Governance, Citizenship, and Accountability: A Pilot Study of People-Centered Development in the Ugandan Health Sector

Joshua L. Greenberg*

Medical School and Department of Economics, University of Michigan, Ann Arbor, MI, USA, ORCID iD: 0000-0002-2928-8111, Twitter: @1JoshGreenberg

December 22, 2022

Abstract

Widespread accountability gaps in the public health sectors of low- and middle-income countries account for many government failures to deliver available, affordable health interventions. In Uganda, local political leaders have the potential to mitigate accountability problems by monitoring service provision at government health centers — but few do so. Limited citizen-politician information flow and inadequate politician training may explain this observation. This pilot study aims to address such constraints to quality healthcare delivery with two governance interventions: (a) quarterly reporting meetings between citizens and local politicians to address health service delivery and (b) politician skills training on monitoring local government health centers. The study is one of a few to evaluate programs directly targeting political economy inefficiencies. Both interventions achieve substantial stakeholder participation and engagement, laying the groundwork for an expanded evaluation with greater statistical power.

Keywords:

JEL Codes:

* *Correspondence Address:* Medical School and Department of Economics, 238 Lorch Hall, 611 Tappan Avenue, University of Michigan, Ann Arbor, MI, USA 48109, Email: jlgr@umich.edu.

1 Introduction

The technologies needed to save millions of lives each year — mostly in low- and middle-income countries (LMICs) — have existed for decades. For example, 5.2 million children across the world continue to die annually (UN IGME 2020), mainly from causes such as pneumonia, diarrhea, and intrapartum-related complications, all of which can be prevented or treated with cost-effective solutions (Chopra et al. 2013; Liu et al. 2015). Why do governments fail to deliver available health interventions, despite their affordability?

1.1 Local Misallocation

Although such delivery failures may be partly driven by resource misallocations at a global level,[1] they also stem largely from local misallocation of resources. These local misallocations may themselves span several levels, from central government expenditures to the transportation and delivery of health supplies to the organization and performance of clinical services at health facilities. The third category of "last-mile" service provision failures plays a prominent role and is rooted in widespread accountability gaps in the public health facilities of poor countries, including health worker absenteeism (Chaudhury et al. 2006), low health worker performance (Leonard and Masatu 2010), diversion of supply and medication inventories, and informal payment in the provision of services (Lewis and Pettersson 2009; McPake et al. 1999). For example, in the Ugandan public health sector, the setting of this study, health worker absenteeism rates exceed 47 percent (World Bank 2016). Such shortfalls constitute fundamental management problems in the healthcare systems of LMICs, leading to significant wastage of available resources.[2]

Strategies for improving care at local health facilities often include training health workers, providing equipment, and introducing quality protocols. However, these primarily tech-

[1] For example, global resource misallocations may include between-country inequities and inefficiencies in public finance, economic growth, and investment, as well as market failures in the pharmaceutical industry.

[2] For more detailed discussion of current health sector performance in Uganda, please see Appendix E.1.

nical interventions may be limited in impact if the same institutional structures among healthcare actors remain. The greater need may be for fundamental reform in the service delivery incentives faced by health workers and political leaders (Pritchett and Woolcock 2004; Bold and Svensson 2013).

1.2 Top-Down Monitoring

One strategy to better align the incentives for healthcare delivery is increased top-down monitoring — namely, through audits. In Indonesia, for instance, Olken (2007) finds that increased frequency of audits by the central government reduces corruption in village road projects. As a penalty-based mechanism to improve healthcare provider incentives, audits may be carried out by higher-level authorities such as the district health management teams that oversee local health facilities in Uganda. However, despite their promise, top-down audits also come with downsides. First, since such measures constitute a central government function, it may not be feasible for the central government to monitor the bureaucrats appointed to perform audits; consequently these individuals themselves may be corrupt (Bardhan and Mookherjee 2006). In addition, top-down approaches may require substantial resources, requiring auditors to visit a large number of facilities in their jurisdictions on a regular basis, and may therefore be difficult to sustain. Even if such audits are effective, they beget another, deeper question: what creates the political incentives to establish an auditing program in the first place?

1.3 People-Centered Approaches

As the fundamental underpinning of political systems, the preferences of the populace shape the political incentives for effective policy. Hence, the solution to ineffective governance might boil down to more effective citizenship. By the same token, interventions that instead target political incentives themselves — through organic, people-centered approaches — could lead to far-reaching reforms by altering the very political system that is responsible

for quality public service provision. Indeed, in addressing the foundational elements of the political system, we may be able to produce a constellation of public service enhancements extending well beyond the establishment of a single intervention, such as auditing. In addition, by altering institutional structures, grassroots, system-driven approaches may provide a mechanism for sustaining service quality improvements in the long-run.

Given these considerations, strategies that pivot around citizens, who reside in proximity to health centers and who are capable of altering the incentives of political leaders, may promise significant impacts on healthcare delivery. Along these lines, the 2004 *World Development Report* (WDR) focuses entirely on the role of citizen participation in assuring accountability in service delivery, and the 2017 WDR centers on the essential role of governance in development (World Bank 2003; 2017). Despite this attention, approaches that give citizens a platform from which they can better influence government service delivery have not been widely prioritized in development policy, and relatively little rigorous evidence exists on suitable interventions.

1.4 Local Political Leaders in Uganda

In Uganda, local political leaders, who are elected democratically, have the potential to mitigate accountability problems by monitoring service provision at local government health centers. However, politicians tend to engage with the health centers on rare occasions — but have the ability to institute important changes when they do. How so? Most fundamentally, Uganda's *Local Governments Act* invests local politicians with responsibilities that include overseeing government employee performance and government service provision within their jurisdictions (Republic of Uganda 2000). In terms of enforcing improved service delivery based on these legal monitoring authorities, several factors make local politicians well-positioned to induce responses from health workers. First, local politicians can report negative performance to district authorities, including the District Health Officer, District Personnel Officer, Chief Administrative Officer, and District Service Commission, the body

responsible for hiring health center employees. If a health worker's performance is sufficiently lacking, politicians (through their jurisdiction's council) can request a transfer of the employee. Such transfers are extremely costly to health workers, with negative consequences for their reputations and networks. According to Raffler (2020), about 50 percent of transfer requests are granted. Second, many local political leaders reside in offices that are proximate to the health centers, giving them easy, consistent access to the facilities. Thus, if local politicians begin to better exercise their monitoring responsibilities, this proximity should make punishment threats against health workers more credible. Third, health workers may also respond to local politicians due to the threat of social sanctions (e.g., humiliation for poor performance) within the community.

Yet, despite all of their abilities to help improve the quality of service delivery, few local leaders engage in regular monitoring activity — exemplifying the gap between the existence and the enforcement of the law that persists in many LMICs (World Bank 2017). Moreover, the constituents of local politicians exert little pressure on them to take action. Why not, given that households and communities might experience significant benefits if their democratically elected leaders adopted more measures to improve public service delivery?

1.5 Theoretical Framework

Economic theory, in conjunction with qualitative information I have collected in the field, points to several possible explanations for the apparent institutional failures in the health sector and political system. Broadly speaking, these explanations cast such accountability deficiencies in health service delivery (vis-à-vis the role of political leaders) as either an incentives/information problem, a human capital problem, or both.

1.5.1 Incentives and Information

According to the incentives and information characterization, two main constraints may lead to suboptimal politician monitoring of health service quality: (1) unobservability of

politician performance by citizens and (2) incomplete information experienced by political leaders related to quality of care at local health centers. The first constraint is grounded in the principal-agent problem that characterizes the relationship between citizens and politicians. Due to the features of this problem (Ross 1973; Hölmstrom 1979), citizens (principals) may be unable to efficiently reward or sanction politicians (agents). Specifically, politician performance — with regard to the health service monitoring activities in which leaders engage — may be largely unobservable. As a result, citizens may be unable to effectively sanction politicians in elections, since they may be unsure of the actions taken by politicians and of whether politicians are to blame for poor health services. Thus, even in the context of democratic elections, citizen incomplete information may render a punishment threat to vote a politician out of office infeasible. The second-best solution for citizens would be to infer politician performance based on the quality of health service provision. However, this measure is likely to be a very noisy signal of politician performance; citizens may never be able to tease out the component of health service quality that can be attributed to politician behavior, again inhibiting them from efficiently determining rewards or sanctions for performance.

Beyond the principal-agent problem, the second constraint — incomplete information on the part of political leaders — arises from the fact that many politicians may seek healthcare elsewhere and, as a result, may be unaware of the accountability problems. Moreover, even if they seek services from the health centers, they may be treated differently and may remain unaware of the average experience of their constituents.

1.5.2 Human Capital

According to the human capital characterization, local politicians may have inadequate training for overseeing health service delivery. Specifically, politicians may be unaware of their rights and responsibilities as elected leaders, along with effective methods for monitoring health service delivery (e.g., gathering information on patient complaints, auditing staff at-

tendance, spot-checking inventory stocks). Beyond methods to collect information on quality of care, politicians also need to understand how to act on the information that they gather. In this regard, they may possess inadequate knowledge of the political and institutional channels through which to pursue improvements in the healthcare system. For instance, at a local level, what strategies can politicians use to work with health center staff to integrate improved practices into facility operations? At a more centralized level, if politicians identify service shortfalls that require resolutions by higher levels of government (e.g., increased budgetary allocations, provision of additional staff), what are the proper administrative procedures that local leaders should pursue to submit requests and to advocate on behalf of their constituencies?

