and/or selection happened on characteristics that change over time. Therefore, we repeat the comparisons between control sub-counties and sub-counties that were allocated to a treatment cell but ended up not being treated from *Table 9*, but now use endline data. For this reason, specifically, instead of simply collecting endline data from the (planned) control sub-counties, we also collected endline data from sub-counties that were supposed to receive a treatment but did not get one. *Table 10*, thus, compares endline outcomes between households that were scheduled to receive a particular treatment but did not end up receiving it to outcomes of households that were assigned to serve as a control for the particular treatment. In the table, we present results for the indices that are also used to summarise impact in *Figure 9*.

We find significant differences between planned but not treated sub-counties and sub-counties that were allocated to the control condition for the agricultural sector. For instance, we find that households that were supposed to receive a sub-county-level baraza treatment but did not get one are 10 percentage points more likely to indicate that they visited an extension office, demonstration site or model farmer. However, because this difference is positive, it could be argued that the OPM seems to have prioritised sub-counties with poorer service delivery in the agricultural sector. As a result, positive results obtained from comparing treated and control groups are likely to underestimate the true impacts of the treatment. It might also result in the fact that some of the positive effects we find turn out to be insignificant.

[9] All district-level barazas were implemented, so we only focus on sub-counties here and in the following sections.

We also find some imbalance when investigating the relative importance of the information component. Here, the imbalance is caused by two variables that measure assistance in marketing. Also, the OPM may have prioritised sub-counties where cooperatives and village marketing committees are less active.[10]

**Table 9: Balance between planned but not treated sub-counties and planned controls**

| | Mean | Sub-county baraza | Information | Deliberation |
|---|---|---|---|---|
| Household size | 6.324 | 0.012 | 0.388* | 0.022 |
| | (2.825) | (0.171) | (0.170) | (0.140) |
| Age of the household head (years) | 46.501 | 0.357 | 0.698 | 0.553 |
| | (14.615) | (0.714) | (0.663) | (0.808) |
| Head of household is woman (1=yes) | 0.191 | 0.008 | -0.019 | -0.003 |
| | (0.393) | (0.017) | (0.016) | (0.017) |
| Head finished primary education (1=yes) | 0.213 | -0.007 | -0.007 | -0.003 |
| | (0.410) | (0.019) | (0.027) | (0.022) |
| Thatched grass roof (1=yes) | 0.298 | -0.002 | 0.000 | -0.036 |
| | (0.457) | (0.029) | (0.024) | (0.027) |
| Traditional mud wall (1=yes) | 0.424 | 0.007 | -0.057 | 0.044 |
| | (0.494) | (0.049) | (0.047) | (0.044) |
| Distance to nearest all-weather road (km) | 0.906 | 0.284** | 0.010 | 0.187 |
| | (0.915) | (0.131) | (0.100) | (0.110) |
| Access to extension (1=yes) | 0.108 | 0.005 | 0.008 | 0.007 |
| | (0.310) | (0.015) | (0.016) | (0.015) |
| VHT in village (1=yes) | 0.854 | -0.007 | -0.01 | -0.015 |
| | (0.353) | (0.035) | (0.028) | (0.028) |
| Number of children in public schools | 2.478 | 0.043 | 0.249+ | 0.076 |
| | (2.074) | (0.112) | (0.115) | (0.100) |
| Number of observations | 12,545 | 4,293 | 7,842 | 8,391 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively.

[10] Note that in all cases, the bias is in a conservative direction, which is likely to lead to an underestimate of the treatment effect. We would be more worried if we found, for example, that households that were scheduled to receive a treatment but did not end up getting it had a 10 per cent lower incidence of visits by extension workers. This may indicate that the OPM selected areas where extension was already stronger than in average areas and this higher incidence of extension visits would erroneously be attributed to the baraza intervention.

**Table 10: Difference between planned but not treated sub-counties and planned controls at endline**

| | Sub-county baraza | Information | Deliberation |
|---|---|---|---|
| Agriculture index | 0.174** | 0.113 | 0.045 |
| | (0.057) | (0.057) | (0.057) |
| Infrastructure index | 0.026 | -0.031 | -0.024 |
| | (0.073) | (0.071) | (0.071) |
| Health index | 0.026 | -0.028 | -0.012 |
| | (0.047) | (0.039) | (0.043) |
| Education index | 0.093 | -0.002 | 0.116 |
| | (0.057) | (0.046) | (0.045) |
| Public service delivery index | 0.161 | 0.004 | 0.075 |
| | (0.083) | (0.070) | (0.069) |
| Number of observations | 1,637 | 2,356 | 2,808 |

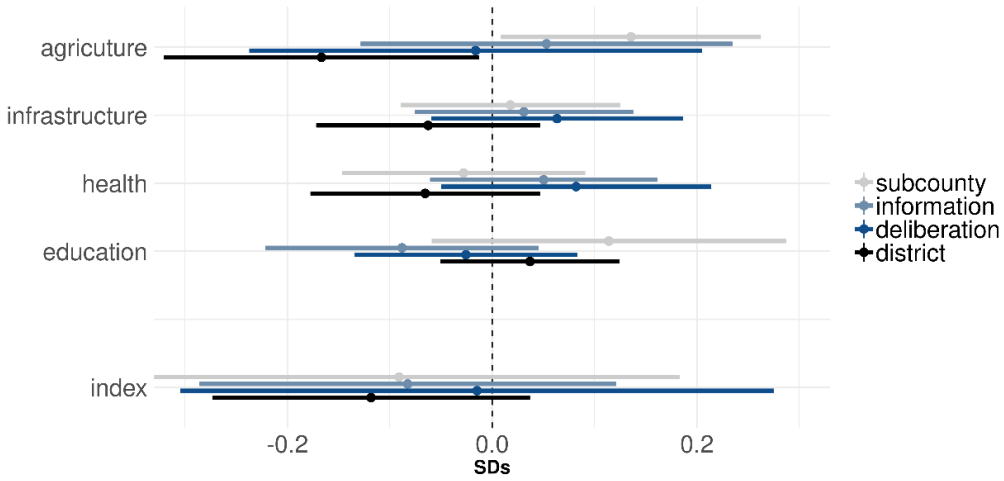
**Note:** column 1 reports difference (and standard errors below) of the sub-county-level baraza intervention; column 2 reports the effect (and standard errors below) of the information component of the baraza intervention; column 3 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively.

Our pre-analysis plan prescribes that if we find evidence of imbalance between planned but untreated sub-counties and planned control sub-counties using endline information, we will try to recover unbiased impact estimates using a matched difference-in-difference estimator. However, we find that baseline outcomes do not predict endline outcomes very well. When autocorrelations are low, there are large improvements in power to be had from using ANCOVA instead of difference-in-differences (McKenzie 2012). Therefore, we deviate from our pre-analysis plan and use matching and estimate ANCOVA models on these pre-processed data.

For the matching, we use Mahalanobis distance with coarsened exact matching, an extremely powerful method of matching (Iacus et al. 2012). We match on (baseline values of) household size, sex of the household head, age of the household head, whether the household head finished secondary education, the logarithm of farm size, housing conditions (iron roof and improved wall), phone ownership, and latitude and longitude. For the coarsened exact matching, custom cut points are defined to construct three age categories, six farm size categories and a five-by-five grid based on GPS coordinates. For the comparison between sub-county-level barazas and district-level barazas, we do not match on GPS coordinates, because this results in too many observations that cannot be matched. Endline data are then merged to the matched dataset and standard ANCOVA models such as those used in the main analysis are estimated. *Section 6.4* discusses the limitations of matching methods.

*Figure 14* provides a summary similar to *Figure 9*. Matching does not change the main conclusions. However, there are some differences between the matched and unmatched results when looking at individual outcomes. We provide detailed results similar to those in *Table 2* up to *Table 5* in Appendix *Tables A.5 to A.8* and provide a brief discussion of the most striking differences here.

**Figure 14: Summary of baraza impact (matched ANCOVA)**




For agriculture, comparing *Table 2* to Appendix *Table A.5*, we see that after matching, the positive impact of the sub-county baraza on the likelihood of receiving seed from the government has disappeared. The positive effect of sub-county barazas on on-farm visits of extension workers is very similar. However, after matching, we also detect a significant and positive effect of sub-county-level barazas on the likelihood that a member of the household visited an extension office, demonstration site or model farmer. This is consistent with the imbalance that we found above: Households in areas where a sub-county baraza was planned but not held reported a 10 per cent higher incidence of visits to extension offices, demonstration sites or model farmers than farmers in areas where no sub-county-level baraza was planned. We also establish that it is more likely that farmer associations and cooperatives are present when a sub-county baraza is held. After matching, both information and deliberation components seem equally important in spurring the formation of cooperatives and associations. The analysis confirms that district-level barazas are less effective than sub-county-level barazas on a range of agriculture-related outcomes.

For infrastructure, the reduction in waiting time at the water source as a result of sub-county barazas ceases to be significant after matching (*Table A.6*). The negative impact of the information component on the likelihood that households treat drinking water also disappears. We now do find a positive and significant effect of the baraza intervention on citizens' participation in water user committees. The effect seems to be driven by the deliberation component.

Further comparing *Table 4* to *Table A.6*, we see that the effect of sub-county-level barazas on the likelihood that VHTs organise public meetings persists. We also see that the somewhat puzzling negative effect of the information component on the likelihood that there is a functioning health management unit at the government health facility disappears after matching. This negative effect is replaced by a positive impact associated with a sub-county-level baraza. Finally, we compare *Table 5* to Appendix *Table A.8* to assess the potential impact of non-random rollout on the results for public service delivery in the education sector. Results are very similar. The negative effect of the information component on the likelihood that inspectors visit the school disappears after matching.

# 5. Cost analysis

## 5.1 Cost information

Cost information was provided by the OPM and is reported in *Table 11*. The OPM implemented 74 of 155 planned baraza forums. As such, the estimation provided here covers only the cost of the 74 barazas that were reported to have been implemented.

**Table 11: Baraza costs in absolute terms**

| | Number of barazas | Average cost (UGX) | Total cost (UGX) | Min. cost (UGX) | Max. cost (UGX) |
|---|---|---|---|---|---|
| District baraza | 7 | 15,325,000 | 107,275,000 | 13,900,000 | 18,500,000 |
| Sub-county baraza | 20 | 12,837,500 | 256,750,000 | 11,200,000 | 14,000,000 |
| Information baraza | 29 | 12,962,500 | 375,912,500 | 11,300,000 | 13,800000 |
| Deliberation baraza | 18 | 12,712,500 | 228,825,000 | 11,100,000 | 14,200,000 |
| Total | 74 | | 968,762,500 | | |

In total, implementing 74 barazas cost the GoU about 968,762,500 Ugandan shillings (UGX).[11] A large share (39%) of this cost originated from sub-county information barazas. Due to the incomplete rollout and the distribution of the target sites from the centre and the resulting large differences in the number of barazas implemented for each type, the different types of barazas vary widely in total costs. However, the average cost of implementing different kinds of sub-county barazas does not differ a lot.

## 5.2 Cost-effectiveness analysis

This analysis compares the costs of the different types of barazas (sub-county-level baraza, information sub-county-level baraza, deliberation sub-county-level baraza and district-level baraza) and their estimated effects. The costs were obtained from the OPM and are reported in *Table 11*. However, assigning monetary values to measures of effects is more challenging. Associating a monetary value to the effect of the baraza on the indices (e.g. to the effect of the sub-county-level barazas on the agricultural index) is difficult, because the indices combine several individual variables. That is why we will focus on individual outcomes and assign monetary values to these individual outcome variables. We selected only individual outcomes for this cost-effectiveness analysis, for which the estimated effect of the baraza intervention was significantly different from zero, as for non-significant outcomes, the estimated benefit of the intervention is zero. Among those significantly affected individual outcomes, we chose the ones for which monetising is relatively feasible, because it is, for example, impossible or at least inappropriate to attach a monetary value to the effect of barazas on the share of households that visit a public health facility to give birth.

A key objective of this study is to compare the effectiveness of barazas organised at lower administrative levels (sub-county) to the effectiveness of barazas organised at a

[11] Exchange rate: 1 USD = 800 UGX.

more aggregate level (district). The level of administrative placement is an important determinant of the cost-effectiveness of the policy intervention: implementing a district-level baraza affects far more people than implementing a sub-county-level baraza, yet a district-level baraza costs only a little more than a sub-county baraza (*Table 11*). As such, organising a district-level baraza could be more cost effective, even though the sub-county-level baraza seems to have a larger impact at first sight.

For instance, in the area of infrastructure, we find that households have to wait on average about 37 minutes. A baraza intervention at the sub-county level reduces this time by about 29 per cent, which corresponds to a reduction of about 11 minutes per household. Assuming a member of the household visits the water source once a day, the intervention saves 3,862 minutes (64 hours) per household per year. On average, 5,100 households live in one sub-county (Uganda Bureau of Statistics 2017, pp.37–38), so this sub-county intervention saves 19,698,393 minutes (328,306 hours) per sub-county per year. For a district-level baraza, the impact is -3.2 per cent, corresponding to 1.18 minutes every time a member of the household goes to the water source, totalling 432.16 minutes (7.20 hours) saved per household per year. However, an average 60,840 households live in one district (ibid.), so this district-level intervention saves 26,292,614 minutes (438,210 hours) per year. Therefore, the district-level intervention saved 109,904 hours more than the sub-county-level intervention. To attach a monetary value to this time difference, we consider the average hourly wage rate of 750 UGX. This results in a difference in impact of 82,428,180 UGX, while the district-level baraza was only 2,487,500 UGX more expensive.

We now consider an example from the agricultural sector, namely access to extension visits at home. About 17.8 per cent of households report an expert visited them. A sub-county baraza increases this by about 5.6 percentage points, so about 23.4 per cent have access to extension at home, which corresponds to 1,193 households. However, the intervention at the district level decreases the access to extension at home by 2.7 percentage points. This means that the sub-county-level intervention is more effective, as long as its benefits outweigh its costs. The 5.6 percentage points increase corresponds to 286 more households having access. From our baseline data, we know that the average household farms 5.8 acres of land, which means that 1,656 acres are affected by the intervention. We use maize in our calculation because maize is an important crop in Uganda, both for home consumption and as a traded commodity due to its relatively high value-to-weight ratio. Average maize yields are about 618 kilograms per acre for the main growing season (Uganda National Household Survey 2005/06). Assuming that access to extension raises yields by 10 per cent (Van Campenhout et al. 2020), this results in 102,370 kilograms more maize produced due to the sub-county intervention. Assuming a bag of 100 kilograms of maize sells at a median price of 60,000 UGX, the monetary benefit of a single sub-county baraza, only considering access to extension, amounts to 61,422,278 UGX, while its average cost is 12,837,500 UGX.

For education, we consider the number of children in public schools to be an important outcome. Because we can only find an effect on public service delivery in the education sector after deleting observations from households in sub-counties or districts where a baraza was held recently, we use this part of the analysis to compare the cost-effectiveness between sub-county-level and district-level barazas. We find that households have on average 1.79 children in school. A baraza intervention at the sub-

county level increases this number by about 0.37 children per household. On average, 5,100 households live in one sub-county, so this sub-county intervention leads to an additional 1,887 children in school. However, the intervention at the district level decreases this number by about 0.02 children per household. This means that the sub-county level intervention is more effective, as long as its benefits outweigh its costs.