Observations from the field confirm that local leaders appear to be largely unaware of the actions that they can take on these fronts. Such human capital constraints likely arise, at least in part, from information asymmetries between politicians and health workers, with the latter acting as gatekeepers of important health facility information and the former being uninformed of their rights and responsibilities as elected representatives. Indeed, Raffler (2020) observes this dynamic in Uganda while studying local government financial accountability.

The remainder of this paper is structured as follows. In Section 2, I describe the overall goals of this study as a pilot testing new approaches to address the health service provision gaps in Uganda. In Section 3, I contextualize the study within the body of existing literature on governance interventions. Following, Section 4 outlines the study's experiment design and describes in detail the study's interventions and data collection. Section 5 presents the main results of the study, including achievements from the implementation of activities, along with the participation and engagement outcomes of the interventions. Then, Section 6 presents an initial, condensed pre-analysis plan for a follow-on study with sufficient statistical power to enable a rigorous quantitative evaluation. Finally, Section 7 discusses the significance of the findings, in regard to both our understanding of the political economy of healthcare

delivery and the potential of the subsequent expanded study.

2 Study Objectives

Carried out in collaboration with the Office of the Prime Minister (OPM) and a non-profit organization, Progressive Health Partnership (PHP), in Uganda, this study consists of a feasibility assessment and pilot randomized controlled trial of the following two governance interventions: (a) increased citizen participation through quarterly feedback, reporting, and accountability meetings with local politicians to discuss health service quality and (b) skills training for politicians on monitoring local government health centers. Designed based on the theoretical framework described in Section 1.5, the OPM's lessons learned from its accountability programs, and PHP's experiences in its partner communities, the interventions have the potential both to produce significant impacts on health service delivery and to be implemented at greater scale through the public sector.

In preparation for a longer-term study with greater statistical power, the primary objective of this pilot has been twofold: to evaluate the feasibility of (1) the interventions and (2) the data collection activities. Through this study, I examine the nature and level of local responsiveness to the interventions by both citizens and political leaders, the feasibility of scaling up the programs, and the qualitative effects and mechanisms of the interventions. Equally important, I evaluate the feasibility of the data collection activities — a key objective given the potential political sensitivities of the survey questions.

For the OPM, this small-scale study has provided an opportunity to test new accountability strategies intended to build on its Baraza Program, a local citizen advocacy initiative that the OPM has been carrying out since 2009. Based on the results of this pilot, the OPM is committed to collaborating on a larger study. Altogether, the pilot serves to provide critical evidence for proof of concept, establishing the framework for a rigorous evaluation.

3 Literature Review

Only a few other randomized evaluations of similarly motivated interventions have been performed. In regard to community organization, Björkman and Svensson (2009) find significant impacts of a participatory intervention in the Ugandan health sector that mobilizes residents to establish health provider monitoring activities implemented and sustained by the community. In encouraging citizen themselves to adopt public sector monitoring practices, this approach centers on the role of the citizen-provider relationship. In contrast, the present study hinges on the relationship between citizens and politicians, the fundamental underpinning of any democratic system of governance. In this context, changes initiated in response to citizen participation should act through improved performance of existing monitoring responsibilities by politicians — not through the establishment of new, parallel monitoring institutions.

Studying corruption, Olken (2007) observes little effect of an "accountability meetings" intervention that again focuses on the citizen-provider rather than the citizen-politician relationship. Because public service providers do not face direct electoral punishment threats from citizens, the citizen-provider approach may be less effective. Pointing to the potential impacts of the meetings intervention in this study, Bidwell, Casey, and Glennerster (2020) find that political information — conveyed through public debates — affects voting behavior, leading to campaign expenditure responses by candidates and greater financial accountability of subsequently elected leaders.

In regard to politician training, Raffler (2020) evaluates the impact of a similar local political leader training and information program in Uganda, which leads to improved political oversight of bureaucrats involved in managing finances and local development projects. Ravanilla (2016) finds that a training workshop combined with performance-based rewards in the Philippines leads more qualified individuals to serve in public office. The present study aims to build on this previous work by not only examining intermediary outcomes such as

politician behavior but also welfare impacts on health. I also examine aspects of job training that have not yet been rigorously tested.

4 Methods

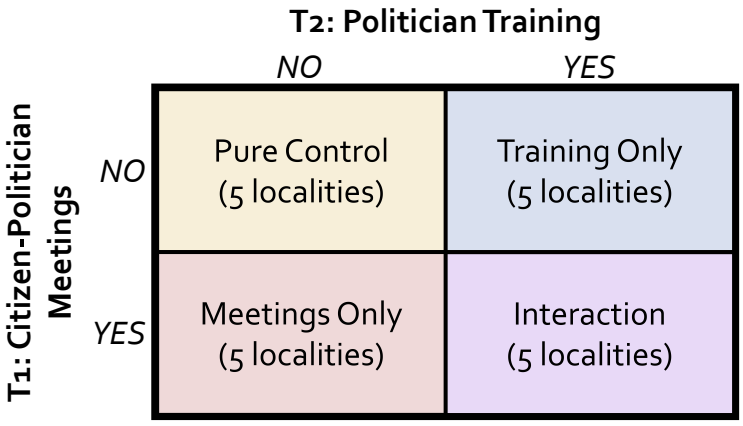
4.1 Experiment Design

Along with the OPM, I collaborated with PHP to implement the interventions in a stratified, cross-randomized, 2×2 design across twenty subcounties/town councils/divisions throughout the Ankole Region of Uganda. As equivalent administrative units across rural, periurban, and urban areas, respectively, each subcounty, town council, and division has a government health center referred to as a Health Center III and is headed by a democratically elected council chairperson. To account for the potential influence of the underlying political environment on intervention effects, I used data from the Electoral Commission of Uganda to stratify the experimental randomization by the margin of victory of council chairpersons in the recent 2016 election. The margin of victory may be a proxy for the level of political competitiveness and possibly for the quality of democratic functioning in each locality. Specifically, I divided the population of subcounties/town councils/divisions into five strata: the first four strata matched the quartiles of the margin of victory variable, while the fifth stratum included all of the localities in which the chairperson ran for office unopposed in the recent election.[3]

Within each of the five strata, I randomly selected four localities, which I then randomized across four groups: pure control, citizen-politician meetings only, politician training only, and interaction. These four arms, encompassing both single and joint implementation of the interventions, are designed to provide insight into key constraints on health service provision, based on the theoretical framework presented above. Although the scale of the pilot does not provide sufficient statistical power for rigorous quantitative analysis of the intervention effects, I implemented the stratified, randomized design to inform qualitative observations

[3]For more information on the study population of localities, please see Appendix A.

FIGURE 1: Experiment Randomization Matrix



Notes: This figure depicts the study's cross-randomized design. The twenty sample localities were evenly divided amongst the four study arms. The interaction arm included both the training and meetings interventions.

and to establish a replicable framework for a larger study.

4.2 Intervention Design and Mechanisms

4.2.1 Citizen-Politician Meetings Program

In the meetings treatment arm, PHP organized quarterly meetings in each locality between the Council Chairperson and citizens to discuss local healthcare quality, creating opportunities for regular feedback and reporting. Within each administrative unit, PHP carried out two meetings — one in the parish/ward that contains the Health Center III and another in a randomly chosen parish/ward.[4] The first location helped to ensure that the most likely users of the health facility participated in a meeting, while the second location served to increase intervention coverage and to make the meetings as accessible and as decentralized as possible. Although PHP originally planned to carry out the intervention over the course of three quarters, the program implementation was paused due to the Covid pandemic after eight of the ten meeting arm localities had been covered in the first round of activities.[5] This

[4]Parishes are lower-level administrative units within subcounties, while wards are lower-level administrative units within town councils and divisions.

[5]The organization and the research team are currently assessing the possible resumption of the study activities.

round took place between January and March 2020.

To foster widespread, representative participation, PHP implemented a systematic and multipronged mobilization procedure, which primarily consisted of making announcements throughout communities with a mobile public address system, hanging announcement posters, sending text message reminders to local leaders, and working with community health workers, village leaders, and other local influencers to advertise each meeting in advance to the population-at-large. To ensure effectively run sessions, trained facilitators mediated the meetings according to standardized guidelines. In brief, each meeting followed a standard agenda, consisting of the following: (1) introductions and establishment of general conduct expectations, (2) small-group discussions along with optional submission of anonymous written comments, (3) open discussion, including a formal response to community concerns from the LC3 Chairperson, (4) presentation of an action plan by the LC3 Chairperson and closing remarks. Given potential political sensitivities, the meeting implementation guidelines emphasized the importance of maintaining neutrality and remaining focused on constructive steps to improve healthcare delivery.

In the context of the theoretical framework outlined in Section 1.5, the citizen-politician meetings intervention aims to address the unobservability problem of the principal-agent relationship, by giving citizens an opportunity to adopt measures to directly observe politician performance. Specifically, the meetings intervention enables citizens to hold political leaders more accountable by gathering regular (quarterly) information on politician behavior vis-à-vis local health facility service delivery. Such political information pathways have been shown to be influential by a range of other studies.[6,7] Findings from my qualitative research also confirm the plausibility of this pathway, as do focus groups in Uganda conducted by

[6]See Pande (2011) for a review. For more recent studies, see, for example, Grossman and Michelitch (2018); Bidwell, Casey, and Glennerster (2020), Gottlieb (2016); Paler (2013); and Fujiwara and Wantchekon (2013).

[7]In the present study, however, it is possible that the meetings may also serve to increase citizen motivation. Disentangling motivation and information effects is a challenge in the literature and would be a suitable topic for future research on the meetings intervention. Additionally, it is possible that the meetings not only may provide citizens a source of information on politician performance but also may help to overcome citizen incomplete information regarding political and health rights.

Larreguy et al. (2017). Overall, targeting the unobservability of the citizen-politician relationship through community meetings offers a promising approach, since the second-best solution to the principal-agent problem — inferring politician performance based on the quality of health service provision — is likely infeasible. Beyond this aim, the meetings intervention also serves to address the second constraint discussed in Section 1.5.1 regarding politician incomplete information on health service quality. Because citizens — from their own experiences with local health facilities — appear to have sufficient knowledge of service quality, the meetings provide an opportunity for them to convey this information to political leaders.[8]

Altogether, these meeting intervention mechanisms suggest that greater citizen-politician interaction could help overcome key barriers to health service quality. Given the potential returns to citizen-politician meetings, what keeps citizens from organizing such meetings in the status quo? Several additional constraints may explain limited political participation by citizens, including the following: (1) organization, (2) information, (3) transaction costs, (4) free-rider effects, and (5) powerlessness. First, citizens likely require robust leadership to help provide structure for collective organization and to offer them a concrete course of action. Without such leadership, citizens may be unlikely to participate. In a perfectly functioning democracy — with perfect information flow between citizens and politicians and complete accountability of political leaders — we might expect the politicians themselves to provide the necessary leadership and resources to organize meetings with their constituents. In imperfect democracies or other political systems, third-party intervention may be required to invigorate collective action. Second, in addition to these organizational constraints, information deficiencies may limit citizen participation. Informational gaps may be as basic as a lack of awareness regarding opportunities to participate in collective action, which citizens

[8]For this channel, note that the role of citizen information on health service provision is not impaired by the observation that information on health service quality is a noisy signal of politician performance in the second-best solution to the principal-agent problem. In the principal-agent setting, the noisiness pertains to the ability of citizens to *attribute* observed health outcomes to politicians. In contrast, in the politician incomplete information mechanism, citizens only need to be generally aware of health service quality, in order to submit complaints to their leaders that spur politician action.

may need for the purposes of advance planning. Citizens may also experience incomplete information regarding their political rights, as well as their right to free, quality public healthcare in Uganda. Even if citizens are aware of their health and political rights, they may experience incomplete information regarding the political channels and systems through which to exercise them. Third, the need for transportation may create a transaction cost that amounts to a physical barrier to citizen participation. Especially in large jurisdictions where major political events are typically held in the center, citizens from more distant areas may be inhibited from participating.

Fourth, since democratic participation is a public good, attendance at such citizen-politician meetings is subject to free-rider effects that discourage participation. Fifth, since individual citizens (particularly non-elite citizens) in the political system are akin to price-takers in a competitive goods market, the complaints of one citizen alone are unlikely to influence the behavior of a political leader or create sufficient politician perceptions of electoral threats. This dynamic amounts to low bargaining power on the part of the average citizen and may be expressed by citizens as a sense of powerlessness — either real or perceived — to influence politician behavior. Taken together, these last two constraints of free-riding and price-taking make collective action an essential element of political processes.[9]

Providing a platform for citizens to engage with their local leaders, the meetings intervention aims to overcome the foregoing constraints. In particular, the community accountability meetings are organized by a non-governmental organization, thereby eliminating leadership and organizational constraints. In addition, the implementation of the meetings is accompanied by substantial mobilization efforts, which serve to alleviate informational barriers to participation. The meetings also take place in two locations within each governing jurisdiction, helping to make the events as accessible as possible. Finally, by bringing many community members together for the common cause of political accountability, the meetings

[9]Yet another constraint that may limit participation is motivation. Even when citizens have all of the necessary information and resources to adopt an accessible course of action, they may need additional persuasion to inspire them to join a collective action effort. For more discussion of motivation-related constraints, please see Chapter 1 of this dissertation.

address the free-rider and powerlessness constraints.[10] Of note, many of the constraints to political participation may disproportionately affect poorer citizens, a group that may be more likely to seek care from government health facilities. If wealthier citizens are less likely to obtain services from the public health centers and can also more easily access political leaders, politicians may develop an inaccurate viewpoint of local healthcare quality. Hence, bringing politicians and the general population together may help to ensure more accurate representation.[11]

4.2.2 Politician Training Program

In the skills training treatment arm, PHP followed a standardized curriculum to train chairpersons on their rights and responsibilities and on methods for monitoring local health facilities and taking action based on their findings. The curriculum addressed topics such as identifying patient complaints, checking for the presence of health workers, monitoring inventories, instituting improvements in local management practices, and using proper legal channels to pursue changes that may require intervention at higher government levels. In the context of the theoretical framework outlined in Section 1.5, the training intervention aims to address the human capital constraints that political leaders may experience.[12] As with the meetings program, PHP originally planned to carry out the training program over the course of three quarters, with a primary training workshop followed by two refresher workshops. However, because of the Covid pandemic, the organization thus far has only been able to complete the first round of training, which took place in January 2020 prior the

[10] One possible constraint on the effectiveness of the accountability meetings is the existence of social networks linking politicians and health workers. However, the inherent aim of the meetings is to generate sufficient citizen pressure through mass action to overcome any such opposing pressures on politicians. The influence of social networks on politicians is also another topic to address in a larger-scale study.

[11] For additional discussion of how the meetings intervention may impact health service provision, particularly in the context of a political environment with dysfunctional electoral processes, please see Appendix D.

[12] As another possible channel for this study arm, the training may not impart any new skills to politicians but may merely serve to remind them about health service delivery shortfalls, leading them to prioritize this area for oversight. As the pilot study is not able to tease apart these two interpretations, this distinction would be a priority for future research.

launch of the meetings program.

4.2.3 Impact Pathways

The programs may produce impacts in two primary areas: health and citizen satisfaction. To produce an impact on health outcomes, each intervention must induce a relatively complex cascade of effects on several intermediary outcomes. Specifically, the programs must first alter political incentives and dynamics. For the meetings intervention, such political effects must begin with strong grassroots participation (also referred to as a 'first-stage' effect), which itself must then be successful in influencing politician behavior. For the training intervention, the program must produce significant learning outcomes, which then also must lead to changes in politician behavior. Under both programs, the effects on political dynamics must in turn affect operational practices at health facilities. Finally, the changes in health center operations must then translate into improved health outcomes. At any stage in this pathway, binding constraints external to the interventions — for example, capacity limitations at health facilities — may thwart the realization of an impact, even if the employed programmatic strategies are effective.

Health outcomes, of course, are not the only contributor to social welfare. By fostering democratic participation, the meetings program in particular may also impact household and community well-being in areas related to self-determination, agency, and empowerment. Such impacts would likely be more intrinsic to the process of program implementation and therefore, compared to potential health impacts, would not require such an involved cascade of effects.

4.3 Data Collection

Carried out by PHP's field interview team, the data collection for the study consisted of a household survey, politician survey, and health facility quality assessment across all study arms, along with a meeting survey in the citizen-politician meetings arm. While both pre-

and post-intervention data collection was originally planned for the household, politician, and health facility surveys, the post-intervention data collection has not yet occurred due to the aforementioned pause in the project activities as a result of the Covid pandemic. However, because the primary objective of the study was feasibility assessment rather than impact evaluation, the round of baseline data collection, combined with nearly one full round of the midline meetings survey, proved sufficient for satisfying the core research goals. All baseline data collection took place between October and November 2019, while the meetings survey occurred between January and March 2020 in conjunction with program implementation.

All data collection instruments were electronically programmed via SurveyCTO and passed the rigor of multiple levels of development and review, including pre-testing with live subjects, quality assurance mechanisms for the fidelity of language translation, and bench testing for programming quality and errors. Each data collection tool was accompanied by an extensive training manual that covered general conduct, interviewer guidelines, data collection procedures, and question-specific instructions and definitions. In addition, for each survey tool, field interviewers went through a tailored, competitive training workshop of 5-8 days in length that concluded with a written examination. Due to the relatively small size of the study and accompanying budget, audits of the data collection process were conducted by the manager and assistant manager of the data collection team. The research received ethical approval from the University of Michigan Institutional Review Board, the Makerere University Research Ethics Committee, and the Uganda National Council for Science and Technology.