For the health sector, we only find an effect on public service delivery after deleting observations from households in sub-counties or districts where a baraza was held recently. So, again, we use this part of the analysis. We look at the waiting time before being attended. In the previous three examples, it depended on the outcome whether a sub-county baraza is superior to a district baraza or not. We now also compare the cost-effectiveness of the information and deliberation components. Both reduce waiting times. Households must wait for 90 minutes on average. The information treatment reduces this time by about 21.9 per cent, which corresponds to a reduction of about 20 minutes per household, every time a member of the household visits a public health facility. Looking at our baseline data, we see that a member of the household visits this kind of facility six times a year, so the information intervention saves 118 minutes per household per year. On average, 5,100 households live in one sub-county, so this information intervention saves 603,200 minutes (10,053 hours) per sub-county per year. For a deliberation baraza, the impact is -23.55 per cent, corresponding to 21 minutes every time a member of the household goes to a public health facility, so 127 minutes saved per household per year. As 5,100 households live in one sub-county, this deliberation intervention saves 648,620 minutes (10,810 hours) per year. The deliberation intervention, thus, saved 757 (4,590) hours more than the information intervention. To attach a monetary value to this time difference, we consider the average hourly wage rate again. This results in a difference of 567,750 UGX, while both types of baraza are similar in costs (12,962,500 UGX versus 12,712,500 UGX).

# 6. Discussion

## 6.1 Introduction

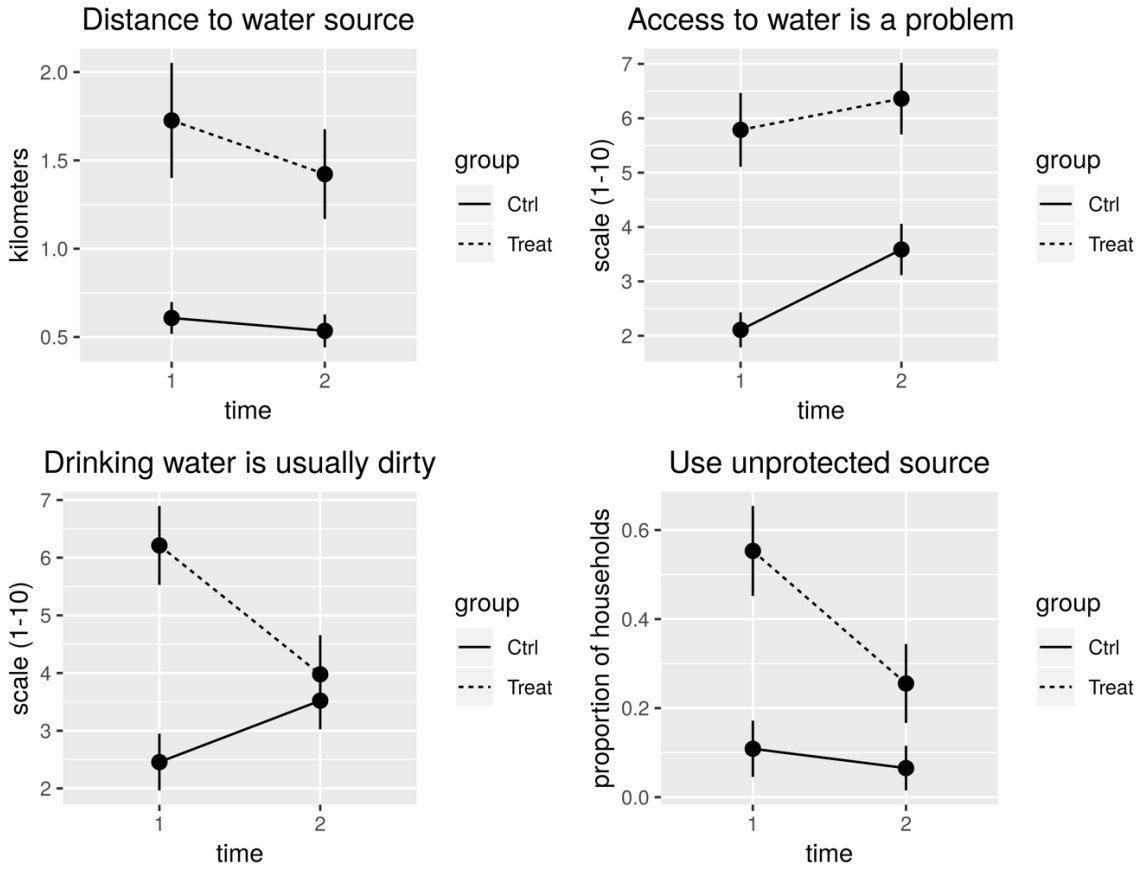
While we do not find that barazas affect public service delivery in general, we do find a variety of interesting effects when we look at individual outcomes and consider heterogeneity in the treatment effects. In light of this, our results confirm some of the likely explanations for why Raffler and colleagues (2018) fail to find significant results on health outcomes in their study. For instance, there are indications that it may take some time for effects to materialise. The endline data in Raffler and colleagues (2018) were collected after 20 months, hence their results may only apply in the short run. In addition, the fact that our results are somewhat more encouraging than those found in Raffler and colleagues (2018) could also be related to the fact that our intervention is organised by the government. Raffler and colleagues (2018) find indications that the presence of sub-county officials during the programming boosted the impact of the intervention on treatment quality in health centres. In line with this, community-based monitoring interventions organised by government might be more effective. This confirms that top-down monitoring may be more important in changing the behaviour of civil servants than bottom-up monitoring by citizens.

This study mostly focuses on the analysis of the endline data of the quantitative component of the impact evaluation. A previous study also provides a less ambitious qualitative exploration of the likely impact of barazas (Van Campenhout et al. 2018). In that study, we find that stakeholders think barazas are useful at improving public service delivery across all sectors, especially if the they take place at the sub-county level. Stakeholders had no difficulty providing examples of changes they felt were the direct result of a baraza being held: projects that were previously dragging were finished or taken up afresh; sub-standard work was redone; and, in some instances, priorities were changed to better align with citizens' needs. A substantial part of these outcomes seemed to derive from the barazas' potential to fix information asymmetries. Focus group discussions revealed civil servants reacted to the consequences of the increased likelihood that sub-standard work would be exposed and politicians responded to electoral considerations, which signals barazas increased bottom-up pressure. There were also indications that barazas boosted community involvement and top-down monitoring.

The diverging results from the qualitative and quantitative analyses may be due to the fact that the baraza programme is a broad intervention that attempts to address a range of issues in a heterogenous setting. It might be that the baraza is effective for some, but not for others. However, if a simple average treatment effect is estimated, the effect may turn out to be insignificant, because it averages over subgroups. For instance, access to water is likely to be more of a problem in remote areas. Even if a baraza increases access to water and reduces waiting times, this may not show up if there is a large group close to the sub-county centre that already has access to water and no additional boreholes were constructed in these areas.

This is illustrated when we link the endline data back to what we learned in the qualitative fieldwork (Van Campenhout et al. 2018). In Bagezza sub-county in Mubende district, drinking water was mentioned as a serious problem and it was discussed extensively during the baraza. When we went back to the sub-county to test the endline tool, it appeared that the government had made whole on its promises and the sub-county now had access to drinking water. To check this, we used baseline and endline data, and simply compared means between Bagezza and a random control sub-county in the neighbourhood (Bwanswa in Kibaale district). The results are presented in *Figure 15*.

**Figure 15: Access to water in two sub-counties**




The figure shows that on all four of the water-related outcomes, Bagezza scored significantly worse than the control sub-county. For instance, in Bagezza, average distance to the nearest water source is 1.75 kilometres, while in Bwanswa, it is only about 650 metres. The average score on the Likert scale used to measure perceptions on the cleanness of drinking water is more than six in Bagezza and only two in Bwanswa. In Bagezza, more than half of all households rely on unprotected drinking water sources and in Bwanswa, only about 10 per cent.

More importantly, we see that the difference between Bagezza and Bwanswa has reduced since the baraza happened in Bagazza. While still significantly higher than in Bwanswa, distance to the water source has reduced to about 1.35 kilometres. Perceptions of access to water become more negative as a result of a general drought in East Africa, but less so in Bagezza than in Bwanswa. The most impressive progress is made in terms of the quality of water. At the time of the endline, there is no more difference between Bagezza and Bwanswa. Use of protected water source increases over time in both sub-counties, but most dramatically in Bagezza.

Results from the section on heterogenous impacts are consistent with this explanation. For instance, we find that distance to water source is affected by the baraza intervention, but only if we restrict the sample to households that live 5 kilometres or more from the sub-county district headquarters. Households living close to the headquarters may already have good access to water, so a baraza may not affect their situation. Failure to take this into account could lead to the conclusion that barazas do not influence access to water.

## 6.2 Policy and programme relevance: evidence uptake and use

At the time of writing this study, we had already presented preliminary findings at the National Monitoring and Evaluation Technical Working Group Workshop on 13 March at the OPM in Kampala, Uganda. The response was very encouraging and the OPM was pleased with the work so far. There were 80 attendees at the meeting, with representatives from the OPM, Kampala City Authority, Economic Policy Research Centre of Makerere University, National Planning Authority, Sustainable Development Goals Secretariat, Ministry of Finance, Planning and Economic Development, Ministry of Gender, Labour and Social Development and about 10 other government authorities/departments. It is clearly too early to know how the evidence generated will be used, but we do feel the OPM holds the study in high esteem.

## 6.3 Challenges and lessons

The mixed results are puzzling, especially given the fact that qualitative research prior to endline data collection suggested real effects from the intervention (Van Campenhout et al. 2018). We suspect that the lack of impact in the quantitative part of the study may be due to the nature of the intervention. Different sub-counties face different challenges, which is reflected in what transpires at the baraza event. For instance, in districts where there are issues related to water, the baraza will mainly revolve around poor service delivery in the infrastructure sector and ways to improve this. In these sub-counties, barazas could affect service delivery in infrastructure, but leave outcomes in other sectors unaffected. In other sub-counties, problems may concentrate in the agricultural sector and impact on infrastructure might be minimal. In other words, the true treatment citizens receive may become hard to discern and could, in fact, be far from the standardised treatments given in RCTs by biophysical scientists. As a result, a focus on the average treatment effect may fail to identify a significant effect, because the impact is averaged over many sub-counties that, in reality, received a ‘different’ type of baraza. Heterogeneity in the treatment will also introduce selection bias, because barazas will tend to focus on areas that are the most problematic (Barrett and Carter 2010). Issues related to non-standardised treatments are confirmed when looking at heterogeneous treatment effects and a case study of access to water in Bagezza sub-county.

## 6.4 Limitations

### *6.4.1 Partial rollout, selection bias and matching*

The primary limitation of this study is that endline data collection after partial rollout might have introduced selection bias. It is possible that from the randomly assigned sub-counties, particular sub-counties were selected to be treated first and the treatment of other sub-counties was postponed. For example, the implementing partner may have started with sub-counties that are close to the capital for logistical reasons, or the OPM may have treated politically favoured sub-counties first and other sub-counties later for political reasons. If sub-counties were selected for logistical or political reasons, or due to other socioeconomic characteristics, treatment is not random, our sample is not representative of the population we intended to analyse and some conclusions of this study may not be correct.

That is why we matched, ex ante, each treated sub-county to a control sub-county that is similar with regard to a range of observable pre-treatment characteristics. We used a range of sub-county characteristics that were likely to be known to OPM staff and may have affected how the intervention was rolled out. These characteristics are used in a probit regression to predict the likelihood that a sub-county was treated. For each treated sub-county, we then matched a potential control sub-county with a likelihood of being treated that is similar to that of the treated sub-county.

Classic matching attempts to reproduce the treatment group among the non-treated to re-establish experimental conditions in a non-experimental setting and relies on observable variables to account for selection. We, on the other hand, dealt with experimental conditions in an experimental setting, even though parts of our experiment were not implemented as planned. Our matching did not attempt to reproduce the treated among the non-treated, but to select matching controls for our treated sub-counties. However, the aim of our and the classic matching methods is equal: line up comparison individuals according to sufficient observable characteristics to remove systematic differences in the evaluation outcome between treated and non-treated, so that the only remaining difference between the two groups is the treatment.

That is why both methods are subject to similar limitations. One main limitation of matching is related to data availability. We cannot be sure that the missing counterfactual, the matching control sub-county, exists in our sample. Some observations might not be matched perfectly, so the estimated parameter is difficult to interpret (Blundell and Dias 2009). Another key limitation of matching is related to our ability to select the right information. We must observe and select the right characteristics to ensure that the unexplained share of the outcome is not related to the treatment decision. Heckman and Navarro (2004) show on one hand how important and on the other how difficult it is to select the appropriate set of variables for matching. If the conditioning set of variables is not right and complete, our estimates are biased (Blundell and Dias 2009). However, if observations are matched well and the right information is used, matching deals well with potential bias.

We acknowledge that the partial rollout is a threat to internal validity that should not be ignored. At the same time, the fact that we started from a cluster RCT design provides a substantial advantage over studies that are based on observational data. For instance, potential selection emanating from partial rollout is restricted to the sub-sample of sub-counties that were assigned to the treated group only, significantly reducing the scope for bias. In addition, we organised the list of sub-counties to be treated that we shared with the implementing partner by treatment group (information, deliberation and combined treatment) and in each treatment group, we listed sub-counties alphabetically. Looking at the list in light of the partial rollout, we get the impression the OPM started at the top of the list and worked its way down. As a result, relatively more sub-counties that were assigned the information treatment were treated and sub-counties towards the top of the list are more likely to be treated. This pattern is confirmed when regressing the likelihood of being treated on the rank of the sub-county. This also suggests that OPM officials did not deliberately select certain sub-counties.

### *6.4.2 Gender*

Our impact evaluation was not designed to answer particular questions about gender and this is a limitation of the study. While women and men were equally able to participate in barazas, we do not have data on who came to the events and cannot say whether women and men were equally present. However, our facilitators were trained to encourage paying attention to the voice of women and minorities during meetings.

As women's priorities might have been raised and addressed less/more often in meetings, we compared female and male perceptions and prioritisations. Enumerators were instructed to interview household heads, so that in our baseline data, 4,714 (38%) were female respondents and 7,831 (62%) were male respondents. Conditioning on the gender of the respondent, we do not find differences between women's and men's priorities for 10 of 14 statements. Using the unpaired two-samples t-test to compare the means of women and men, we found that men's perceptions were significantly different from women's for the following statements:

a. Staff at government health centres or hospitals are rude to patients.
b. Children's learning outcomes are poor.
c. Agricultural inputs supplied by the government are of poor quality.
d. Agricultural extension agents are not aware of enterprises or agricultural inputs relevant for farmers.

For all four statements, women agreed significantly less than men, indicating that they were less concerned about the issues. Because we cannot find an issue women prioritise more than men (about which women are significantly more concerned than men), we cannot study whether such an issue received more or less attention during and after the baraza. However, these differences and the lack thereof do not align with our expectations. It seems strange that women do not prioritise any issue more than men. There are several potential explanations for this surprising finding.

First, prior studies in different contexts find that women respond in a more socially desirable fashion than men (Bernardi 2006; Chung and Monroe 2003; Hebert et al. 1995). Social desirability is the tendency of an individual to avoid criticism and convey an image in line with social norms (Hebert et al. 1995). The social desirability response bias refers to the tendency of individuals to overreport socially desirable aspects and underreport undesirable aspects (Zerbe and Paulhus 1987). If respondents perceive it as socially desirable not to agree with statements that point to problems in public service delivery (i.e. not to criticise the government), this bias could explain why women responded to agree less than men.

Second, selection bias could drive the result: the women in our sample are often single household heads who are not necessarily representative of the rest of the female population.