4.3.1 Outcomes of Interest

Based on the health impact pathway described in Section 4.2.3, the data collection for this study is designed to measure political outcomes, health services outcomes, and individual-level health impacts. Political outcomes — which demonstrate intervention success in altering local political dynamics — include participation in meetings with local leaders, citizen

perceptions of politicians, perceptions of health center quality by politicians, and changes in job practices by local leaders. Health services outcomes — which demonstrate the extent to which political changes alter facility practices — include a collection of measures that fall into one of four categories: (1) healthcare utilization, (2) healthcare quality, (3) intervention delivery and adoption, and (4) patient satisfaction. Health impacts include child mortality, weight-for-age, and mid-upper arm circumference.[13] In addition, apart from the health impact pathway, the data collection includes measures of citizen satisfaction to capture possible impacts stemming from enhanced self-determination, agency, and empowerment.

For the purpose of the present pilot study, with limited statistical power to evaluate all of the foregoing outcomes, I primarily focus on demonstrating the 'first-stage' effects of the interventions. Toward this end, key outcomes include measures of politician and citizen participation in the meetings, such as attendance, gender inclusiveness, elite vs. non-elite participation, and problems reported; feedback from households regarding their participation in the meetings; measures of politician participation in the training program; and feedback from political leaders regarding the quality of the training program.

In a longer-term study with greater statistical power, the aforementioned data collection will enable not only an evaluation of impact but also an analysis of which constraints on health service provision appear to be most active in the intervention mechanisms. In such a study, I will also aim to differentiate between alternative possible impact pathways that could be fostered by the interventions.

4.3.2 Household Survey

The household survey randomly sampled households from villages across the control and treatment localities. The survey lasted 1-1.5 hours and covered topics including following: agency, empowerment, and well-being; citizen knowledge, perceptions, participation, and satisfaction; community relations; healthcare utilization and patient satisfaction; child

[13] For more detailed discussion of the health outcomes of interest, focusing in particular on current healthcare delivery gaps in Uganda and the potential intervention impact pathways, please see Appendix E.2.

health, including anthropometric measurements; and household demographics. To define the catchment area for the survey, I carried out the following procedure. First, recall that PHP conducted the meetings intervention in two locations within each of the meeting treatment localities. One meeting took place in the parish/ward containing the Health Center III, while the other took place in another randomly chosen parish/ward. These two locations within each locality formed the basis of the household survey sampling. However, because the meeting implementation locations only applied to the meetings treatment group, the pure control and politician training only groups required an equivalent design to guide the household survey sampling process. Therefore, in such non-meeting localities, the data collection team similarly identified the parish/ward containing the Health Center III and randomly selected another parish/ward.

Following, for each selected parish/ward location, the data collection team identified a cluster of three central villages in which the survey sampling would take place. For the parish/ward containing the Health Center III, the first of the three villages was defined as the village containing the Health Center III. The second and third villages chosen were the nearest neighbors to the first village. For the other parish/ward that did not include the Health Center III, the first of three villages was defined as the village containing the parish/ward administrative headquarters, while the remaining two villages were again the nearest neighbors. Appendix B contains the determination criteria for the nearest neighboring villages. At the conclusion of the village selection process, each main study locality contained two clusters of three villages each, which formed the sampling area for the household survey.

The sampling procedure then proceeded according to the urbanization level of each locality. For the three of the twenty localities in the sample that were more densely populated town councils or divisions, the field interview team employed a random route method to sample households, as pre-survey enumeration of all of the resident households was impractical due to budgetary constraints. In accordance with this procedure, interviewers began at the center of each three-village cluster (either the Health Center III or the local government ward

headquarters) and then followed instructions that led them on a random route to sample households, while always remaining within the boundaries of the three villages. Immediately after sampling, field interviewers scheduled the survey administration with the household.

For the other seventeen, less densely populated localities, PHP's data collection team, in advance of the baseline household survey, carried out enumeration of all households in the sampling area to form the sampling frame. To complete the household enumeration process, the data collection team visited the village chairperson in each of the sample villages to request a copy of the village's household registry. Members of the data collection team then reviewed the registry with the chairperson, verifying that the list of households was comprehensive, accurate, and up-to-date. To be eligible for the survey, households were required to have one or more children under five as well as a respondent above 18 years of age. Therefore, as an initial eligibility screen, the data collection team asked the chairperson whether each household contained one or more children under five. In general, village chairpersons are knowledgeable about the households in their community. However, the question responses included a "Don't know" option. All households with "Yes" or "Don't know" responses for the child under five question were included in the sampling frame. In all cases, eligibility was confirmed directly with the household at the time of survey administration.

From the sampling frame, I then randomly selected households for the baseline survey, along with a list of replacements in case of ineligible, untraceable, or non-participating households. If an interviewer visited a household for the survey and determined that the respondent and/or household was temporarily absent, the household would be revisited. If a household was visited three times and an eligible and competent respondent remained unavailable each time, the household was dropped from the sample and replaced. To meet the sample goal of 500 households, twenty-five households from each locality were initially selected for the sample. To ensure that each cluster of sample villages with the localities would be roughly equally represented in the sample, the twenty-five households were divided into thirteen households from each Health Center III village cluster and twelve households

from each non-Health Center III village cluster. Ultimately, thanks to the productivity of the field interview team, the survey sample included 553 households.

4.3.3 Politician Survey

The politician survey lasted 1-1.5 hours and covered topics including the following: job background, knowledge, perceptions, and practices; health sector knowledge and perceptions; healthcare utilization; job-related activities in the health sector; constituent perceptions and interactions; and community relations. The sample for the survey was pre-defined, consisting of all of the Council Chairpersons for the twenty administrative units included in the study. Securing interview appointments with these local political leaders proved to be a relatively difficult process. In the end, the final survey sample consisted of sixteen of the twenty chairpersons, as the remaining subjects either refused to participate, could not be contacted, or were deceased.

4.3.4 Health Facility Quality Assessment

The health facility quality assessment consisted of an unannounced audit visit along with a more extended questionnaire, which was scheduled at the time of the initial audit visit. For the unannounced visit, the interviewer arrived at the health facility prior to the official opening time of 8:00AM and then discretely recorded observational data regarding the facility's operations, including the actual opening time. If and when the health center opened, the interviewer then conducted a personnel audit by comparing the facility's official personnel roster with the actual staff attendance. The scheduled questionnaire portion of the quality assessment lasted 2-3 hours and covered topics including the following: service delivery, patient experience, monitoring and supervision, community relations, infrastructure, equipment, drugs and supplies, healthcare utilization, and use of locally allocated healthcare funds. The equipment-related questions involved physical verification of the availability and functionality of selected items. Similarly, questions about inventories focused on a list of

tracer medicine and supplies and included physical verification of the available units, along with collection of administrative stock records. As with the politician survey, the sample for the health facility quality assessment was pre-defined, consisting of all of the Health Center III's in the twenty administrative units included in the study. As all of the health facilities consented to participate in the survey, the final sample consisted of twenty health centers.

4.3.5 Meeting Survey

The meeting survey involved a data collector-observer who recorded information about each community meeting as it was conducted, covering the following topics: attendance figures, according to different sub-populations (e.g., male vs. female, elite vs. non-elite); participation levels, again according to different sub-populations; and meeting discussion content.

5 Results

Both of the interventions yielded important results, demonstrating the promise of a follow-on study with greater statistical power. The key results fall into two categories: (1) achievements from program implementation and (2) outcomes of the interventions.

5.1 Implementation Outputs

From the standpoint of implementation, the pilot study has been a major success. In particular, the pilot has provided an opportunity to develop key program and research materials and protocols that will lay the foundation for an expanded study in the future. In the course of developing these materials, our research team adhered to stringent quality assurance and quality control measures — based on the idea that if the project sought to improve institutional quality and accountability, then we must exemplify those same priorities ourselves. Toward this end, implementation materials have included a wide range of standard operating procedures and reporting systems to guide each activity. This section highlights some of the

most critical outputs.

5.1.1 Research Implementation

As described above, research materials produced under the pilot have included extensive household, politician, health facility, and meeting surveys, all programmed electronically and accompanied by comprehensive training manuals, workshops, and examinations. In addition, our research team has developed in-field data auditing tools, along with a large body of computer code to facilitate the sampling process, carry out regular data quality checks, and perform final data processing. We have also created various reporting templates for the data collection team, as well as comprehensive checklists for all major activities, such as interviewer training and fieldwork, to ensure efficient implementation of the research.

In practice, the pilot has provided strong proof of concept for the feasibility of the data collection activities. Despite the political sensitivities of the research, the data collection exceeded expectations with the relative ease of its implementation. The household survey had a refusal rate of only 1.3 percent, and respondents expressed few reservations about the questions asked of them. While the respondents for the politician survey were somewhat challenging to secure for an appointment, many of them became much more enthusiastic about their participation once they learned further about the topics that the questionnaire addressed. The lessons learned from working with this group of political leaders to foster their participation in the research will be especially valuable for future data collection. The health facility quality assessment also initially posed concerns for our research team, in light of both its level of depth and its focus on health worker performance. However, as all health centers participated in the questionnaire, the survey was a significant success. To avoid disrupting the normal work flow of health facilities, our team developed approaches to adapt to the facility activities in carrying out the data collection. Ultimately, the questionnaire content proved to be well tailored to Health Center III's, capturing important information about their operations. For all of the foregoing surveys, our team of field interviewers also

reported that they felt respondents answered questions with relative honesty, an encouraging indicator of the quality of participation that we were able to secure. In addition, we experienced no adverse events in the course of any of the data collection. Lastly, the meeting survey presented its own challenges given the difficulties of collecting a wide breadth of observational data in real-time. However, with logistical tools developed by our research team and a clear plan of action for the conduct of the meetings, the survey progressed smoothly. Altogether, the quality of the data collection instruments, combined with the corresponding implementation protocols that we have created, will help enable an efficient transition to an expanded study.

5.1.2 Program Implementation

For the implementation of the two interventions, our team developed a similar breadth of tools. First, for the meetings intervention, we developed a handbook for meeting facilitators entitled *Citizen-Chairperson Dialogue Meetings: Implementation Guidelines*. The manual reviews the objectives of the meetings and the roles of meeting facilitators, presents background information on local government and health facility operations, establishes guidelines for the meeting mobilization process, and describes the process of meeting implementation, covering topics such as the logistics, discussion content, and meeting management.

For the training intervention, we developed two additional handbooks to guide its implementation. First, for the chairperson-trainees, we developed the *Health Sector Performance Initiative Health Leadership Manual*, available in both English and the local language of Runyankore. This manual for trainees presents a curriculum covering the LC3 Chairperson's roles and responsibilities in the government, detailed background information on health sector operations and related government stakeholders, techniques for monitoring local health service quality, and guidelines for taking action to improve healthcare delivery. Accompanying the trainee manual, our team developed the *Health Sector Performance Initiative Training Guide* for the facilitators of the training workshop. The training facilitation manual provides an

extended series of educational activities to carry out in presenting the contents of the trainee handbook at the training workshop. The facilitation manual is also accompanied by slides for the presentation of all of the material.

Alongside the development of program implementation guidelines, we also established a system of 'activity-based reporting,' through which we designed reporting requirements around each component activity of the interventions. Each report template contains detailed instructions with step-by-step procedures for the corresponding activity, a main report body, and an item-by-item checklist to guide the preparation for and implementation of the activity. By documenting and accounting for the activities so extensively, we minimized the error rate over the course of implementation. The checklists, which cover many of the mundane and easily forgettable tasks involved in program implementation, also served to free up the energy of the implementation team to focus on more intangible or discretionary tasks that also contribute to program quality.

Altogether, the collection of program guidelines and reporting tools help to ensure standardized, efficient, and high-quality implementation of the activities across different localities and facilitators. The protocols also make the interventions highly replicable, which will help to facilitate a larger-scale study subsequent to this pilot. Moving forward, in preparation for scale-up and based on lessons learned, our team will make additional tweaks to all of the implementation materials to further fine-tune them to the program activities.

5.2 Program Outcomes

5.2.1 Politician Training

The training intervention brought together the LC3 chairpersons from the ten training localities for a three-day workshop that covered the curricular materials described in Section 5.1.2. Initially, our program team was uncertain of what to expect from the training; in fact, we could not even be confident that the political leaders would attend the workshop. However,

in terms of both participation and learning outcomes, the program recorded achievements that exceeded expectations.

In particular, as the workshop drew the trainees away from their day jobs, and as a majority of the politician-trainees had to travel a significant distance to the training venue, the attendance of nine of ten chairpersons from the training treatment arm was itself a significant achievement. The tenth chairperson who did not attend sent another local leader as a delegate. While trainees sometimes circulated in and out of the workshop activities, we always achieved a participation rate of at least eighty percent throughout the workshop. More concretely, the training proved constructive on two main levels: (1) knowledge acquisition by the local leaders and (2) cross-community exchange of local government and health sector experiences. First, while the politician-trainees had some background knowledge on their general roles and responsibilities, the training filled critical knowledge gaps related to the operations of the health sector, the roles of political leaders in monitoring service provision, and remedies for performance problems in healthcare delivery. Many attendees remarked that they knew that they were capable of playing key roles to improve life in their communities, but that no other programs had empowered them with the knowledge and training to do so. Based on the comments from the trainees, it is clear that LC3 chairpersons occupy a unique but neglected position in their communities from which they can help foster more accountable local governance and service delivery. Second, in addition to the core training objectives, the workshop fostered a rich exchange of experiences between all of the chairpersons in attendance. By sharing lessons from their own communities, the trainees had a valuable opportunity to learn from one another, complementing the workshop curriculum.

At the conclusion of the training, the facilitators solicited feedback from the politician-trainees regarding the workshop. All trainees offered overwhelmingly positive comments, on topics ranging from the pedagogical methods employed to the content presented. They expressed significant gratitude for the opportunity to participate, along with enthusiasm to return for future workshops. In addition, they suggested that the government ultimately

integrate such workshops into standard orientation activities for political leaders. Altogether, the observations that our program team made during the training and the feedback that we received afterward evinced the value of the activities. While thus far, due to the Covid pandemic, it has only been possible to carry out one of three rounds of the politician training intervention, the achievements to date have been sufficient to demonstrate the significant promise that this program holds to impact health service provision.

5.2.2 Community Accountability Meetings

The meetings intervention proved similarly successful, achieving strong stakeholder engagement. Table 1 displays summary statistics on meeting attendance and participation. The data come from a sample of sixteen meetings across eight localities, due to the Covid-related pause in the intervention implementation before the two remaining localities in the meetings treatment arm could be covered. Based on population figures from the baseline household survey sampling frame, an average of 19.7 percent of households in each local catchment area sent one or more members to the meetings, with this measure of participation ranging from 12.5 percent to 33.0 percent across the meeting locations. On average, meetings attracted more than one hundred attendees, with the vast majority of these participants being members of the general community as opposed to individuals with official government leadership positions. In addition, these community members primarily came from the non-elite portion of the population, with members of the elite accounting for an average of only 1.2 percent of attendance by general community members.

In terms of gender representation, an average of 62.8 percent of the general community members present at each meeting were women. While the program aimed to elicit participation from typically underrepresented groups like women, it is also possible that the preferences of women — who in Uganda occupy the traditional role of managing household health matters, and who more frequently come into contact with the local health centers — make them more likely to attend the meetings. The discussions at the meetings reflected

TABLE 1: Meeting Attendance and Participation Summary Statistics

	Mean
Proportion Catchment Area Households Attended	0.197
Total Attendance	109.688
Proportion General Community Members	0.956
Proportion Politicians	0.039
Proportion Civil Servants	0.004
Total General Community Member Attendance	105.125
Proportion Male	0.372
Proportion Female	0.628
Proportion Elite	0.012
Total Comments by General Community Members	123.188
Proportion Male	0.389
Proportion Female	0.611
Proportion Elite	0.037

Notes: This table displays summary statistics for the meetings intervention. The main attendance figures are based on the number of individuals present, as opposed to the number of households represented. Each mean is an average across the sixteen meetings in the sample.

the composition of community members present, with men and women contributing meeting comments at rates commensurate with their level of representation. While elite community members appear to make comments at a higher rate than their level of representation, only 3.7 percent of comments on average are attributed to elites. This observation implies that the meetings do not suffer from elite capture, a major concern in local government affairs.

Qualitatively, the discussions at the meetings were vibrant, with ordinary citizens openly and passionately voicing their complaints about health service delivery. The most common complaints included health worker absenteeism and shortage, limited working days/hours of the local health center, lack of drugs and supplies, charges or bribes for services, and disrespect by health workers. In regard to the meetings themselves, community members expressed significant appreciation for the activities, indicating that it was uncommon to hold such gatherings and to be given a platform from which to express their concerns.

Alongside the observational data from the meetings, I also use the baseline household survey data linked with attendance data to assess the predictors of household attendance at the meetings. Table 2 presents the results of a linear probability model regressing meeting

attendance on a range of baseline dimensions reported by households. Overall, relatively few covariates are significant predictors of household attendance. First, households that report greater satisfaction with the community are, for each level of satisfaction measured on a 1-5 Likert scale, 7.4 percentage points less likely to attend meetings. Second, households that report that households like their own are able to influence the decisions made by local government officials in the community are 14.0 percentage points more likely to attend meetings. Third, households that report having sent a member to one or more Local Council III meetings over the past 12 months are 23.6 percentage points less likely to attend the accountability meetings. As the Local Council III is the top governing body in each locality and often attracts the participation of elites, this large effect may reflect an aversion of local elites to attend the meetings, perhaps because they perceive them (correctly) as events for the general community. Of note, education and economic status are not significant predictors of meeting attendance. These findings suggest that households from a range of backgrounds were attracted to the meetings — an encouraging result in light of the program's goals related to representation.