Third, this result could indicate that the way we define women's priorities and perceptions is debatable. However, our study was not designed to answer particular gender-related questions and, therefore, we lack a better way to find out what women perceive to be important. That is why we cannot sufficiently test whether issues women prioritise are more or less likely to be addressed during and after barazas.

# 7. Conclusions and recommendations

To improve governance and public service delivery, the GoU organises community advocacy forums – popularly known as barazas – where citizens receive information from government officials and get the opportunity to directly engage with them. In 2015, we designed a study aimed at evaluating the effectiveness of these forums. The evaluations set out to answer four research questions:

1. What is the impact of the baraza as implemented by the OPM?
2. What is the relative effectiveness of the information component of a baraza?
3. What is the relative effectiveness of the deliberation component of a baraza?
4. What is the impact of district-level barazas?

Baseline data on more than 12,500 households spread over almost 250 sub-counties in about 40 districts throughout Uganda were collected and the OPM started implementing barazas following our protocol.

The OPM faced various complications that affected the timely rollout of the barazas, including budgetary constraints and disruptions related to the general election of 2016. This resulted in the decision to collect endline data after partial rollout. Various strategies were followed to diagnose and reduce the consequences of potential selection bias introduced by this partial rollout.

To answer the four questions mentioned above, we analysed a set of carefully selected variables, declared in a preregistered analysis plan and combined in indices. In this confirmatory analysis, we focus on five indices corresponding to the four main sectors – agriculture, infrastructure, health and education – and one overall index. We do not find a significant impact of the baraza programme on overall public service delivery. There are some indications that sub-county-level barazas affected the agricultural sector, but the difference is only significant at the 10 per cent level.

While we do not find that the baraza programme affects public service delivery in general, we do find a variety of interesting effects when we look beyond the indices and analyse individual outcomes. In this second part of our analysis, which is more exploratory in nature, we find that in the agricultural sector, sub-county-level barazas significantly expand access to extension. We also see an increase in the likelihood that farmers received improved seeds from the government. This is consistent with the positive effect sub-county-level barazas seem to have on the likelihood that farmer associations or groups are formed in a village and an increase in the number of such institutions that are assisted by NAADS/Operation Wealth Creation. There are also some improvements in public school infrastructure after a sub-county-level baraza and a small reduction in waiting time at the water source.

We assess whether citizens interact more with politicians and service providers due to baraza meetings and find mixed results. The baraza intervention does not affect the likelihood that citizens participate in elections. However, the information component of a sub-county-level baraza reduces in-kind contributions, yet increases cash contributions. Further, the baraza interventions changed citizens' perceptions of a range of problems.

The lack of significant impact of barazas on public service delivery indices surprises us, especially because qualitative research prior to endline data collection suggested real

effects of the intervention (Van Campenhout et al. 2018). That is why we have investigated potential explanations and run a series of robustness checks. We find a variety of interesting effects when considering heterogeneity in the treatment effects. First, the slow rollout of the baraza programme introduces sub-county heterogeneity in the time that passed between treatment and endline data collection. Our analysis suggests that the effects of the interventions dissipate over time as enthusiasm fades, plans are abandoned and promises forgotten. Second, we asked sub-county officials whether they remembered that a baraza took place, used this variable to check for heterogeneous treatment effects and found generally larger impacts of the intervention. Officials who recalled the baraza might be intrinsically more motivated or, alternatively, the information we received from the OPM might be inaccurate and some sub-counties that were not treated were indicated as being treated. Third, we consider heterogeneity related to remoteness, because barazas may have different effects on households that live further from the sub-county headquarters. There are indications that the intervention is particularly effective for households in more remote locations. Fourth, we reran the analysis with the subset of households that indicated they were aware of the concept of barazas. The results suggest that, in particular, households that are well-informed and interested in public service provision cash in on barazas.

These mixed results are puzzling, especially because previous qualitative research suggests a real impact of the baraza programme (Van Campenhout et al. 2018). The nature of our interventions might be an explanation for this. Different sub-counties face different challenges, so different issues are discussed during baraza events. As a result, subjects may receive different treatments and not standardised treatments comparable to the ones given in RCTs in biophysics. That is why a focus on average treatment effects may fail to find significant effects, because the impact is averaged over many sub-counties that, in reality, received ‘different’ barazas. A case study in Bagezza sub-county confirms these issues related to non-standardised treatments.

Because barazas are designed to affect a broad range of public service outcomes, comparing the cost-effectiveness of the different types of barazas does not result in one straightforward conclusion. For some outcomes, sub-county-level interventions seem to be more cost-effective than district-level interventions, whereas for others the opposite holds. Similarly, in some cases, information barazas are more cost-effective, while in other cases deliberation barazas are. However, baraza interventions have an impact on many households and are inexpensive, so the rate of return is substantial even if treatment effects are small.

Taking into account the impact of the baraza intervention on individual public service delivery outcomes, the heterogeneity in the treatment effects and our concerns regarding non-standardised treatments, we recommend baraza meetings even though they do not have a measurable effect on our preregistered indices. We do not conclude that sub-county-level barazas are more effective than district-level barazas or the other way around. Barazas at the sub-county level seem to have a larger effect on some outcomes, while barazas at the district level appear to affect other variables. We do not find many significant impacts of district-level barazas, but this might be due to insufficient statistical power. We would, therefore, recommend a mix of both approaches. Furthermore, both the information and the deliberation components of a sub-county-level baraza seem to be important. We hence recommend the implementation of full barazas, especially

because they are only slightly more expensive than information and deliberation barazas. Finally, because our heterogeneity analysis suggests that the effects of the interventions dissipate over time, baraza meetings should not be held only once but several times; for instance, every two years.

The primary limitation of this study is that endline data collection after partial rollout might have introduced selection bias. We acknowledge that this is a threat to internal validity, but believe that the fact that we started from a cluster RCT design still provides a substantial advantage over studies that are based on observational data. We also provide an extensive investigation into the possibility that the results – or lack thereof – are driven by selection bias.

# Appendixes

Note: the tables in the appendixes show the district-level baraza versus sub-county-level baraza comparison and not the district-level baraza versus no baraza (control) comparison.

## Descriptive statistics and balance tables

In *Table A.1*, we test for balance between the treatment groups at baseline following the initial design of the experiment. 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 six household members and about 30 per cent of sampled households live in a house with a thatched grass roof. In the second column, we report differences between baseline characteristics of households that will receive a sub-county-level combined information and deliberation baraza, and those that will not be exposed to any baraza. We cannot reject the null hypothesis that households in these two groups are similar for all but one of the characteristics in *Table A.1*. We do find that, at baseline, households assigned to a sub-county-level baraza live further from the nearest all-weather road and this difference is significant at the 5 per cent significance level. When comparing households that were exposed to a sub-county-level information baraza to households that did not receive a sub-county-level information baraza (third column), we see that that households are slightly larger in the former group and the difference is significant at the 5 per cent level. The average household has 2–3 children attending a public school. We also find a slight pre-treatment imbalance on this outcome for the information treatment, but the difference is only significant at the 10 per cent level.

In the fourth column of *Table A.1*, we report differences between households that were exposed to a sub-county deliberation baraza and households that were 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 were exposed to a district-level baraza and households that were exposed to a sub-county-level baraza that combined both information and deliberation components. We see that household heads in the first group are slightly older than in the latter group. Furthermore, the share of households that report there is a VHT in their village is also slightly higher in the treatment group. In both cases, judged by the cluster robust standard errors, the differences are significant at a 10 per cent level. However, it is well known that when the clusters are few in number (say, 30 or less), the cluster robust standard error is downwardly biased and tends to over-reject the null of no effect. Indeed, we find that the differences are not significant when randomisation inference is used. Overall, out of 40 comparisons, we find that 2 differences are significant at the 5 per cent level and 1 is significant at the 10 per cent level, which is what one would expect to find due to chance alone. As such, we conclude that the initial randomisation was successful.

In *section 4.3*, we provide additional balance tests to investigate whether the partial rollout of the intervention introduced selection bias.

During both baseline and endline, we collected some data at a more aggregate level. We visited sub-county headquarters and interviewed one politician and one civil servant. For completeness, we also provide a balance table for these data. Results are in *Table A.2*.

Despite the small sample size, the various sub-groups seem to be balanced on a range of characteristics.

**Table A1: Orthogonality tests**

| | Mean | Sub-county baraza | Information | Deliberation | Jurisdictional tier |
|---|---|---|---|---|---|
| Household size | 6.324 | 0.021 | 0.304* | -0.003 | 0.246 |
| | | (0.142) | (0.133) | (0.125) | (0.248) |
| Age of the household head (years) | 46.501 | 0.736 | 0.464 | 0.725 | 1.427 |
| | (14.615) | (0.681) | (0.594) | (0.714) | (0.802) |
| Head of household is woman (1=yes) | 0.191 | 0.012 | -0.014 | 0.004 | -0.013 |
| | (0.393) | (0.014) | (0.013) | (0.015) | (0.016) |
| Head finished primary education (1=yes) | 0.213 | -0.007 | -0.02 | -0.003 | -0.026 |
| | (0.410) | (0.017) | (0.020) | (0.020) | (0.027) |
| Thatched grass roof (1=yes) | 0.298 | -0.001 | 0.009 | -0.032 | 0.011 |
| | (0.457) | (0.026) | (0.025) | (0.023) | (0.036) |
| Traditional mud wall (1=yes) | 0.424 | 0.021 | -0.025 | 0.038 | -0.034 |
| | (0.494) | (0.043) | (0.040) | (0.039) | (0.104) |
| Distance to nearest all-weather road (km) | 0.906 | 0.167* | 0.106 | 0.147 | -0.192 |
| | (0.915) | (0.106) | (0.095) | (0.092) | (0.138) |
| Access to extension (1=yes) | 0.108 | 0.002 | 0.004 | 0.007 | 0.009 |
| | (0.310) | (0.014) | (0.012) | (0.014) | (0.017) |
| Village health team in village (1=yes) | 0.854 | 0.000 | 0.006 | 0.025 | 0.07 |
| | (0.353) | (0.031) | (0.026) | (0.026) | (0.036) |
| Number of children in public schools | 2.478 | 0.044 | 0.165+ | 0.038 | 0.139 |
| | (2.074) | (0.095) | (0.091) | (0.089) | (0.155) |
| Number of observations | 12,545 | 5,193 | 10,241 | 10,241 | 4,949 |

Note: column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports differences (and standard errors below) between baseline characteristics of households that received a district-level combined information and deliberation baraza, and those that did not receive any baraza; **, * and + denote significance at the 1%, 5% and 10% levels, respectively.

**Table A2: Balance table for sub-county-level data**

| | Mean | Sub-county baraza | Information | Deliberation | Jurisdictional tier |
|---|---|---|---|---|---|
| Frequency of executive committee meetings | 0.983 | -0.03+ | 0.00 | -0.03+ | 0.003 |
| | (0.128) | (0.016) | (0.002) | (0.016) | (0.02) |
| Proportion of health budget that has not been received | 17.236 | -6.836 | -3.403 | 1.168 | -0.726 |
| | (26.086) | (4.475) | (4.807) | (4.796) | (4.913) |
| Lengths of other all-weather roads | 65.244 | -14.968 | -16.807 | -13.466 | -8.139 |
| | (69.357) | (12.366) | (10.76) | (12.133) | (11.795) |
| Proportion of households with electricity | 17.154 | 4.259 | 2.132 | 1.862 | -5.823 |
| | (19.552) | (3.349) | (2.556) | (3.25) | (4.192) |
| Number of male crop extension staff/agents | 0.913 | 0.259 | -0.08 | 0.031 | -0.098 |
| | (0.583) | (0.167) | (0.119) | (0.136) | (0.076) |
| Proportion of households using improved seeds | 41.293 | -0.964 | 0.167 | 2.171 | -5.102 |
| | (26.748) | (3.534) | (3.889) | (3.758) | (3.314) |
| Number of HC2s | 3.428 | 0.039 | 0.458 | -0.3 | -0.846 |
| | (3.34) | (0.757) | (0.777) | (0.63) | (0.604) |
| Number of nurses/nursing assistants in-place in HC2s | 6.015 | 0.322 | 1.262 | 0.924 | 3.003** |
| | -3.734 | (0.7) | (0.983) | (0.74) | (0.839) |
| Student enrolment in government secondary schools | 733.866 | 29.312 | 178.374 | 106.592 | -72.357 |
| | (694.694) | (142.906) | (178.474) | (145.499) | (97.122) |
| Number of observations | 262 | 102 | 168 | 168 | 102 |

**Note:** HC2 = health centre 2; column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively.

**Table A3: Impact of barazas on participation in elections**

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| Hold any political/traditional positions? | 0.303 | 0.018 | -0.028 | -0.017 | -0.051* |
| | (0.46) | (0.019) | (0.021) | (0.03) | (0.023) |
| Voted in LC1 elections? | 0.926 | -0.014 | 0.000 | 0.011 | -0.017 |
| | (0.261) | (0.017) | (0.013) | (0.014) | (0.019) |
| Voted in LC3 elections? | 0.884 | 0.025 | 0.016 | 0.027 | -0.014 |
| | (0.32) | (0.028) | (0.02) | (0.024) | (0.031) |
| Voted in LC5 elections? | 0.898 | -0.002 | 0.011 | 0.004 | 0.008 |
| | (0.302) | (0.024) | (0.017) | (0.025) | (0.029) |
| Voted in the presidential election? | 0.932 | -0.008 | 0.003 | 0.003 | 0.018 |

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| | (0.252) | (0.018) | (0.012) | (0.016) | (0.018) |
| Voted in parliamentary elections? | 0.922 | -0.011 | 0.001 | 0.002 | 0.032 |
| | (0.269) | (0.022) | (0.014) | (0.019) | (0.027) |
| Voted in party leaders' elections? | 0.752 | -0.01 | -0.043 | -0.01 | -0.003 |
| | (0.432) | (0.039) | (0.033) | (0.039) | (0.056) |
| Political Participation Index | 0.000 | 0.006 | -0.034 | 0.004 | -0.043 |
| | (0.646) | (0.053) | (0.041) | (0.049) | (0.066) |
| Number of observations | 6,700 | 2,390 | 4,266 | 4,266 | 2,379 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively.

## Sub-county-level analysis

In addition to household surveys, we conducted surveys with government officials as respondents. Obviously, sample sizes are much smaller here, so results should be interpreted with this caveat in mind. In each sub-county, we interviewed two officials: the highest-ranking politician (the LC3) and the highest-ranking civil servant (sub-county chief). Sometimes, the deputy was interviewed. We have 261 observations in this dataset.

*Agriculture*

We start again with agriculture. We report results in *Table A.4.1*. As in previous tables, the first column shows sample averages, with standard deviations in brackets below. In the second column, we report differences in outcomes between sub-counties that received a typical sub-county-level baraza and sub-counties that did not receive any baraza. In the third column, we report differences between outcomes of sub-counties where an information baraza was organised and outcomes of sub-counties that were not exposed to an information baraza. In the fourth column, we report differences between outcomes of sub-counties where a deliberation baraza was organised and outcomes of sub-counties that were not exposed to a deliberation baraza. Finally, in the fifth column we directly compare sub-counties that received a sub-county-level baraza to sub-counties that were exposed to a district-level baraza.

Government officials report that, on average, 14.3 per cent of the agricultural budget was not received. We do not find evidence that the baraza intervention affected this percentage.