TABLE 2: Predictors of Meeting Attendance by Households

	Attended Meeting
Satisfaction with Health	0.000
	(0.029)
Satisfaction with Community	-0.074**
	(0.029)
Able to Influence Local Government	0.140*
	(0.081)
Able to Improve Government Health Services in Community	-0.028
	(0.091)
Attended Village Meetings	0.005
	(0.072)
Attended LCIII Meetings	-0.236**

	(0.095)
Community Would Join Together to Improve Health Services	-0.088
	(0.083)
Chances of Community Pressuring Chairperson	-0.029
	(0.032)
Have Sufficient Information to Evaluate Local Health Services	-0.068
	(0.075)
Drugs and Supplies Usually Available	0.096
	(0.080)
Health Center Staff Normally Attend Work	0.109
	(0.119)
Travel Time to Health Center	-0.000
	(0.001)
Number Health Center Visits in Past 12 Months	-0.002
	(0.010)
Number Visits to Other Providers in Past 12 Months	-0.009
	(0.011)
Number Household Members	-0.013
	(0.015)
Respondent Highest Grade in School	0.016
	(0.012)
Household Head Highest Grade in School	-0.015
	(0.012)
Economic Index	0.009
	(0.022)
Mean Dependent Variable	0.416
Observations	231

Notes: This table displays the results of the meeting attendance predictors analysis. Heteroskedasticity-robust standard errors are shown in parentheses below the regression coefficient estimates. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level.

Aside from attendance by households, the meetings also achieved strong participation

by the chairpersons. In general, chairpersons found the meetings to provide an appealing opportunity to meet with their constituents. However, in two of the sample localities, the chairpersons declined to participate, providing explanations that insinuated their expectation of a monetary inducement for attending the meetings. In a third locality, the chairperson had to forego one of the two meetings due to an emergency. Although the sample size is small, these outcomes amount to an intervention adoption rate by chairpersons of roughly 70 percent. I interpret this adoption rate as strong, even though chairperson attendance would ideally be perfect, given the central role of the chairpersons in the program. In addition, even in the absence of chairpersons, meetings remained at least partially productive. In all of the instances in which the chairpersons did not attend, they delegated the vice chairpersons to attend the meetings in their place. Meetings involving the vice chairpersons were more challenging, as the deputies had less authority and also often less knowledge compared to the chairpersons. Nevertheless, our program team felt that these meetings were still helpful in bringing community members together to address important health matters. Furthermore, while it has not been possible to observe the effects of repeated meetings due to the Covid-related activity stoppage, the program team felt that the non-participating chairpersons likely would have become more cooperative as more meetings were implemented. The team additionally reported that chairpersons who were also part of the study's training arm tended to be more cooperative, suggesting complementarity between the meeting and training programs.[14] Moving forward, our program team will explore further strategies to induce maximal attendance by chairpersons.[15]

Altogether, the varied findings on the meetings intervention demonstrate the program's success in terms of both level of attendance and level of representation. In line with the original program objectives, the meetings induced significant participation from a diverse

[14]This observation may be due to the lessons that the chairpersons learned during the training, but it may also be due to the prior relationship that our staff had established with them as a result of the training — or a combination of both.

[15]We do not, however, envision offering any monetary inducements, as we do not believe such an approach would be aligned with the program's focus on accountability and on the use of existing public resources more effectively.

group of households. As with the training program, the outcomes provide strong evidence in support of the meeting program's potential to ultimately impact health. In addition, in communities where the LC3 chairperson was also among the group of leaders who received the training intervention, our program implementation team reported that the meetings seemed more effective, with the chairpersons suggesting more specific and more practical courses of action to address healthcare delivery gaps. This observation confirms the possibility of interactive effects between the two interventions.

6 Preliminary Pre-Analysis Plan for At-Scale Study

Based on the results of the pilot, this section presents a preliminary, condensed pre-analysis plan for a follow-on study with sufficient statistical power to rigorously evaluate the interventions. In general, the at-scale study will employ the same research design that has been described above for the pilot. To allow sufficient time to observe measurable effects on health services, the expanded study's intervention implementation period will last for approximately two years in between the baseline and endline data collection.

6.1 Research Questions

Along with quantitatively evaluating the overall impact of the programs, the at-scale study will enable several additional lines of inquiry. Specifically, the study will be designed to answer the following research questions:

1. *What are the causal effects of the meetings intervention, the training intervention, and the two interventions implemented together?*

   This question will answer the most fundamental objective of the study and will elucidate the impact of the programs on politician behavior, health facility operations, health outcomes, and citizen satisfaction. A greater effect from the jointly implemented program may suggest complementarity between the two interventions.

2. *What are the mechanisms of the interventions? What underlying constraints do the programs alleviate?*

   Based on the hypothesized theoretical framework and intervention mechanisms, this question will provide insight on the specific governance and institutional dynamics at play in the Ugandan health sector. The answers to this question may also point to additional promising points of intervention.

3. *Do the effects of the interventions depend on the competitiveness or functioning of the surrounding political environment in a given locality?*

   This question will use the study's stratified design — based on recent margin of victory as a proxy for political competitiveness — to shed light on the influence of underlying political factors in producing a program impact.

4. *Do the interventions produce any negative externalities for localities neighboring the treated localities?*

   If, in response to the interventions, local politicians begin lobbying more effectively on behalf of their communities to higher-level governing bodies (e.g., the district government), it is possible that resources allocated to the intervention localities will trade off with resources allocated to the non-intervention localities within the same legal jurisdiction. By carrying out data collection on resource allocations made to both treated and untreated localities within the same overall jurisdiction, the study will help to answer this question.

5. *Within the meetings arm, do changes in politician behavior depend on the number of persons who attend the meetings?*

   Stemming from the observations in Section 4.2.1 that collective action plays a key role in political processes, this question will provide direct evidence on the influence of collective action on politician behavior. Specifically, I will further randomize the localities

within the meetings intervention group across different pre-meeting mobilization activities, based on approaches such as the encouragement messages studied in Chapter 1. The random assignment of mobilization procedures will create exogenous variation in the size of the meetings, enabling a credible evaluation of this question.

6.2 Statistical Framework

To answer the study's primary question about the causal effects of the interventions, I will estimate the following regression equation:

$$Y_{ij} = \alpha + \beta_m M_j + \beta_s S_j + \beta_{ms} MS_j + u_{ij}, \tag{1}$$

where Y_{ij} denotes the outcome variable of interest for household i in locality j, M_j is an indicator for the meetings only treatment, S_j is an indicator for the skills training only treatment, and MS_j is an indicator for the joint treatment. The coefficients of interest — β_m, β_s, and β_{ms} — represent the intention-to-treat effect of the respective intervention. To support the internal validity of the experiment, I will test for pre-program balance and differential attrition between the control and treatment groups. All statistical analysis will use heteroskedasticity-robust standard errors adjusted for clustering at the locality level.[16]

6.3 Power Calculations

The data from the pilot enable power calculations to determine the necessary scale of the expanded cluster randomized controlled trial. Given the wide scope of data collected for the pilot, the data for many of the key outcomes that will be used in the at-scale study have not yet been fully processed. As more outcome data become available for analysis, I will carry out further power calculations, which will be presented in a complete version of the pre-analysis plan. At present, I perform power calculations for the following two citizen

[16] In a complete version of the pre-analysis plan, I will present fully specified hypothesis tests as well as fully specified regressions to evaluate the additional research questions outlined above.

satisfaction outcomes from the household survey: satisfaction with health and satisfaction with community, both of which are measured on a 1-5 Likert scale.

For each outcome of interest, the trial will test the null hypothesis $H_0 : \mu_T = \mu_C$, where μ_T and μ_C represent the means of the outcome for a given treatment group and the control group, respectively. Let $\text{var}(\mu_T) = \sigma_T^2$ and $\text{var}(\mu_C) = \sigma_C^2$, and let the minimum detectable effect $d = \mu_T - \mu_C$. Based on this setup, I specify all of the power calculations for two-sided t tests between two populations, with significance level $\alpha = 0.05$ and power $1 - \beta = 0.8$ (where $\beta = 0.2$ is the type II error probability). For a cluster randomized controlled trial with clusters of equal size, the required sample size per arm is given by the following equation (Hayes and Bennett 1999; Hemming et al. 2011; Hemming and Marsh 2013):

$$n = (\sigma_T^2 + \sigma_C^2)\left\{\frac{(z_{\alpha/2} + z_\beta)^2}{d^2}\right\}\Big\{1 + (m-1)\rho\Big\}, \tag{2}$$

where the z values are determined using the inverse cumulative standard normal distribution at the specified probabilities, m is the fixed cluster size, and ρ is the intracluster correlation (ICC). In the foregoing equation, the quantity $1 + (m-1)\rho$ is known as the design effect, while the quantity $(\sigma_T^2 + \sigma_C^2)\left\{\frac{(z_{\alpha/2}+z_\beta)^2}{d^2}\right\}$ gives the sample size that would be required under individual randomization (Donner, Birkett, and Buck 1981; Donner 1998). Accompanying equation (2), the required number of clusters per arm is given by

$$k = \lceil \frac{n}{m} \rceil + 1, \tag{3}$$

where the addition of one cluster is made to permit the use of the t distribution (Hayes and Bennett 1999).

For the power calculations, I set the mean of each variable to μ_C and then choose d, which implies μ_T. In addition, I assume that $\sigma_T^2 = \sigma_C^2$. After obtaining the minimum sample size and minimum number of clusters using equations (2) and (3), respectively, I make two adjustments. First, I round up the sample size so that it is a multiple of the cluster size.

Second, I increase both the cluster size and the total sample size by ten percent to account for possible attrition between baseline and endline data collection.

In regard to the outcomes in question, the satisfaction with health variable has a mean of 3.546, a standard deviation of 1.367, and an ICC of 0.019, while the satisfaction with community variable has a mean of 3.932, a standard deviation of 1.331, and an ICC of 0.000. For each of these outcomes, I consider a minimum detectable effect corresponding to a ten percent increase over the mean. Setting $m = 30$ and applying all of the values to equation (2), I find that satisfaction with health is the limiting variable, requiring a larger sample size for the minimum detectable effect. Based on equations (2) and (3) alone, the study requires a sample size of 360 households across 13 clusters for each arm. After incorporating the additional aforementioned adjustments in the computation, I arrive at the final result of 429 households across 13 clusters for each experimental arm. Given the study's four arms, these figures imply a total sample size of 1,716 households across 52 clusters. This projected sample size is relatively small. In reality, power calculations for the additional outcome variables of interest that have not yet been fully processed will likely yield a substantially larger required sample size.

7 Discussion

Many studies in development economics focus on the introduction of "novel" interventions implemented in clean, idealized settings; but the generalizability of these studies in terms of translating such interventions into national programs depends on background political economy factors (Acemoglu 2010; Deaton 2010). With potentially broad implications, this study instead directly targets political economy inefficiencies and structural governance factors that likely have pervasive incentive effects across multiple service sectors. Rather than introducing a new, often externally supported technology, the approach here focuses on using existing resources more efficiently and may be an especially promising strategy for sustained

development.

In this fashion, the primary aim of this study has been to test the feasibility of two programs — community accountability meetings and politician training — along with corresponding data collection activities. Toward this end, the results demonstrate significant potential for both interventions to improve health service provision. For each of the programs, I record substantial 'first-stage' effects, with strong, representative attendance and participation achieved by the meetings and important learning outcomes registered by the training. In addition, despite the complexities of many of the research instruments, all of the data collection activities proceeded smoothly. Beyond these outcomes, our team has also used the pilot to develop a rich set of implementation protocols and tools, creating a strong foundation for future work.

Building on these accomplishments and using the data from the pilot, I present an initial pre-analysis plan for a follow-on study conducted at-scale with sufficient statistical power. The pre-analysis plan outlines the main research questions of the scaled study, establishes the overall statistical framework for the measurement of program effects, and provides power calculations for the required sample size. The plan, which will be developed in greater depth as time goes on, offers a concrete roadmap for a high-quality quantitative evaluation.

Altogether, this study investigates supply-side barriers to quality healthcare delivery, contributing to the broader development economics literature on health interventions. Within this literature, many studies focus on demand-side barriers to health, addressing "puzzles" related to why poor households do not invest in health interventions that seem to produce significant private returns (Dupas 2011). While demand-side barriers represent one part of the equation, supply-side factors are likely another essential side of the story. Indeed, it is possible that poor health service quality itself contributes to low demand for health services. By examining both politician capacity-building and citizen political participation in the context of the health sector, this study sheds light on the nature of governing institutions and important areas for reform. In addition, the meetings intervention in particular — by giving

citizens a platform to demand improved health services, which in turn may lead to improved healthcare provision — connects both ends of the supply and demand equation.

At a macro level, multiple studies in economics highlight the role of institutions in development (Acemoglu and Robinson 2001; Hall and Jones 1999; Acemoglu, Johnson, and Robinson 2002). However, because of the aggregated nature of data, along with data quality concerns, reverse causality and omitted variable bias often complicate inference about the causal effect of institutions on development. The research design of both this pilot and the planned at-scale study, as randomized controlled trials using micro-level data, serves to provide more credible evidence on the impact of institutions on development.

The essence of the two programs in this study — with their focus on institutional quality — is to remedy the current state of health and governance affairs by better operationalizing the law. How can we close the gap between "what is" and "what should be"? In establishing the framework for a larger, subsequent trial that rigorously evaluates the interventions and their associated mechanisms of action, this study contributes toward this key priority.

Acknowledgments

I thank Hoyt Bleakley, John Bound, Fred Muhumuza, Edward Norton, and Dean Yang for their feedback on this paper. I also thank Jordan Bateisibwa, Sam Twesigye, Joram Amanya, and Alex Niwagaba for the instrumental roles they played in the data collection. Equally important, I am grateful to Timothy Lubanga, Gonzaga Mayanja, and Joseph Muserero of the Office of the Prime Minister for helping to foster the activities that made this work possible. Finally, I thank Christine Convery for her financial management of project activities.

Funding

Research reported in this publication was supported by the Abdul Jameel Latif Poverty Action Lab (J-PAL) (grant number GR-0947), the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health (award number T32HD007339), the National Science Foundation (grant number 1824290), the U.S. Department of Education (grant number P022A180037), the National Institute of General Medical Sciences of the National Institutes of Health (award number T32GM007863), the University of Michigan Department of Economics (Jordan and Kim Dickstein African Initiative Fund Research Award, Summer Research Apprenticeship), the University of Michigan Medical School Global REACH Program (Global Health Project Grant and Global Health Travel Grant), the University of Michigan Population Studies Center (Weinberg Research Prize), and the University of Michigan Rackham Graduate School (Rackham International Research Award). The content of the article is solely the responsibility of the author and does not necessarily represent the official views of any of the funders.

Disclosures

Alongside his academic position, the author serves as the Chief Executive Officer and as a board member of Progressive Health Partnership (PHP), the organization that carried out

the intervention and data collection activities for this article. The author's relationship with PHP is unpaid.

Ethics Statement

Research reported in this publication was approved by the University of Michigan Institutional Review Board, the Makerere University Research Ethics Committee, and the Uganda National Council for Science and Technology.

References

Abuya, Timothy, Charlotte E Warren, Nora Miller, Rebecca Njuki, Charity Ndwiga, Alice Maranga, Faith Mbehero, Anne Njeru, and Ben Bellows. 2015. "Exploring the Prevalence of Disrespect and Abuse during Childbirth in Kenya." *PloS One* 10 (4): e0123606.

Acemoglu, Daron. 2010. "Theory, General Equilibrium, and Political Economy in Development Economics." *Journal of Economic Perspectives* 24 (3): 17–32.

Acemoglu, Daron, Simon Johnson, and James A Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117 (4): 1231–1294.

Acemoglu, Daron, and James A Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.

Awortwi, Nicholas. 2010. "The Past, Present, and Future of Decentralisation in Africa: A Comparative Case Study of Local Government Development Trajectories of Ghana and Uganda." *International Journal of Public Administration* 33 (12-13): 620–634.

Bardhan, Pranab, and Dilip Mookherjee. 2006. "Decentralisation and Accountability in Infrastructure Delivery in Developing countries." *Economic Journal* 116 (508): 101–127.

Bidwell, Kelly, Katherine Casey, and Rachel Glennerster. 2020. "Debates: Voting and Expenditure Responses to Political Communication." *Journal of Political Economy* 128 (8): 2880–2924.

Björkman, Martina, and Jakob Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-based Monitoring in Uganda." *Quarterly Journal of Economics* 124 (2): 735–769.

Bold, Tessa, and Jakob Svensson. 2013. "Policies and Institutions for Effective Service Delivery: The Need of a Microeconomic and Micropolitical Approach." *Journal of African Economies* 22 (Supplement 2): ii16–ii38.

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

Chopra, Mickey, Elizabeth Mason, John Borrazzo, Harry Campbell, Igor Rudan, Li Liu, Robert E Black, and Zulfiqar A Bhutta. 2013. "Ending of Preventable Deaths from Pneumonia and Diarrhoea: An Achievable Goal." *Lancet* 381 (9876): 1499–1506.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48 (2): 424–455.

Devas, Nick, and Ursula Grant. 2003. "Local Government Decision-Making—Citizen Participation and Local accountability: Some Evidence from Kenya and Uganda." *Public Administration and Development* 23 (4): 307–316.

Dizon-Ross, Rebecca, Pascaline Dupas, and Jonathan Robinson. 2017. "Governance and the Effectiveness of Public Health Subsidies: Evidence from Ghana, Kenya and Uganda." *Journal of Public Economics* 156:150–169.

Donner, Allan. 1998. "Some Aspects of the Design and Analysis of Cluster Randomization Trials." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 47 (1): 95–113.

Donner, Allan, Nicholas Birkett, and Carol Buck. 1981. "Randomization by Cluster: Sample Size Requirements and Analysis." *American Journal of Epidemiology* 114 (6): 906–914.

Dupas, Pascaline. 2011. "Health Behavior in Developing Countries." *Annual Review of Economics* 3 (1): 425–449.