We then look at officials' perceptions of problems in the agricultural sector. Over the past year, officials received on average 2.9 complaints related to agricultural service provision. The number of complaints seemed to reduce after a sub-county-level baraza took place. As in the household questionnaire, officials were also asked to rate their agreement with various statements. We do not find that barazas affect perceptions on

input quality. However, officials in sub-counties with deliberation barazas report that there is increased transparency in how farmers are selected to receive agricultural inputs. Officials in sub-counties with information barazas agree less with the assertion that extension agents rarely visit. Officials in sub-counties with deliberation barazas are also more of the opinion that extension agents are aware of what their customers want. The above seems to suggest perceptions became more positive in the agricultural sector after sub-county-level barazas, but it is unclear whether it is the information or deliberation component that is driving this result.

Turning to outcomes, we investigate the effect of barazas on access to extension at home, as reported by officials. Recall that when analysing household data, we found that significantly more households in areas that received a sub-county-level baraza were visited by an expert at home. Analysing the responses of government officials, we find that the number of male crop extension agents is about one person higher in areas where a deliberation baraza took place. There is also a significant difference in the number of male crop extension agents when directly comparing sub-county- to district level barazas, with more staff available after a district-level baraza. The baraza intervention did not affect the number of female crop extension staff/agents. We also find substantial reductions in the number of demonstration sites as a result of sub-county-level barazas. This is surprising, given that in the household-level data, there is some evidence of increased visits to extension offices, demonstration sites and model farmers, especially after matching to reduce potential bias introduced by the partial rollout (*Table A.5*).

Looking at the use of modern inputs, we find that for both fertilisers and improved planting material, there is a negative and significant difference between areas exposed to a district-level baraza and areas exposed to a sub-county-level baraza. We also see that the percentage of households in the sub-county that reportedly used improved seed or fertiliser is higher in areas that received a sub-county baraza than in control sub-counties, but the difference is not significant.

Using household survey data, we find that the proportion of households that received improved seed from the government extension system is significantly higher in areas where a sub-county-level baraza took place. Asking government officials about the frequency of improved seed distribution, we do not find a significant difference between sub-counties with a sub-county-level baraza and control sub-counties. However, we do see that the frequency of improved seed distribution is about 0.4 higher in areas where an information baraza took place. We also see that the frequency of improved breeds of cattle, goat, pig and poultry distribution is higher in areas that were exposed to a sub-county-level baraza. Grievances related to the distribution of seed and livestock (goats and milk cows) were often encountered during qualitative work.

**Table A4: Impact on agriculture (sub-county-level analysis)**

| | Mean | Sub-county baraza | Information | Deliberation | Jurisdictional tier |
|---|---|---|---|---|---|
| **Political effort** | | | | | |
| Agricultural budget that has not been received (in %)‡ | 14.284 | -3.446 | -1.991 | 7.389 | -6.375 |
| | (29.046) | (6.443) | (5.41) | (9.309) | (3.906) |
| **Perception** | | | | | |
| Number of complaints‡ | 2.945 | -2.874+ | -1.565 | -1.44 | 0.824 |
| | (7.686) | (1.636) | (1.751) | (1.781) | (0.81) |
| 'Agricultural inputs supplied by the government are of poor quality.' | 5.669 | 0.254 | -0.24 | 1.088 | -0.313 |
| | (3.179) | (0.666) | (0.659) | (0.867) | (0.667) |
| 'Lack of transparency in how farmers are selected to receive agricultural inputs from govt.' | 5.225 | -0.76 | 0.054 | -1.423+ | 0.201 |
| | (3.244) | (0.611) | (0.684) | (0.781) | (0.886) |
| 'Agricultural extension agents rarely visit.' | 5.199 | 0.895 | -1.355+ | -1.063 | -0.313 |
| | (3.194) | (0.665) | (0.76) | (0.766) | (0.469) |
| 'Agricultural extension agents are not aware of needs relevant to farmers.' | 4.483 | -0.125 | -0.792 | -1.527+ | -0.402 |
| | (3.069) | (0.557) | (0.72) | (0.799) | (0.41) |
| **Outcomes** | | | | | |
| Number of male crop extension agents‡ | 1.123 | -0.207 | 0.248 | 1.047+ | 0.205* |
| | (1.233) | (0.133) | (0.309) | (0.517) | (0.089) |
| Number of female crop extension agents‡ | 0.36 | 0.382 | 0.206 | 0.357 | -0.321 |
| | (1.049) | (0.236) | (0.235) | (0.39) | (0.24) |
| Number of demonstration sites‡ | 3.157 | -2.249* | -1.276 | -2.461* | 2.977* |
| | (6.235) | (1.029) | (1.257) | (1.071) | (1.244) |
| Households using purchased fertilisers (in %)‡ | 28.11 | 5.129 | -1.79 | -8.809 | -19.808* |
| | (29.98) | (7.588) | (7.367) | (7.905) | (7.195) |
| Households using improved seeds (in %)‡ | 44.136 | 5.288 | 5.393 | -2.399 | -14.907* |
| | (29.9) | (5.316) | (6.45) | (7.193) | (5.63) |
| Households using pesticides/herbicides/fungicides (in %)‡ | 46.195 | 3.096 | 7.286 | -9.217 | -5.65 |
| | (29.289) | (6.205) | (5.916) | (7.779) | (7.302) |
| Households using improved livestock breeds (in %)‡ | 23.131 | 2.499 | 0.394 | -7.672 | -15.533 |
| | (25.196) | (10.283) | (6.568) | (6.279) | (11.049) |
| Frequency of improved seed distribution‡ | 1.932 | 0.000 | 0.393+ | 0.496 | 0.879* |
| | (1.776) | (0.223) | (0.219) | (0.293) | (0.378) |
| Frequency of improved breed distribution | 1.053 | 0.602+ | 0.097 | 0.492 | -0.668* |
| | (1.112) | (0.33) | (0.179) | (0.328) | (0.282) |
| Frequency of fertiliser/manure distribution | 0.524 | -0.1 | 0.284 | -0.215 | 0.316 |

| | Mean | Sub-county baraza | Information | Deliberation | Jurisdictional tier |
|---|---|---|---|---|---|
| | (1.399) | (0.178) | (0.542) | (0.201) | (0.291) |
| Frequency of pesticide/herbicide/fungicide distribution | 0.557 | 0.634 | -0.036 | -0.174 | -0.447 |
| | (1.206) | (0.389) | (0.156) | (0.185) | (0.408) |
| Number of observations | 262 | 102 | 168 | 168 | 102 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; ‡ indicates that missing observations were interpreted as zero; † indicates that we did not control for the baseline value.

*Infrastructure*

We now turn to infrastructure-related outcomes as reported by sub-county officials (*Table A.4.2*). We start with perception. For infrastructure, we also recorded the number of complaints. We find a reduction of -5.1 water-related complaints after a sub-county-level baraza. It seems that the deliberative component is the main driver behind this result. Moreover, government officials were asked to report their agreement with two water infrastructure-related statements. When asked whether '*Access to a drinking water source is a serious problem*', government officials in sub-counties that received a sub-county-level baraza agreed significantly more with this statement. We find a similar effect for the statement '*Drinking water is usually dirty*'. Perhaps, sub-county-level barazas made officials more sensitive to this issue.

We also include some questions related to road infrastructure. Looking at the household data, we do not find that the baraza programme reduces the average distance of households to the nearest all-weather road. This is in line with our findings from surveying government officials. However, when officials were asked to report their agreement with the statement '*Availability/Access to all-weather roads is a serious problem*', officials in sub-counties that received a typical sub-county-level baraza agreed significantly less with this statement.

In the household-level analysis, we learned that the difference in distance to the primary water source during the dry season is never significantly different from zero, but that there is a significant reduction in the time one has to wait at the water source in areas that were exposed to a sub-county-level baraza intervention. Government officials reported on a range of different water sources. We find that there are 14 more boreholes in sub-counties where an information baraza was organised. We further find a reduction of the number of protected springs in sub-counties that were exposed to a sub-county-level baraza. At the same time, we see an increase in the number of protected springs in areas that were exposed to a district-level baraza, resulting in a difference of 7.3, which is significant at the 1 per cent level.

The increase in the number of boreholes in sub-counties where an information baraza was organised and the decrease in the number of protected springs in sub-counties that

received a sub-county-level baraza could be the reason the difference in distance to the primary water source of households during the dry season is never significantly different from zero, because the two effects offset each other. Because the positive impact on the number of boreholes is much larger than the negative impact on the number of protected springs, this could be an explanation for the significant reduction in waiting time at the water source households report in areas that were exposed to the sub-county-level baraza intervention.

**Table A5: Impact on infrastructure (sub-county-level analysis)**

| | Mean | Sub-county baraza | Information | Deliberation | Jurisdictional tier |
|---|---|---|---|---|---|
| **Perception** | | | | | |
| Number of complaints‡ | 4.966 | -5.144* | -0.943 | -5.17+ | 1.12 |
| | (9.145) | (2.449) | (2.256) | (2.51) | (1.267) |
| 'Access to a drinking water source | 5.953 | 1.623* | 0.203 | 0.702 | -0.854 |
| is a serious problem.' | (3.057) | (0.719) | (0.688) | (0.817) | (0.697) |
| 'Drinking water is usually dirty.' | 5.025 | 1.729* | -0.164 | 1.166 | -2.373* |
| | (3.116) | (0.687) | (0.675) | (0.914) | (0.85) |
| 'Availability/Access to all-weather | 6.784 | -1.137+ | -0.502 | 0.177 | -0.296 |
| roads is a serious problem.' | (2.774) | (0.558) | (0.542) | (0.648) | (0.887) |
| **Outcomes** | | | | | |
| Lengths of tarmac roads‡ | 3.393 | -4.317 | -4.182 | -5.61+ | -0.988 |
| | (9.545) | (2.817) | (2.578) | (3.257) | (0.678) |
| Lengths of other all-weather | | | | | |
| roads‡ | 50.255 | 8.485 | 1.654 | 40.143 | 6.108 |
| | (74.271) | (11.343) | (10.359) | (30.295) | (12.236) |
| Number of boreholes‡ | 16.763 | 2.601 | 14.467** | 7.371 | 1.133 |
| | (26.371) | (2.843) | (5.079) | (4.808) | (3.602) |
| Number of protected springs‡ | 9.275 | -5.287+ | 1.943 | 1.918 | 7.334** |
| | (19.549) | (2.767) | (3.304) | (4.727) | (2.1) |
| Number of protected dug/shallow | | | | | |
| wells‡ | 3.585 | -2.046 | -0.152 | 1.824 | 2.462+ |
| | (7.965) | (1.729) | (1.397) | (3.116) | (1.268) |
| Number of unprotected | | | | | |
| dug/shallow wells‡ | 8.513 | -0.094 | 1.885 | 8.02 | -4.073 |
| | (18.575) | (3.589) | (3.569) | (7.406) | (3.315) |
| Number of piped/gravity flows† | 12.234 | 7.473 | 2.04 | -4.963 | -18.955+ |
| | (37.886) | (11.924) | (13.147) | (7.795) | (10.346) |
| Number of observations | 262 | 102 | 168 | 168 | 102 |

**Note:** Column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; ‡ indicates that missing observations were interpreted as zero; † indicates that we did not control for the baseline value.

*Health*

We now study health-related outcomes and report them in *Table A6*. We find that, on average, 14.1 per cent of the health budget has not been received and this proportion is independent of treatment groups. While we see that the number of health-related complaints reduces for most comparisons, the only significant difference is the one between sub-counties with a deliberation baraza and sub-counties without, the former receiving on average two complaints less. For perception, absenteeism seems to be less of a problem in areas that experience a district-level baraza than in areas that were exposed to a sub-county-level baraza.

We then investigate outcomes related to VHTs. Using the household data, we find that the share of households that report a VHT is present in their village and the likelihood that individuals participate as VHT members is not affected by the baraza intervention. Responses of government officials are in line with household responses. In fact, the number of VHT members reduces by 14.6 in sub-counties with a sub-county-level baraza. The number of VHT members is independent of the other treatment groups.

Using household data, we find that access to public health facilities was independent of the treatment groups. Also, we do not find that barazas reduced the distance to the nearest government health facility. Here, we look at the number of health centres in the sub-counties as reported by the officials. The number of health centres 2 (HC2s) is not significantly different for comparisons of the different sub-county-level barazas. However, when directly comparing sub-counties that received a sub-county-level baraza to those that were exposed to a district-level baraza, we do find a significant increase of 0.3 centres. Furthermore, the number of health centres 3 (HC3s) increases by 0.3 in sub-counties that were exposed to a sub-county-level baraza. Both are significant at the 5 per cent level.

After looking at the number of HC2s and HC3s, we want to take a closer look at the situation inside these public health facilities. We start with staffing in HC2s, differentiating between clinical officers, nurses and birth attendants. Interestingly, we find that nurses and birth attendants are present in higher numbers after a district-level baraza than after a sub-county-level baraza. In fact, there are some indications that a sub-county-level baraza leads to a reduction in staff.

Equipment in HC2s matters for service delivery. The number of HC2s with a safe drinking water source, laboratory tests, a medical waste pit, or HIV/AIDS guidance and counselling services is independent of the treatment groups. In sub-counties that received a sub-county-level baraza, the number of HC2s with immunisation facilities, the number of HC2s with outpatient services and the number of HC2s with family planning services all increase, while the number of HC2s with in-patient care reduces. Administrative placement also seems to matter for equipment in HC2s. If we compare sub-counties that received a sub-county-level baraza to sub-counties that were exposed to a district-level baraza, there are 0.1 more HC2s with electricity, 0.4 percentage points more with staff houses for all relevant employees and 0.2 more with in-patient care in areas that were exposed to a higher-level baraza.

We also look at staffing in HC3s. The number of doctors, clinical officers, medical assistants, nurses, nursing assistants and laboratory technicians on payroll in HC3s are

all independent of the treatment groups. The number of midwives on payroll in HC3s is lower in sub-counties that received a district-level baraza compared to sub-counties that were exposed to a sub-county-level baraza. The number of in-patient care beds in HC3s, and the number of HC3s with electricity and with a mortuary/cold room are independent of the treatment groups. In sub-counties that received a sub-county-level baraza, HC3s are more likely to have a safe drinking water source, provide laboratory tests, provide immunisation services, provide more outpatient services and provide more family planning services and antenatal care. Sub-counties that received a baraza are also more likely to have appropriate medical waste disposal facilities and offer HIV/AIDS guidance and counselling services. These outcomes seem to be mostly driven by the information component.

**Table A6: Impact on health sector (sub-county-level analysis)**

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| **Political effort** | | | | | |
| Proportion of health budget that has not been received‡ | 14.089 | -1.429 | 8.521 | 5.522 | -6.923 |
| | (28.051) | (6.472) | (6.868) | (9.498) | (4.753) |
| **Perception** | | | | | |
| Number of complaints‡ | 1.75 | -1.641 | -1.016 | -1.986+ | 0.545 |
| | (5.719) | (1.363) | (1.285) | (1.115) | (0.623) |
| 'Access to a government health centre or hospital is a serious problem.' | 6.64 | -0.426 | 0.194 | -0.761 | -0.378 |
| | (3.026) | (0.895) | (0.62) | (0.657) | (0.781) |
| 'Government health centres or hospitals do not have relevant medicines.' | 6.826 | 0.553 | -0.127 | 0.505 | -0.434 |
| | (2.818) | (0.52) | (0.644) | (0.627) | (0.37) |
| 'Staff at government health centres or hospitals are rude to patients.' | 4.394 | 0.797 | -0.047 | -0.664 | -0.597 |
| | (2.821) | (0.605) | (0.61) | (0.66) | (0.739) |
| 'Medical staff at government health centres or hospitals are often absent.' | 4.411 | 0.757 | -0.049 | -0.976 | -1.683* |
| | (2.943) | (0.565) | (0.587) | (0.685) | (0.666) |
| **Outcomes** | | | | | |
| Number of villages with VHTs‡ | 43.225 | -8.924 | -4.327 | -1.829 | -1.831 |
| | (24.644) | (6.351) | (6.282) | (6.739) | (3.768) |
| Number of VHT members‡ | 73.907 | -14.644+ | -5.301 | 11.937 | -2.738 |
| | (40.67) | (7.936) | (8.061) | (11.889) | (10.935) |
| Number of HC2s‡ | 1.097 | -0.124 | 0.112 | 0.14 | 0.34+ |
| | (1.229) | (0.181) | (0.154) | (0.162) | (0.153) |
| Number of clinical officers on payroll in HC2s‡ | 0.343 | -0.174 | -0.052 | 0.024 | 0.135 |
| | (0.925) | (0.22) | (0.148) | (0.262) | (0.121) |
| Number of nurses/nursing assistants on payroll in HC2s‡ | 2.047 | -1.273* | 0.002 | -0.873 | 1.409** |
| | (3.103) | (0.49) | (0.493) | (0.594) | (0.417) |
| Number of birth attendants on payroll in HC2s‡ | 0.445 | -0.32* | 0.088 | 0.615 | 0.269** |
| | (1.142) | (0.146) | (0.165) | (0.478) | (0.072) |
| Number of HC3s‡ | 0.801 | 0.307* | 0.333* | -0.016 | -0.292 |
| | (0.67) | (0.125) | (0.127) | (0.117) | (0.208) |

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| Number of in-patient care beds in HC3s‡ | 5.839 | -0.451 | 0.472 | 0.01 | -1.609 |
| | (7.359) | (1.6) | (1.721) | (1.871) | (1.636) |
| Number of doctors on payroll in HC3s† | 0.063 | -0.048 | -0.015 | -0.044 | 0.053 |
| | (0.312) | (0.096) | (0.097) | (0.118) | (0.066) |
| Number of clinical officers on payroll in HC3s‡ | 1.042 | 0.193 | 0.228 | 0.164 | -0.238 |
| | (1.122) | (0.224) | (0.271) | (0.297) | (0.257) |
| Number of medical assistants on payroll in HC3s‡ | 0.39 | -0.158 | -0.048 | -0.104 | 0.012 |
| | (0.937) | (0.18) | (0.172) | (0.226) | (0.202) |
| Number of nurses/nursing assistants on payroll in HC3s‡ | 2.826 | 0.579 | 1.063 | 0.951 | -0.625 |
| | (3.072) | (0.687) | (0.826) | (0.997) | (0.793) |
| Number of midwives on payroll in HC3s‡ | 1.691 | -0.194 | -0.063 | -0.529 | -0.745** |
| | (2.205) | (0.403) | (0.432) | (0.458) | (0.218) |
| Number of laboratory technicians on payroll in HC3s‡ | 0.936 | 0.497 | 0.339 | -0.236 | -0.508 |
| | (1.126) | (0.309) | (0.251) | (0.227) | (0.293) |
| Number of HC2s with electricity‡ | 0.237 | 0.243 | -0.183+ | 0.151 | 0.136* |
| | (0.533) | (0.24) | (0.104) | (0.156) | (0.056) |
| Number of HC2s with safe drinking water source‡ | 0.564 | 0.738 | 0.122 | 0.246 | -0.394 |
| | (1.084) | (0.47) | (0.21) | (0.251) | (0.266) |
| Number of HC2s with staff houses for all relevant employees‡ | 0.72 | 0.945 | 0.135 | 0.17 | 0.359+ |
| | (1.587) | (0.642) | (0.229) | (0.145) | (0.16) |
| Number of HC2s with laboratory tests‡ | 0.53 | -0.128 | -0.007 | -0.313 | 0.268 |
| | (1.008) | (0.086) | (0.178) | (0.197) | (0.204) |
| Number of HC2s with immunisation facilities‡ | 0.886 | 0.823** | 0.117 | -0.139 | 0.26 |
| | (1.156) | (0.264) | (0.166) | (0.174) | (0.263) |
| Number of HC2s with in-patient care‡ | 0.242 | -0.169+ | -0.071 | -0.089 | 0.235* |
| | (0.712) | (0.088) | (0.11) | (0.107) | (0.103) |
| Number of HC2s with outpatient services‡ | 1 | 0.661* | 0.221 | 0.187 | 0.034 |
| | (1.203) | (0.243) | (0.134) | (0.172) | (0.161) |
| Number of HC2s with family planning services‡ | 0.911 | 0.576* | 0.098 | 0.136 | 0.113 |
| | (1.169) | (0.238) | (0.155) | (0.207) | (0.188) |
| Number of HC2s with medical waste pit‡ | 0.826 | 0.401 | 0.33 | 0.08 | 0.326 |
| | (1.174 | (0.355) | (0.205) | (0.233) | (0.362) |
| Number of HC2s with HIV/AIDS guidance and counselling‡ | 0.856 | 0.375 | -0.018 | -0.305 | 0.092 |
| | (1.169) | (0.32) | (0.176) | (0.216) | (0.226) |
| Number of HC3s with electricity‡ | 0.525 | 0.299 | 0.13 | 0.009 | -0.04 |
| | (0.635) | (0.402) | (0.139) | (0.163) | (0.098) |
| Number of HC3s with safe drinking | 0.568 | 0.997* | 0.216 | 0.005 | -0.116 |

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| water source‡ | | | | | |
| | (0.632) | (0.176) | (0.141) | (0.146) | (0.136) |
| Number of HC3s with staff houses for all relevant employees‡ | 0.525 | -0.025 | -0.203 | -0.226 | -0.289* |
| | (1.128) | (0.752) | (0.266) | (0.228) | (0.103) |
| Number of HC3s with laboratory tests‡ | 0.797 | 0.827** | 0.165 | -0.133 | 0.113 |
| | (0.821) | (0.106) | (0.155) | (0.139) | (0.257) |
| Number of HC3s with immunisation facilities‡ | 0.818 | 0.659** | 0.463* | 0.022 | -0.059 |
| | (0.838) | (0.127) | (0.199) | (0.114) | (0.299) |
| Number of HC3s with in-patient care‡ | 0.674 | 0.037 | 0.29+ | -0.066 | 0.177 |
| | (0.69) | (0.14) | (0.158) | (0.169) | (0.206) |
| Number of HC3s with out-patient services‡ | 0.775 | 0.632* | 0.317** | 0.022 | -0.089 |
| | (0.694) | (0.136) | (0.105) | (0.12) | (0.314) |
| Number of HC3s with family planning services‡ | 0.839 | 0.786* | 0.125 | -0.141 | -0.066 |
| | (0.825) | (0.097) | (0.162) | (0.149) | (0.31) |
| Number of HC3s with antenatal care‡ | 0.792 | 0.659** | 0.275* | 0.000 | -0.032 |
| | (0.687) | (0.127) | (0.105) | (0.119) | (0.289) |
| Number of HC3s with maternity wards‡ | 0.763 | 0.221 | 0.343+ | -0.054 | 0.047 |
| | (0.705) | (0.179) | (0.171) | (0.171) | (0.272) |
| Number of HC3s with placenta pit‡ | 0.725 | 0.608* | 0.158 | -0.038 | 0.104 |
| | (0.669) | (0.135) | (0.104) | (0.134) | (0.248) |
| Number of HC3s with medical waste pit‡ | 0.742 | 0.66* | 0.254* | 0.035 | 0.019 |
| | (0.712) | (0.131) | (0.113) | (0.144) | (0.192) |
| Number of HC3s with HIV/AIDS guidance and counselling‡ | 0.78 | 0.586* | 0.322** | 0.073 | -0.058 |
| | (0.71) | (0.145) | (0.109) | (0.122) | (0.26) |
| Number of HC3s with mortuary/cold room‡ | 0.097 | 0.018 | 0.217 | 0.005 | -0.011 |
| | (0.572) | (0.065) | (0.173) | (0.061) | (0.072) |
| Number of HC3s with isolation room for special cases‡ | 0.191 | 0.136 | 0.29+ | -0.023 | 0.000 |
| | (0.805) | (0.143) | (0.171) | (0.072) | (0.176) |
| Number of observations | 262 | 102 | 168 | 168 | 102 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; ‡ indicates that missing observations were interpreted as zero; † indicates that we did not control for the baseline value.

*Education*

Now we assess the impact of the baraza intervention on education outcomes as reported by sub-county government officials. Like other sectors, about 14 per cent of the budget

has not been received and this proportion did not change as a result of the baraza intervention. The number of education-related complaints also remains stable over sub-groups.

We find that student enrolment in government primary schools and government secondary schools is also not affected by the baraza intervention, which is in line with the household-level analysis. However, there seems to be an effect on dropout rates. The dropout rate for girls in primary schools is 13.2 percentage points lower in sub-counties where an information baraza was organised and 12.1 percentage points lower in sub-counties that received a district-level baraza, compared to sub-counties that were exposed to a sub-county-level baraza. Also, the dropout rate for boys in primary schools is 6.8 percentage points lower in sub-counties with an information baraza than in sub-counties without.

From the household data, we learned that the baraza programme did not have an impact on the distance to a government-operated primary or secondary school. Looking at government officials' data, we see that the baraza programme did not affect the number of government primary schools or the number of government secondary schools, which could explain why the baraza programme did not affect the distance to government-operated schools.

The number of teachers on payroll in government primary schools is not affected by the baraza intervention. The number of teachers on payroll in government secondary schools is: There are, on average, 8.6 more secondary school teachers on payroll in government secondary schools in sub-counties that received a sub-county-level baraza and this effect seems to come mostly from the participation component. Comparing sub-counties that received a sub-county-level baraza to sub-counties that were exposed to a district-level baraza, we find that secondary schools in the latter group had significantly fewer teachers. However, care needs to be taken when interpreting these results because of the limited number of observations.

According to the household-level data, the baraza intervention does not seem to affect whether the school has an SMC. In line with this result, the number of government schools (primary or secondary) with an active SMC is not affected by the baraza intervention, according to government officials.

Finally, government officials were asked their opinion on four problems that stakeholders often mentioned. The intervention does not significantly affect agreement with the statements '*Access to a government primary school is a serious problem*' and '*Children's learning outcomes in government schools are poor*'. When government officials were asked whether '*Teachers in government schools are often absent*', they agreed significantly more in sub-counties that received a sub-county-level baraza. Officials in sub-counties that were exposed to a district-level baraza were less of the opinion that absenteeism was a problem than officials in sub-counties that were exposed to a sub-county-level baraza.

**Table A7: Impact on education sector (sub-county-level analysis)**

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| **Political effort** | | | | | |
| Number of government schools with active SMC‡ | 7.398 | -0.986 | 1.595 | -3.395 | -1.913 |
| | (7.595) | (2.005) | (1.749) | (2.029) | (1.529) |
| Proportion of education budget that has not been received‡ | 13.881 | 1.339 | 6.384 | 4.831 | -2.378 |
| | (27.959) | (6.539) | (5.463) | (7.896) | (3.619) |
| **Perception** | | | | | |
| Number of complaints‡ | 2.411 | -0.247 | 1.190 | -0.24 | -0.441 |
| | (4.495) | (0.929) | (1.139) | (0.883) | (0.644) |
| 'Access to a government primary school is a serious problem.' | 4.225 | -0.380 | 0.203 | -1.054 | -0.496 |
| | (2.996) | (0.708) | (0.683) | (0.895) | (0.441) |
| 'Teachers in government schools are often absent.' | 4.318 | 1.119+ | 0.331 | 0.63 | -1.96** |
| | (2.767) | (0.654) | (0.536) | (0.639) | (0.556) |
| 'Children's learning outcomes in government schools are poor.' | 7.542 | -0.256 | 0.403 | -0.005 | -0.077 |
| | (2.383) | (0.533) | (0.556) | (0.612) | (0.553) |
| **Outcomes** | | | | | |
| Dropout rate for girls in primary schools | 35.045 | 1.085 | -13.228** | -8.808 | -12.104+ |
| | (22.047) | (6.169) | (4.478) | (5.346) | (5.658) |
| Dropout rate for boys in primary schools | 26.247 | -0.088 | -6.827+ | -6.069 | -6.938 |
| | (19.533) | (4.799) | (4.013) | (4.849) | (4.564) |
| Number of government primary schools‡ | 8.737 | -1.099 | -0.443 | 0.194 | -0.747 |
| | (6.31) | (1.008) | (1.039) | (1.09) | (1.095) |
| Student enrolment in government primary schools‡ | 4458.78 | -2860.48 | -531.81 | -2814.12 | 341.71 |
| | (9512.47) | (2952.79) | (2561.43) | (3333.18) | (572.89) |
| Number of teachers on payroll in government primary schools‡ | 58.386 | 6.690 | 17.379 | 9.720 | 5.919 |
| | (60.045 | (11.897) | (15.201) | (16.079) | (9.742) |
| Number of government secondary schools‡ | 0.996 | -0.328 | 0.163 | -0.320 | 0.064 |
| | (1.472) | (0.314) | (0.342) | (0.339) | (0.121) |
| Student enrolment in government secondary schools‡ | 388.453 | 81.974 | 67.362 | 27.858 | -41.100 |
| | (546.069) | (147.351) | (116.486) | (125.191) | (122.726) |
| Number of teachers on payroll in government secondary schools‡ | 8.737 | 8.616* | 0.313 | 6.141+ | -8.896* |
| | (13.826) | (3.963) | (2.633) | (3.487) | (2.969) |
| Number of observations | 262 | 102 | 168 | 168 | 102 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; ‡ indicates that missing observations were interpreted as zero; † indicates that we did not control for the baseline value.

**Table A8: Impact of barazas on agricultural outcomes (matched analysis of variance [ANOVA])**

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| Household used inorganic fertilisers?† | 0.314 | 0.004 | 0.061 | -0.03 | -0.021 |
| | (0.464) | (0.038) | (0.053) | (0.066) | (0.044) |
| Household used improved seed?† | 0.441 | 0.006 | -0.019 | -0.097+ | -0.106+ |
| | (0.497) | (0.043) | (0.058) | (0.048) | (0.052) |
| Received improved seed from govt? | 0.146 | -0.001 | 0.011 | 0.024 | -0.063+ |
| | (0.353) | (0.032) | (0.041) | (0.056) | (0.032) |
| Household used agro-chemicals? | 0.577 | 0.048 | 0.011 | -0.007 | -0.042 |
| | (0.494) | (0.057) | (0.058) | (0.06) | (0.06) |
| Household used improved livestock inputs? | 0.27 | 0.07 | 0.053 | -0.006 | -0.046 |
| | (0.444) | (0.044) | (0.046) | (0.043) | (0.042) |
| Did an agricultural expert visit your home?† | 0.212 | 0.054+ | 0.023 | 0.038 | -0.107** |
| | (0.409) | (0.031) | (0.05) | (0.064) | (0.029) |
| Visited extension office/demo site/model farm? † | 0.306 | 0.077+ | 0.087 | 0.036 | -0.088* |
| | (0.461) | (0.043) | (0.053) | (0.062) | (0.038) |
| Are officials aware of extension demand? | 0.832 | 0.024 | 0.03 | -0.021 | -0.006 |
| | (0.374) | (0.03) | (0.032) | (0.034) | (0.033) |
| Not consulted for extension content? | 0.295 | 0.007 | 0.032 | -0.005 | -0.083+ |
| | (0.456) | (0.032) | (0.042) | (0.048) | (0.04) |
| Are farmer associations/groups in this village? | 0.394 | 0.012 | -0.04 | -0.033 | -0.108+ |
| | (0.489) | (0.052) | (0.036) | (0.042) | (0.056) |
| Farmer groups supported by govt?† | 0.381 | 0.090* | 0.07 | 0.073 | -0.085 |
| | (0.486) | (0.04) | (0.059) | (0.053) | (0.048) |
| Received help in marketing from govt?† | 0.194 | 0.082+ | 0.037 | 0.013 | -0.099* |
| | (0.396) | (0.04) | (0.049) | (0.051) | (0.04) |
| Received help in marketing from coop?† | 0.073 | 0.029 | -0.006 | 0.003 | 0.011 |
| | (0.26) | (0.028) | (0.023) | (0.023) | (0.029) |
| Number of observations | 6,703 | 666 | 1,568 | 1,584 | 1,517 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; † indicates that outcome was included in index.

**Table A9: Impact of barazas on infrastructure (matched ANOVA)**

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| Household uses unprotected water source† | 0.195 | -0.006 | 0.035 | 0.036 | -0.056 |
| | (0.396) | (0.051) | (0.055) | (0.047) | (0.057) |
| Distance to water source (km)† | 0.778 | -0.056 | -0.036 | -0.061 | 0.031 |
| | (0.572) | (0.06) | (0.058) | (0.076) | (0.069) |
| Waiting time at source (min.)† | 3.188 | 0.002 | -0.003 | -0.303 | 0.126 |
| | (1.638) | (0.207) | (0.18) | (0.227) | (0.145) |
| Is there a water user committee in the village?† | 0.579 | 0.043 | 0.056 | 0.007 | -0.005 |
| | (0.494) | (0.06) | (0.046) | (0.057) | (0.061) |
| Is member of water user committee? | 0.168 | 0.080** | 0.018 | 0.063* | -0.047 |
| | (0.374) | (0.026) | (0.029) | (0.025) | (0.035) |
| Water user committee holds public meetings? | 0.431 | 0.028 | 0.051 | 0.044 | 0.012 |
| | (0.495) | (0.057) | (0.047) | (0.056) | (0.051) |
| Satisfied with quality of drinking water? | 0.594 | 0.02 | 0.043 | -0.095 | -0.006 |
| | (0.491) | (0.066) | (0.058) | (0.058) | (0.05) |
| Treat water before drinking (boil or treat)? | 0.593 | -0.005 | -0.08 | -0.024 | -0.028 |
| | (0.491) | (0.054) | (0.059) | (0.055) | (0.071) |
| Distance to nearest all-weather road (km)† | 3.102 | 0.211 | -0.167 | -0.268 | -0.262 |
| Number of observations | 6,703 | 578 | 1,461 | 1,440 | 1,400 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; † indicates that outcome was included in index.

**Table A10: Impact of barazas on the health sector (matched ANOVA)**

| | Mean | Sub-county baraza | Infor-mation | Delibe-ration | Juris-dictional tier |
|---|---|---|---|---|---|
| Seek treatment for fever in public health facility† | 0.696 | -0.018 | 0.015 | 0.007 | 0.061 |
| | (0.460) | (0.046) | (0.044) | (0.051) | (0.076) |
| Go to public health facility to give birth† | 0.828 | -0.007 | -0.013 | -0.016 | -0.046 |
| | (0.377) | (0.048) | (0.039) | (0.043) | (0.069) |
| Is there a VHT in village?† | 0.891 | 0.007 | 0.034 | 0.033 | -0.026 |
| | (0.312) | (0.033) | (0.024) | (0.034) | (0.027) |
| Member of VHT? | 0.127 | 0.007 | 0.001 | 0.002 | -0.052+ |
| | (0.333) | (0.026) | (0.024) | (0.019) | (0.024) |
| VHT organises any public meetings? | 0.415 | 0.090+ | 0.031 | 0.025 | -0.103+ |
| | (0.493) | (0.051) | (0.047) | (0.053) | (0.052) |
| Distance to nearest govt health facility (km)† | 4.033 | 0.149 | -0.144 | -0.172 | -0.341 |
| | (1.283) | (0.202) | (0.263) | (0.241) | (0.318) |
| Any members sick? | 0.646 | 0.025 | 0.004 | 0.061 | 0.018 |
| | (0.478) | (0.036) | (0.042) | (0.048) | (0.036) |
| Number of days ill? | 2.486 | 0.044 | -0.148 | -0.04 | 0.033 |
| | (2.157) | (0.190) | (0.214) | (0.260) | (0.125) |
| Number of days school/work missed due to illness† | 2.176 | 0.08 | -0.086 | 0.023 | 0.031 |
| | (1.987) | (0.166) | (0.180) | (0.239) | (0.098) |
| Waiting time before being attended (min.)† | 4.763 | -0.128 | -0.24 | -0.278+ | 0.014 |
| | (0.987) | (0.107) | (0.152) | (0.144) | (0.112) |
| Has visited traditional health practitioner? † | 0.283 | -0.051 | 0.013 | 0.044 | -0.001 |
| | (0.450) | (0.049) | (0.039) | (0.036) | (0.037) |
| Patient was examined by in-charge/doctor | 0.432 | 0.093 | 0.087 | -0.088 | -0.056 |
| | (0.496) | (0.061) | (0.052) | (0.066) | (0.054) |
| Time of examination | 3.415 | 0.054 | -0.022 | 0.001 | -0.12 |
| | (0.758) | (0.095) | (0.122) | (0.099) | (0.092) |
| Paid anything | 0.2 | 0.033 | 0.004 | 0.025 | -0.018 |
| | (0.401) | (0.039) | (0.036) | (0.061) | (0.032) |
| Received meds in hospital | 0.677 | 0.021 | -0.038 | 0.053 | 0.024 |
| | (0.468) | (0.041) | (0.037) | (0.040) | (0.064) |
| Had to buy meds outside of hospital | 0.955 | 0.002 | -0.007 | 0.015 | -0.042* |
| | (0.207) | (0.026) | (0.046) | (0.035) | (0.017) |
| Satisfied with services at hospital | 0.642 | 0.047 | -0.078 | -0.069 | -0.055 |
| | (0.480) | (0.039) | (0.048) | (0.053) | (0.036) |
| Number of observations | 6,703 | 326 | 786 | 789 | 771 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; † indicates that outcome was included in index.

**Table A11: Impact of barazas on education (matched ANOVA)**

| | Mean | Sub-county baraza | Information | Deliberation | Jurisdictional tier |
|---|---|---|---|---|---|
| Number of children in UPE or USE† | 1.696 | 0.253 | -0.078 | 0.045 | 0.041 |
| | (1.842) | (0.159) | (0.143) | (0.136) | (0.143) |
| Distance to public school (km) † | 1.424 | -0.018 | -0.031 | -0.055 | 0.066 |
| | (0.707) | (0.114) | (0.118) | (0.09) | (0.081) |
| Has complete boundary fence?† | 0.416 | 0.106 | -0.085 | -0.096 | -0.054 |
| | (0.493) | (0.07) | (0.075) | (0.062) | (0.087) |
| Has electricity? | 0.352 | 0.195** | 0.004 | -0.045 | -0.091 |
| | (0.478) | (0.058) | (0.052) | (0.061) | (0.054) |
| Has water facility?† | 0.677 | 0.094 | -0.029 | 0.032 | -0.103* |
| | (0.468) | (0.067) | (0.079) | (0.082) | (0.039) |
| Has PTA? | 0.959 | 0.005 | -0.033 | -0.054 | 0.003 |
| | (0.198) | (0.01) | (0.023) | (0.049) | (0.016) |
| Has SMC?† | 0.934 | 0.050+ | -0.011 | -0.054 | 0.027 |
| | (0.248) | (0.029) | (0.027) | (0.052) | (0.037) |
| Informed about SMC?† | 0.877 | -0.028 | -0.029 | -0.067 | -0.028 |
| | (0.328) | (0.047) | (0.04) | (0.059) | (0.022) |
| Inspectors visited schools?† | 0.730 | 0.078 | -0.024 | -0.054 | 0.058 |
| | (0.444) | (0.065) | (0.055) | (0.076) | (0.067) |
| Number of observations | 6,703 | 285 | 582 | 625 | 612 |

**Note:** column 1 reports sample means (and standard deviations below); column 2 reports effect (and standard errors below) of the sub-county-level baraza intervention; column 3 reports the effect (and standard errors below) of the information component of the baraza intervention; column 4 reports the effect (and standard errors below) of the deliberation component of the baraza intervention; column 5 reports the effect (and standard errors below) of the administrative placement of the baraza intervention; **, * and + denote significance at the 1%, 5% and 10% levels, respectively; † indicates that outcome was included in index.

## References


Alhassan, RK, Nketiah-Amponsah, E, Spieker, N, Arhinful, DK and De Wit, TFR, 2016. Assessing the Impact of Community Engagement Interventions on Health Worker Motivation and Experiences with Clients in Primary Health Facilities in Ghana: A Randomized Cluster Trial. *PloS One*, 11(7), e0158541.

Ananthpur, K, Malik, K and Rao, V, 2014. *The Anatomy of Failure: An Ethnography of a Randomized Trial to Deepen Democracy in Rural India.* The World Bank, Development Research Group, Poverty and Inequality Team.

Anderson, ML, 2008. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*, 103(484), pp.1481–1495.

Arkedis, J, Creighton, J, Dixit, A, Fung, A, Kosack and S, Levy, D, 2019. *Can Transparency and Accountability Programs Improve Health? Experimental Evidence from Indonesia and Tanzania.*

Banerjee, A, Deaton, A and Duflo, E, 2004. Wealth, health, and health services in rural Rajasthan. *American Economic Review*, 94(2), pp.326–330.

Banerjee, AV, Duflo, E and Glennerster, R, 2008. Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System. *Journal of the European Economic Association*, 6(2–3), pp.487–500.

Barrett, CB and Carter, MR, 2010. The power and pitfalls of experiments in development economics: Some non-random reflections. *Applied Economic Perspectives and Policy*, 32(4), pp.515–548.

Beath, A, Fotini, C and Ruben, E, 2013. *Do elected councils improve governance? Experimental evidence on local institutions in Afghanistan.* The World Bank.

Benin, S, Nkonya, E, Okecho, G, Pender, J, Nahdy, S and Mugarura, S, 2007. *Assessing the impact of the National Agricultural Advisory Services (NAADS) in the Uganda rural livelihoods.* International Food Policy Research Institute.

Berman, E, Callen, M, Condra, LN, Downey, M, Ghanik, T and Isaqzadeh, M, 2017. *Community Monitors vs. Leakage: Experimental Evidence from Afghanistan.*

Bernardi, RA, 2006. Associations between Hofstede's cultural constructs and social desirability response bias. *Journal of Business Ethics*, 65(1), pp.43–53.

Bernheim, BD and Whinston, M, 1986. Common Agency. *Econometrica*, 54(4), pp.923–42.

Bertsimas, D, Johnson, M and Kallus, N, 2015. The power of optimization over randomization in designing experiments involving small samples. *Operations Research*, 63(4), pp.868–876.

Beuermann, DW and Amelina, M, 2014. *Does Participatory Budgeting Improve Decentralized Public Service Delivery?* Inter-American Development Bank.

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

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

Blundell, R and Dias, MC, 2009. Alternative approaches to evaluation in empirical microeconomics*. Journal of Human Resources*, 44(3), pp.565–640.

Bradley, J and Igras, S, 2005. Improving the quality of child health services: participatory action by providers. *International Journal for Quality in Health Care*, 17, pp.391–399.

Burki, SJ, Perry, G, Dillinger, W, Griffin, C, Gutman, J, Rojas, F, Webb, S and Winkler, D, 1999. *Beyond the center: Decentralizing the state.* The World Bank.

Capuno, JJ and Garcia, MM, 2010. Can Information about Local Government Performance Induce Civic Participation? Evidence from the Philippines. *The Journal of Development Studies*, 46(4), pp.624–643.

Chung, J and Monroe, GS, 2003. Exploring social desirability bias. *Journal of Business Ethics*, 44(4), pp.291–302.

De Souza Leão, L and Eyal, G, 2019. The rise of randomized controlled trials (RCTs) in international development in historical perspective. *Theory and Society*, 48(3), pp.383–418.

Diaz-Cayeros, A, Magaloni, B and Ruiz-Euler, A, 2014. Traditional Governance, Citizen Engagement, and Local Public Goods: Evidence from Mexico. *World Development*, 53, pp.80–93.

Donato, K and Mosqueira, AG, 2016. Power to the people? A replication study of a community-based monitoring programme in Uganda. *3ie Replication Papers* 11, pp.92–108.

Dunning, T, Grossman, G, Humphreys, M, Hyde, S, McIntosh, C and Nellis, G eds., 2019. *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I (Cambridge Studies in Comparative Politics)*. Cambridge: Cambridge University Press.

Francis, P and James, R, 2003. Balancing rural poverty reduction and citizen participation: The contradictions of Uganda's decentralization program. *World Development*, 31(2), pp.325–337.

Fujiwara, T and Wantchekon, L, 2013. Can informed public deliberation overcome clientelism? Experimental evidence from Benin. *American Economic Journal: Applied Economics*, 5(4), pp.241–255.

Gerber, A and Green, D, 2012. *Field Experiments: Design, Analysis, and Interpretation.* WW Norton.

Gilens, M, 2001. Political Ignorance and Collective Policy Preferences. *American Political Science Review*, 95(2), pp.379–396.

Giné, X, Khalid, S and Mansuri, G, 2018. *The Impact of Social Mobilization on Health Service Delivery and Health Outcomes – Evidence from Rural Pakistan.* World Bank Group, Development Research Group, Finance and Private Sector Development Team & Poverty and Equity Global Practice Group.

Goeree, JK and Yariv, L, 2011. An Experimental Study of Collective Deliberation. *Econometrica*, 79(3), pp.893–921.

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

Goncalves, S, 2013. The Effects of Participatory Budgeting on Municipal Expenditures and Infant Mortality in Brazil. *World Development*, 53, pp.94–110.

Grossman, G, Rodden, J, Tausanovich, Z and Han, A, 2017. *Can Text Messages Improve Local Governance? An Impact Evaluation of the U-Bridge Programme in Uganda.* USAID.

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

Gullo, S, Galavotti, C, Kuhlmann, AS, Msiska, T, Hastings, P and Marti, CN, 2017. Effects of a social accountability approach, CARE's Community Score Card, on reproductive health-related outcomes in Malawi: A cluster-randomized controlled evaluation. *PloS One*, 12.

Heckman, J and Navarro, S, 2004. Using Matching, Instrumental Variables and Control Functions to Estimate Economic Choice Models. *Review of Economics and Statistics*, 86(l), pp.30–57.

Hebert, JR, Clemow, L, Pbert, L, Ockene, IS and Ockene, JK, 1995. Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *International Journal of Epidemiology*, 24(2), pp.389–398.

Humphreys, M, De la Sierra, RS and Van der Windt, P, 2013. Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration. *Political Analysis*, 21(1), pp.1–20.

Humphreys, M, De la Sierra, RS and Van der Windt, P, 2014. *Social and economic impacts of Tuungane: Final report on the effects of a community driven reconstruction programme in the Democratic Republic of Congo.* London: International Initiative for Impact Evaluation (3ie).

Humphreys, M and Weinstein, JM, 2012. *Policing Politicians: Citizen Empowerment and Political Accountability in Uganda Preliminary Analysis.* Columbia Universities. (Unpublished manuscript).

Iacus, SM, King, G and Porro, G, 2012. Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20(1), pp.1–24.

Kasy, M, 2016. Why experimenters might not always want to randomize, and what they could do instead. *Political Analysis*, 24(3), pp.324–338.

Kosec, K and Wantchekon, L, 2020. Can information improve rural governance and service delivery? *World Development*, 125, 104376.

McKenzie, D, 2012. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2), pp.210–221.

Molina, E, 2014. *Can Bottom-Up Institutional Reform Improve Service Delivery?* Inter-American Development Bank.

Muralidharan, K, Romero and M, Wüthrich, K. 2019. *Factorial designs, model selection, and (incorrect) inference in randomized experiments.* Tech. rep., National Bureau of Economic Research.

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

OPM, 2013. *Implementation of the Baraza Initiative.* Progress Report. Tech. rep., Office of the Prime Minister (OPM).

Pandey, P, Goyal, S and Sundararaman, V, 2009. Community participation in public schools: impact of information campaigns in three Indian states. *Education Economics*, 17(3), pp.355–375.

Ram, K, 2013. Git can facilitate greater reproducibility and increased transparency in science. *Source Code for Biology and Medicine*, 8(1), p.7.

Raffler, P, Posner, DN and Parkerson, D, 2018. *The weakness of bottom-up accountability: Experimental evidence from the Ugandan health sector.* (Unpublished manuscript).

Reinikka, R and Svensson, J, 2004. Local capture: evidence from a central government transfer program in Uganda. *The Quarterly Journal of Economics*, 119(2), pp.679–705.

Steiner, S, 2007. Decentralisation and poverty: conceptual framework and application to Uganda. Public Administration and Development. *The International Journal of Management Research and Practice*, 27(2), pp.175–185.

Timmons, JF and Garfias, F, 2015. Revealed Corruption, Taxation, and Fiscal Accountability: Evidence from Brazil. *World Development*, 70, pp.13–27.

Touchton, M and Wampler, B, 2014. Improving social well-being through new democratic institutions. *Comparative Political Studies*, 47(10), pp.1442–1469.

Uganda Bureau of Statistics, 2017. *The National Population and Housing Census 2014 – National Analytical Report.* Kampala, Uganda.

Van Campenhout, B, Bizimungu, E, Smart, J and Kabunga, N, 2018. Impact pathways of a participatory local governance initiative in Uganda: a qualitative exploration. *Development in Practice*, 28(8), pp.1046–1056.

Van Campenhout, B, Spielman, DJ and Lecoutere, E, 2020. Information and Communication Technologies to Provide Agricultural Advice to Smallholder Farmers: Experimental Evidence from Uganda. *American Journal of Agricultural Economics*.

Waddington, H, Sonnenfeld, A, Finetti, J, Gaarder, M, John, D and Stevenson, J, 2019. *Does incorporating participation and accountability improve development outcomes? Meta-analysis and framework synthesis. 3ie Systematic Review,* 43. London: International Initiative for Impact Evaluation (3ie).

Zerbe, WJ and Paulhus, DL, 1987. Socially Desirable Responding in Organizational Behavior: A Reconception. *Academy of Management Journal*, 12(2), pp.250–264.

## Other publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://3ieimpact.org/evidence-hub/publications/impact-evaluations

*Impacts of supportive feedback and nonmonetary incentives on child immunisation in Ethiopia,* 3ie Impact Evaluation Report 134. Demilew, A, Girma, M, McElwee, E, Datta, S, Barofsky, J and Disasa, T, 2021.

*Impacts of electronic case management systems on court congestion in the Philippines,* 3ie Impact Evaluation Report 133. Orbeta, AC, Jr, Paqueo, VB and Siddiqi, B, 2021.

*Impacts of judicial reform in small claims procedures on court congestion in the Philippines,* 3ie Impact Evaluation Report 132. Orbeta, AC, Jr, Paqueo, VB and Siddiqi, B, 2021.

*Impacts of judicial reform in criminal case procedures on court congestion in the Philippines,* 3ie Impact Evaluation Report 131. Orbeta, AC, Jr, Paqueo, VB and Siddiqi, B, 2021.

*Impacts of the Stimulate, Appreciate, Learn and Transfer community engagement approach to increase immunization coverage in Assam, India*, 3ie Impact Evaluation Report 130. Pramanik, S, Ghosh, A, Goswami, A, Das, T, Albert, S, Forth, P and Nanda, R, 2020.

*Impacts of a novel mHealth platform to track maternal and child health in Udaipur, India,* 3ie Impact Evaluation Report 129. Nagar, R, Ambiya, MS, Singh, P, Abdullah, H, Banshiwal, V, Stone, L, Manjanatha, D, Venkat, P, Purawat, D, Supatkar, V, Singh, A, Dalal, S and Shahnawaz, M, 2020.

*Impact evaluation of the National Rural Livelihoods Project,* 3ie Impact Evaluation Report 128. Kochar, A, Barooah, B, Jain, C, Singh, G, Closepet, N, Narayanan, R, Sarkar, R and Shah, R, 2020.

*Impacts of engaging communities through traditional and religious leaders on vaccination coverage in Cross River State, Nigeria,* 3ie Impact Evaluation Report 127. Oyo-Ita, A, Bosch-Capblanch, X, Ross, A, Hanlon, P, Oku, A, Esu, E, Ameh, S, Oduwole, B, Arikpo, D and Meremikwu, M, 2020.

*Evaluating the impact of interventions to improve full immunisation rates in Haryana, India,* 3ie Impact Evaluation Report 126. Banerjee, A, Chandrasekhar, A, Duflo, E, Dalpath, S, Floretta, J, Jackson, M, Kannan, H, Schrimpf, A and Shrestha, M, 2020.

*Impacts of community-led video education to increase vaccination coverage in Uttar Pradesh, India,* 3ie Impact Evaluation Report 125. Gurley, N, Shearer, J, Srivastava, Y, Mahapatra, S and Desmond, M, 2020.

*Impact of creative capacity building of local innovators and communities on income, welfare and attitudes in Uganda,* 3ie Impact Evaluation Report 124. Nkonya, E, Bashaasha, B, Kato, E, Bagamba, F and Danet, M, 2020.

*Impact evaluation of the integrated soil fertility management dissemination programme in Burkina Faso*, 3ie Impact Evaluation Report 123. A, Frölich, M, Koussoubé, E, Maïga, E and Varejkova, T, 2020.

*The effect of demonstration plots and the warehouse receipt system on integrated soil fertility management adoption, yield and income of smallholder farmers: a study from Malawi's Anchor Farms*, 3ie Impact Evaluation Report 122. Michelson, H, Barrett, C, Palm, C, Maertens, A, Mhango, W and Chirwa, E, 2020.

*Impacts of linking savings group to formal financial service providers and strengthening their internal group insurance mechanism in Zambia*, 3ie Impact Evaluation Report 121. Frölich, M and Nguyen, PL, 2020.

*Promoting latrine use in rural Karnataka using the risks, attitudes, norms, abilities and self-regulation (RANAS) approach*, 3ie Impact Evaluation Report 120. Friedrich, M, Balasundaram, T, Muralidharan, A, Raman, VR and Mosler, H-J, 2020.

*Impacts of low-cost interventions to improve latrine use and safe disposal of child faeces in rural Odisha, India*, 3ie Impact Evaluation Report 119. Caruso, BA, Sclar, GD, Routray, P, Nagel C, Majorin, F, Sola, S, Koehne, W, DeShay, R, Udaipuria, S, Williams, R and Clasen, T, 2020.

*Improving households' attitudes and behaviours to increase toilet use (HABIT) in Bihar, India*, 3ie Impact Evaluation Report 118. Viswanathan, S, Saith, R, Chakraborty, A, Purty, N, Malhotra, N, Singh, P, Mitra, P, Padmanabhan, V, Datta, S, Harris, J, Gidwani, S, Williams, R, Florence, E and Daniel, S, 2020.

*Rebuilding the social compact: urban service delivery and property taxes in Pakistan*, 3ie Impact Evaluation Report 117. Khwaja, AI, Haq, O, Khan, AQ, Olken, B and Shaukat, M, 2020.

*Rural institutional innovation: can village courts in Bangladesh accelerate access to justice and improve socio-economic outcomes?* 3ie Impact Evaluation Report 116. Mattsson, M and Mobarak, AM, 2020.

*Using big data to evaluate the impacts of transportation infrastructure investment: the case of subway systems in Beijing*, 3ie Impact Evaluation Report 115. Li, S and Liu, Y, 2020.

*Community toilet use in Indian slums: willingness-to-pay and the role of informational and supply side constraints*, 3ie Impact Evaluation Report 113. Armand, A, Augsburg, B, Bancalari A and Trivedi B, 2020.

*Impacts, maintenance and sustainability of irrigation in Rwanda*, 3ie Impact Evaluation Report 112. Byiringo, E, Jones M, Kondylis F, Loeser J, Magruder, J and Ndahimana, C, 2020.

*Continuous Emissions Monitoring Systems (CEMS) in India*, 3ie Impact Evaluation Report 111. Greenstone, M, Pande, R, Ryan, N and Sudarshan A, 2020.

*Evaluating the impacts of the Dar es Salaam Bus Rapid Transit System*, 3ie Impact Evaluation Report 110. Morten, M, Bryan, G, Siddiqi, B, Balboni, C, 2020.

*Access to safe drinking water: experimental evidence from new water sources in Bangladesh*, 3ie Impact Evaluation Report 109. Cocciolo, S, Ghisolfi, S, Habib, A, Rashid, SMA and Tompsett, A, 2020.

*Impact of alternate wetting and drying on farm incomes and water savings in Bangladesh*, 3ie Impact Evaluation Report 108. Chakravorty, U, Dar, MH, Emerick, K, 2020.

*The effects of vouchers for essential household items on child health, mental health, resilience and social cohesion among internally displaced persons in the Democratic Republic of Congo*, 3ie Impact Evaluation Report 107. Quattrochi, J, Bisimwa, G, Thompson, T, van der Windt, P and Voors, M, 2020.

*Measuring impacts of conservation interventions on human well-being and the environment in Northern Cambodia*, 3ie Impact Evaluation Report 106. Clements, T, Neang, M, Milner-Gulland, EJ and Travers, H, 2020.

*The 5 Star Toilet Campaign: improving toilet use in rural Gujarat,* 3ie Impact Evaluation Report 105. Chauhan, K, Schmidt, WP, Aunger, R, Gopalan, B, Saxena, D, Yashobant, S, Patwardhan, V, Bhavsar, P, Mavalankar, D and Curtis, V, 2020.

*How education about maternal health risk can change the gender gap in the demand for family planning in Zambia*, 3ie Impact Evaluation Report 104. Ashraf, N, Field, E, Voena, A and Ziparo, R, 2019.

In search of the holy grail: can unconditional cash transfers graduate households out of poverty in Zambia?, Impact Evaluation Report 103. Handa, S, Tembo, G, Natali, L, Angeles, G and Spektor, G, 2019.

*Increasing HIV self-testing and linkage to care for partners of women in antenatal care in Uganda,* Impact Evaluation Report 102. Wanyenze, R, Buregyeya, E, Matovu, J, Kisa, R, Kagaayi, J, Vrana-Diaz, C, Malek, A, Musoke, W, Chemusto, H, Mukama, S and Korte, J, 2019.

*Improving the quality of care for children with acute malnutrition in Uganda*, 3ie Impact Evaluation Report 101. Marzia, L, Wanzira, H, Lochoro, P and Putoto, G, 2019.

*Impacts of increasing community resilience through humanitarian aid in Pakistan*, 3ie Impact Evaluation Report 100. Avdeenko, A and Frölich, M, 2019.

*Impacts of community monitoring of socio-environmental liabilities in the Ecuadorian and Peruvian Amazon,* 3ie Impact Evaluation Report 99. Pellegrini, L, 2019.

*Increasing HIV testing demand among Kenyan truck drivers and female sex workers*, 3ie Impact Evaluation Report 98. Kelvin, E, George, G, Mwai, E, Kinyanjui, S, Inoti, S, Chetty, T, Strauss, M, Romo, M, Oruko, F, Odhiambo J, Nyaga, E, Mantell, J and Govender, K, 2019.

*Impacts of community stakeholder engagement interventions in Ugandan oil extractives,* 3ie Impact Evaluation Report 97. Parker, R, Coleman, E, Manyindo, J, Schultz, B and Mukuru, E, 2019.

*The impacts of formal registration of businesses in Malawi,* 3ie Impact Evaluation Report 96. Campos, F, Goldstein, M and McKenzie, D, 2019.

*Unpacking the determinants of entrepreneurship development and economic empowerment for women in Kenya*, 3ie Impact Evaluation Report 95. McKenzie, D, Puerto, S and Odhiambo, F, 2019.

*Impacts of key provisions in Ghana's Petroleum Revenue Management Act,* 3ie Impact Evaluation Report 94. Edjekumhene, I, Voors, M, Lujala, P, Brunnschweiler, C, Owusu, CK and Nyamekye, A, 2019.

*Using information to break the political resource curse in natural gas management in Mozambique,* 3ie Impact Evaluation Report 93. Armand, A, Costa, AI, Coutts, A, Vicente, P and Vilela, I, 2019.

*Harnessing transparency initiatives to improve India's environmental clearance process for the mineral mining sector,* 3ie Impact Evaluation Report 92. Pande, R and Sudarshan, A, 2019.

*Impacts of removing user fees for maternal health services on universal health coverage in Kenya,* 3ie Impact Evaluation Report 91. Abuya, T, Dennis, M, Matanda, D, Obare, F and Bellows, B, 2018.

*Impact of voice reminders to reinforce harvest aggregation services training for farmers in Mali*, 3ie Impact Evaluation Report 90. Osei, RD, Dzanku, FM, Osei-Akoto, I, Asante, F, Hodey, LS, Adu, PN, Adu-Ababio, K and Coulibaly, M, 2018.

*Impacts of Breakthrough's school-based gender attitude change programme in Haryana, India,* 3ie Impact Evaluation Report 89. Jayachandran, S, Jain, T and Dhar, D, 2018.

*Hotspot interventions at scale: the effects of policing and city services on crime in Bogotá, Colombia,* 3ie Impact Evaluation Report 88. Blattman, C, Green, D, Ortega, D and Tobón, S, 2018.

*Impact evaluation of the Philippine Special Program for Employment of Students,* 3ie Impact Evaluation Report 87. Beam, E, Linden, L, Quimbo, S and Richmond, H, 2018.

*Community-based distribution of oral HIV self-testing kits: experimental evidence from Zambia,* 3ie Impact Evaluation Report 86. Hensen, B, Ayles, H, Mulubwa, C, Floyd, S, Schaap, A, Chiti, B, Phiri, M, Mwenge, L, Simwinga, M, Fidler S, Hayes, R, Bond, V and Mwinga, A, 2018.

*Evaluating the economic impacts of rural banking: experimental evidence from southern India,* 3ie Impact Evaluation Report 85. Field, E and Pande, R, 2018.

*Direct provision versus facility collection of HIV tests: impacts of self-testing among female sex workers in Uganda.* 3ie Impact Evaluation Report 84. Ortblad, K, Musoke, DK, Ngabirano, T, Oldenburg, C and Bärnighausen, T, 2018.

*Increasing female sex worker HIV testing: effects of peer educators and HIV self-tests in Zambia*, 3ie Impact Evaluation Report 83. Chanda, MM, Ortblad, KF, Mwale, M, Chongo, S, Kanchele, C, Kamungoma, N, Fullem, A, Bärnighausen, T and Oldenburg, CE, 2018.

*Community delivery of antiretroviral drugs: a non-inferiority matched-pair pragmatic cluster-randomized trial in Dar es Salaam, Tanzania*, 3ie Impact Evaluation Report 82. Francis, JM, Geldsetzer, P, Asmus, G, Ulenga, N, Ambikapathi, R, Sando, D, Fawzi, W and Bärnighausen, T, 2018.

*Nourishing the future: targeting infants and their caregivers to reduce undernutrition in rural China,* 3ie Impact Evaluation Report 81. Cai, J, Luo, R, Li, H, Lien, J, Medina, A, Zhou, H and Zhang, L, 2018.

*Impacts of the World Food Programme's interventions to treat malnutrition in Niger.* 3ie Impact Evaluation Report 80. Brück, T, Ferguson, NTN, Ouédraogo, J and Ziegelhöfer, Z, 2018.

*Impact evaluation of the World Food Programme's moderate acute malnutrition treatment and prevention programmes in Sudan.* 3ie Impact Evaluation Report 79. Guevarra, E, Mandalazi, E, Balegamire, S, Albrektsen, K, Sadler, K, Abdelsalam, K, Urrea, G and Alawad, S, 2018.

*Impact evaluation of WFP's programs targeting moderate acute malnutrition in humanitarian situations in Chad.* 3ie Impact Evaluation Report 78. Saboya, M, Rudiger, J, Frize, J, Ruegenberg, D, Rodríguez Seco, A and McMillon, C, 2018.

*Improving midday meal delivery and encouraging micronutrient fortification among children in India,* 3ie Impact Evaluation Report 77. Shastry, GK, Berry, J, Mukherjee, P, Mehta, S and Ruebeck, H, 2018.

*Evaluation of infant development centres: an early years intervention in Colombia,* 3ie Impact Evaluation Report 76. Andrew, A, Attanasio, O, Bernal, R, Cordona, L, Krutikova, S, Heredia, DM, Medina, C, Peña, X, Rubio-Codina, M and Vera-Hernandez, M, 2018.

*Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone.* 3ie Impact Evaluation Report 75. Cilliers, J, Dube, O and Siddiqi, B, 2018.

*Impact evaluation of the Menabe and Melaky development programme in Madagascar,* 3ie Impact Evaluation Report 74. Ring, H, Morey, M, Kavanagh, E, Kamto, K, McCarthy, N, Brubaker, J and Rakotondrafara, C, 2018.

*Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya,* 3ie Impact Evaluation Report 73. Bonilla, J, McCarthy, N, Mugatha, S, Rai, N, Coombes, A and Brubaker, J, 2018.

*Impact and adoption of risk-reducing drought-tolerant rice in India*, 3ie Impact Evaluation Report 72. Yamano, T, Dar, MH, Panda, A, Gupta, I, Malabayabas, ML and Kelly, E, 2018.

*Poverty and empowerment impacts of the Bihar Rural Livelihoods Project in India,* 3ie Impact Evaluation Report 71. Hoffmann, V, Rao, V, Datta, U, Sanyal, P, Surendra, V and Majumdar, S 2018.

*How should Tanzania use its natural gas? Citizens' views from a nationwide Deliberative Poll,* 3ie Impact Evaluation Report 70. Birdsall, N, Fishkin, J, Haqqi, F, Kinyondo, A, Moyo, M, Richmond, J and Sandefur, J, 2018.

*Impact evaluation of the conditional cash transfer program for secondary school attendance in Macedonia,* 3ie Impact Evaluation Report 69. Armand, A and Carneiro, P, 2018.

*Age at marriage, women's education, and mother and child outcomes in Bangladesh,* 3ie Impact Evaluation Report 68. Field, E, Glennerster, R, Nazneen, S, Pimkina, S, Sen, I and Buchmann, N, 2018.

*Evaluating agricultural information dissemination in western Kenya*, 3ie Impact Evaluation Report 67. Fabregas, R, Kremer, M, Robinson, J and Schilbach, F, 2017.

*General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia,* 3ie Impact Evaluation Report 66. Filipski, M, Taylor, JE, Abegaz, GA, Ferede, T, Taffesse, AS and Diao, X, 2017.

*Impact of the Uddeepan programme on child health and nutrition in India,* 3ie Impact Evaluation Report 65. Kochar, A, Sharma, A and Sharma, A, 2017.

*Evaluating oral HIV self-testing to increase HIV testing uptake among truck drivers in Kenya,* 3ie Impact Evaluation Report 64. Kelvin, EA, Mwai, E, Romo, ML, George, G, Govender, K, Mantell, JE, Strauss, M, Nyaga, EN and Odhiambo, JO, 2017.

*Integration of EPI and paediatric HIV services for improved ART initiation in Zimbabwe*, 3ie Impact Evaluation Report 63. Prescott, M, Boeke, C, Gotora, T, Mafaune, HW, Motsi, W, Graves, J, Mangwiro, A and McCarthy, E, 2017.

*Increasing male partner HIV testing using self-test kits in Kenya,* 3ie Impact Evaluation Report 62. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

*Evaluating the impact of community health worker integration into prevention of mother-to-child transmission of HIV services in Tanzania,* 3ie Impact Evaluation Report 61. Nance, N, McCoy, S, Ngilangwa, D, Masanja, J, Njau, P and Noronha, R, 2017.

*Using HIV self-testing to promote male partner and couples testing in Kenya,* 3ie Impact Evaluation Report 60. Thirumurthy, H, Omanga, E, Obonyo, B, Masters, S and Agot, K, 2017.

*Increasing male partner HIV self-testing at antenatal care clinics in Kenya,* 3ie Impact Evaluation Report 59. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

*Impact of free availability of public childcare on labour supply and child development in Brazil*, 3ie Impact Evaluation Report 58. Attanasio, O, Paes de Barros, R, Carneiro, P, Evans, D, Lima, L, Olinto, P and Schady, N, 2017.

*Estimating the effects of a low-cost early stimulation and parenting education programme in Mexico,* 3ie Impact Evaluation Report 57. Cardenas, S, Evans, D and Holland, P, 2017.

*The Better Obstetrics in Rural Nigeria study: an impact evaluation of the Nigerian Midwives Service Scheme,* 3ie Impact Evaluation Report 56. Okeke, E, Glick, P, Abubakar, IS, Chari, AV, Pitchforth, E, Exley, J, Bashir, U, Setodji, C, Gu, K and Onwujekwe, O, 2017.

*The Productive Safety Net Programme in Ethiopia: impacts on children's schooling, labour and nutritional status*, 3ie Impact Evaluation Report 55. Berhane, G, Hoddinott, J, Kumar, N and Margolies, A, 2016.

*The impact of youth skills training on the financial behaviour, employability and educational choice in Morocco,* 3ie Impact Evaluation Report 54. Bausch, J, Dyer, P, Gardiner, D, Kluve, J and Mizrokhi, E, 2016.

*Using advertisements to create demand for voluntary medical male circumcision in South Africa,* 3ie Impact Evaluation Report 53. Frade, S, Friedman, W, Rech, D and Wilson, N, 2016.

*The use of peer referral incentives to increase demand for voluntary medical male circumcision in Zambia,* 3ie Impact Evaluation Report 52. Zanolini, A, Bolton, C, Lyabola, LL, Phiri, G, Samona, A, Kaonga, A and Harsha Thirumurthy, H, 2016.

*Using smartphone raffles to increase demand for voluntary medical male circumcision in Tanzania,* 3ie Impact Evaluation Report 51. Mahler, H and Bazant, E, 2016.

*Voluntary medical male circumcision uptake through soccer in Zimbabwe,* 3ie Impact Evaluation Report 50. DeCelles, J, Kaufman, Z, Bhauti, K, Hershow, R, Weiss, H, Chaibva, C, Moyo, N, Braunschweig, E, Mantula, F, Hatzold, K and Ross, D, 2016.

*Measuring the impact of SMS-based interventions on uptake of voluntary medical male circumcision in Zambia,* 3ie Impact Evaluation Report 49. Leiby, K, Connor, A, Tsague, L, Sapele, C, Koanga, A, Kakaire, J and Wang, P, 2016.

*Assessing the impact of delivering messages through intimate partners to create demand for voluntary medical male circumcision in Uganda,* 3ie Impact Evaluation Report 48. Semeere, AS, Bbaale, DS, Castelnuovo, B, Kiragga, A, Kigozi, J, Muganzi, A, Kambugu, A and Coutinho, AG, 2016.

*Optimising the use of economic interventions to increase demand for voluntary medical male circumcision in Kenya,* 3ie Impact Evaluation Report 47. Thirumurthy, H, Omanga, E, Rao, SO, Murray, K, Masters, S and Agot, K, 2016.

*The impact of earned and windfall cash transfers on livelihoods and conservation in Sierra Leone,* 3ie Impact Evaluation Report 46. Bulte, E, Conteh, B, Kontoleon, A, List, J, Mokuwa, E, Richards, P, Turley, T and Voors, M, 2016.

*Property tax experiment in Pakistan: Incentivising tax collection and improving performance,* 3ie Impact Evaluation Report 45. Khan, A, Khwaja, A and Olken, B, 2016.

*Impact of mobile message reminders on tuberculosis treatment outcomes in Pakistan,* 3ie Impact Evaluation Report 44. Mohammed, S, Glennerster, R and Khan, A, 2016.

*Making networks work for policy: Evidence from agricultural technology adoption in Malawi,* 3ie Impact Evaluation Report 43. Beaman, L, BenYishay, A, Fatch, P, Magruder, J and Mobarak, AM, 2016.

*Estimating the impact and cost-effectiveness of expanding access to secondary education in Ghana,* 3ie Impact Evaluation Report 42. Dupas, P, Duflo, E and Kremer, M, 2016.

*Evaluating the effectiveness of computers as tutors in China, 3*ie Impact Evaluation Report 41. Mo, D, Bai, Y, Boswell, M and Rozelle, S, 2016.

*Micro entrepreneurship support programme in Chile,* 3ie Impact Evaluation Report 40. Martínez, CA, Puentes, EE and Ruiz-Tagle, JV, 2016.

*Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh,* 3ie Impact Evaluation Report 39. Barham, T, Kuhn, R, Menken, J and Razzaque, A, 2016.

*Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection and malnutrition in India,* 3ie Impact Evaluation Report 38. Clasen, T, Boisson, S, Routray, P, Torondel, B, Bell, M, Cumming, O, Ensink, J, Freeman, M and Jenkins, M, 2016.

*Evaluating the impact of vocational education vouchers on out-of-school youth in Kenya,* 3ie Impact Evaluation Report 37. Hicks, JH, Kremer, M, Mbiti, I and Miguel, E, 2016.

*Removing barriers to higher education in Chile: evaluation of peer effects and scholarships for test preparation,* 3ie Impact Evaluation Report 36. Banerjee, A, Duflo E and Gallego, F, 2016.

*Sustainability of impact: dimensions of decline and persistence in adopting a biofortified crop in Uganda,* 3ie Impact Evaluation Report 35. McNiven, S, Gilligan, DO and Hotz, C 2016.

*A triple win? The impact of Tanzania's Joint Forest Management programme on livelihoods, governance and forests*, 3ie Impact Evaluation Report 34. Persha, L and Meshack, C, 2016.

*The effect of conditional transfers on intimate partner violence: evidence from Northern Ecuador,* 3ie Impact Evaluation Report 33. Hidrobo, M, Peterman, A and Heise, L, 2016.

*The effect of transfers and preschool on children's cognitive development in Uganda,* 3ie Impact Evaluation Report 32. Gillian, DO and Roy, S, 2016.

*Can egovernance reduce capture of public programmes? Experimental evidence from India's employment guarantee,* 3ie Impact Evaluation Report 31. Banerjee, A, Duflo, E, Imbert, C, Mathew, S and Pande, R, 2015.

*Improving maternal and child health in India: evaluating demand and supply strategies,* 3ie Impact Evaluation Report 30. Mohanan, M, Miller, G, Forgia, GL, Shekhar, S and Singh, K, 2016.

*Smallholder access to weather securities in India: demand and impact on production decisions,* 3ie Impact Evaluation Report 28. Ceballos, F, Manuel, I, Robles, M and Butler, A, 2015.

*What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi,* 3ie Impact Evaluation Report 27. Baird, S, Chirwa, E, McIntosh, C and Özler, B, 2015.

*Validation of hearing screening procedures in Ecuadorian schools,* 3ie Impact Evaluation Report 26. Muñoz, K, White, K, Callow-Heusser, C and Ortiz, E, 2015.

*Assessing the impact of farmer field schools on fertilizer use in China,* 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G, 2015.

*The SASA! study: a cluster randomised trial to assess the impact of a violence and HIV prevention programme in Kampala,* Uganda, 3ie Impact Evaluation Report 24. Watts, C, Devries, K, Kiss, L, Abramsky, T, Kyegombe, N and Michau, L, 2014.

*Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts,* 3ie Impact Evaluation Report 23. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R, 2015.

*A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana,* 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M, 2015.

*Shelter from the storm: upgrading housing infrastructure in Latin American slums*, 3ie Impact Evaluation Report 21. Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R, 2015.

*Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme*, 3ie Impact Evaluation Report 20. Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P, 2015.

*A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China,* 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E, 2014.

*Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation,* 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K, 2015.

*Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua*, 3ie Impact Evaluation Report 17. Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME, 2014.

*The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India,* 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M, 2014.

*A youth wage subsidy experiment for South Africa,* 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V, 2014.

*Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone,* 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S, 2014.

*Scaling up male circumcision service provision: results from a randomised evaluation in Malawi,* 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R, 2014.

*Targeting the poor: evidence from a field experiment in Indonesia,* 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R, 2014.

*An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India,* 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B, 2014.

*Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India,* 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N, 2013.

*No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia*, 3ie Impact Evaluation Report 9. Ashraf, N, Bandiera, O and Jack, K, 2013.

*Paying for performance in China's battle against anaemia*, 3ie Impact Evaluation Report 8. Zhang, L, Rozelle, S and Shi, Y, 2013.

*Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo*, 3ie Impact Evaluation Report 7. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P, 2013.

*The impact of daycare on maternal labour supply and child development in Mexico*, 3ie Impact Evaluation Report 6. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E, 2014.

*Impact evaluation of the non-contributory social pension programme 70 y más in Mexico*, 3ie Impact Evaluation Report 5. Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M, 2014.

*Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India,* 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A, 2013.

*The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone*, 3ie Impact Evaluation Report 3. Casey, K, Glennerster, R and Miguel, E, 2013.

*A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana*, 3ie Impact Evaluation Report 2. Burwen, J and Levine, DI, 2012.

*The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique*, 3ie Impact Evaluation Report 1. Martinez, S, Naudeau, S and Pereira, V, 2012.

Ineffective monitoring and weak accountability mechanisms have greatly affected the realization of the benefits of decentralization in Uganda. To address this, the government initiated community advocacy forums or citizen barazas. The objective was to enhance public involvement in holding the government accountable for service delivery, especially, in relation to the resources spent. The authors of this report conducted a cluster randomised control trial to understand the impact of these barazas on public service delivery and to inform policymakers about the effectiveness of this initiative.

**Impact Evaluation Series**

International Initiative for Impact Evaluation
215-216, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

**www.3ieimpact.org**