Filmer, Deon, and Lant H Pritchett. 2001. "Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Educational Enrollments in States of India." *Demography* 38 (1): 115–132.

Fujiwara, Thomas, and Leonard Wantchekon. 2013. "Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin." *American Economic Journal: Applied Economics* 5 (4): 241–55.

Gottlieb, Jessica. 2016. "Greater Expectations: A Field Experiment to Improve Accountability in Mali." *American Journal of Political Science* 60 (1): 143–157.

Grossman, Guy, and Kristin Michelitch. 2018. "Information Dissemination, Competitive Pressure, and Politician Performance between Elections: A Field Experiment in Uganda." *American Political Science Review* 112 (2): 280–301.

Hall, Robert E, and Charles I Jones. 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114 (1): 83–116.

Hayes, RJ, and S Bennett. 1999. "Simple Sample Size Calculation for Cluster-Randomized Trials." *International journal of epidemiology* 28 (2): 319–326.

Hemming, Karla, Alan J Girling, Alice J Sitch, Jennifer Marsh, and Richard J Lilford. 2011. "Sample Size Calculations for Cluster Randomised Controlled Trials with a Fixed Number of Clusters." *BMC Medical Research Methodology* 11 (1): 1–11.

Hemming, Karla, and Jen Marsh. 2013. "A Menu-Driven Facility for Sample-Size Calculations in Cluster Randomized Controlled Trials." *The Stata Journal* 13 (1): 114–135.

Hölmstrom, Bengt. 1979. "Moral Hazard and Observability." *Bell Journal of Economics*: 74–91.

Kakumba, Umar. 2010. "Local Government Citizen Participation and Rural Development: Reflections on Uganda's Decentralization System." *International Review of Administrative Sciences* 76 (1): 171–186.

Kiyaga-Nsubuga, John, and Yasin Olum. 2009. "Local Governance and Local Democracy in Uganda." *Commonwealth Journal of Local Governance*, no. 2: 26–43.

Kyaddondo, David, Kidza Mugerwa, Josaphat Byamugisha, Olufemi T Oladapo, and Meghan A Bohren. 2017. "Expectations and Needs of Ugandan Women for Improved Quality of Childbirth Care in Health Facilities: A Qualitative Study." *International Journal of Gynecology & Obstetrics* 139 (S1): 38–46.

Larreguy, Horacio, Benjamin Marx, Otis Reid, and Christopher Blattman. 2017. *A Market Equilibrium Approach to Reduce the Incidence of Vote-Buying: Evidence from Uganda*. Working Paper.

Leonard, Kenneth L, and Melkiory C Masatu. 2010. "Using the Hawthorne Effect to Examine the Gap between a Doctor's Best Possible Practice and Actual Performance." *Journal of Development Economics* 93 (2): 226–234.

Lewis, Maureen, and Gunilla Pettersson. 2009. *Governance in Health Care Delivery: Raising Performance*. Policy Research Working Paper 5074. Washington, DC: World Bank.

Liu, Li, Shefali Oza, Daniel Hogan, Jamie Perin, Igor Rudan, Joy E Lawn, Simon Cousens, Colin Mathers, and Robert E Black. 2015. "Global, Regional, and National Causes of Child Mortality in 2000–13, with Projections to Inform Post-2015 Priorities: An Updated Systematic Analysis." *Lancet* 385 (9966): 430–440.

McPake, Barbara, Delius Asiimwe, Francis Mwesigye, Mathias Ofumbi, Lisbeth Ortenblad, Pieter Streefland, and Asaph Turinde. 1999. "Informal Economic Activities of Public Health Workers in Uganda: Implications for Quality and Accessibility of Care." *Social Science & Medicine* 49 (7): 849–865.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–249.

Paler, Laura. 2013. "Keeping the Public Purse: An Experiment in Windfalls, Taxes, and the Incentives to Restrain Government." *American Political Science Review* 107 (4): 706–725.

Pande, Rohini. 2011. "Can Informed Voters Enforce Better Governance? Experiments in Low-Income Democracies." *Annual Review of Economics* 3 (1): 215–237.

Pritchett, Lant, and Michael Woolcock. 2004. "Solutions when the Solution is the Problem: Arraying the Disarray in Development." *World Development* 32 (2): 191–212.

Raffler, Pia. 2020. *Does Political Oversight of the Bureaucracy Increase Accountability? Field Experimental Evidence from a Dominant Party Regime*. Working paper.

Ravanilla, Nico. 2016. "Attracting Good People into Public Service: Evidence from a Field Experiment in the Philippines." Unpublished.

Republic of Uganda. 2000. "Laws of Uganda." Chap. 243: The Local Governments Act. Kampala, Uganda: Republic of Uganda.

Ross, Stephen A. 1973. "The Economic Theory of Agency: The Principal's Problem." *American Economic Review* 63 (2): 134–139.

Sharma, Gaurav, Matthews Mathai, Kim E Dickson, Andrew Weeks, G Justus Hofmeyr, Tina Lavender, Louise Tina Day, Jiji Elizabeth Mathews, Sue Fawcus, Aline Simen-Kapeu, et al. 2015. "Quality Care during Labour and Birth: A Multi-Country Analysis of Health System Bottlenecks and Potential Solutions." *BMC Pregnancy and Childbirth* 15 (2): S2.

Tripp, Aili Mari. 2010. *Museveni's Uganda: Paradoxes of Power in a Hybrid Regime*. Vol. 6. Lynne Rienner Publishers Boulder, CO.

UBOS and ICF. 2018. *Uganda Demographic and Health Survey 2016*. Country Report. Kampala, Uganda and Rockville, Maryland: Uganda Bureau of Statistics and ICF International Inc.

UN IGME. 2020. *Levels and Trends in Child Mortality: Report 2020*. Technical Report. New York, NY: UN Inter-agency Group for Child Mortality Estimation.

World Bank. 2016. *Health Service Delivery in Uganda*. Technical Report AUS5359. Washington, DC: World Bank.

— . 2003. *World Development Report 2004: Making Services Work for Poor People*. Washington, DC: World Bank.

— . 2017. *World Development Report 2017: Governance and the Law*. Washington, DC: World Bank. doi:10.1596/978-1-4648-0950-7.

APPENDIX

A. Study Population of Localities

This section details how the population of localities — from which the sample localities were drawn — was defined. While the locality population included subcounties/town councils/divisions throughout the Ankole Region of Uganda, three modifications were made to the full list of localities in this region to establish the population. First, due to budgetary constraints, localities that were prohibitively far from PHP's main office in Mbarara City were excluded from the population. Second, as some localities in the Ankole Region either had no Health Center III at all (contrary to government standards) or had higher-level government health facilities, they were also excluded to maintain uniformity of the local health service infrastructure vis-à-vis the study interventions and data collection. Third, due to an increasing trend of administrative unit division by the government in Uganda, the boundaries of some localities had changed since the 2016 election. Because boundary changes would likely render the 2016 margin of victory variable (which was used for the stratification) inaccurate, any localities that had experienced such changes were excluded from the population. Some of these units also included newly formed localities, which accounted for the absence of Health Center III's in some cases.

B. Sampling Area Village Selection Criteria

As described in Section 4.3.2, the sampling area consisted of clusters of three villages within each parish/ward included in the study. The first village in each cluster was defined as the village containing either the Health Center III or the parish/ward administrative headquarters, depending on the parish/ward in question. The remaining two villages were the nearest neighbors to the first village. The nearest neighboring villages were defined as the villages "of greatest accessibility" to the first village. Specifically, enumerators received the following guidelines to determine accessibility.

The term “of greatest accessibility” is intended to specify those villages that are most accessible to the Central Village. In determining the most accessible villages, you will need to exercise appropriate judgement. You should adhere to the following criteria, listed in order of importance, to determine accessibility.

1. *Time:* The length of an average trip from one village to the other should be the primary determinant of accessibility. Often, time will be related to the physical distance separating two villages. However, in some unusual cases—for instance, where the terrain is difficult or where the walking paths or roads are not direct—two villages could be closer to each other in terms of straight-line physical distance but further from each other in terms of time.

2. *Distance:* If judgements related to time are unclear, physical distance can help distinguish which villages are most accessible to the Central Village. The shorter the physical distance, the more accessible the village is.

3. *Borders:* Depending on the geographic shape of the villages, it is possible for multiple villages to border the Central Village. Some bordering villages may have longer shared borders than others. In general, villages with longer bordering areas should be considered more accessible to the Central Village.

4. *Population:* If after consideration of the above three criteria it remains difficult to differentiate between the level of accessibility of two or more villages, population should be used to “break the tie.” You may gather a rough estimate of the populations of each village and favor the village with greater population.

Upon meeting the LC1 Chairperson, use your skills to identify the two nearest neighboring villages, working to ensure that the LC1 Chairperson does not give you biased responses. For instance, in considering the abovementioned criteria, you may ask questions such as the following:

- Where do I go to next after this village? And then where?
- Which other villages are the greatest users of the Health Center III?
- If there is a meeting at the Parish/Ward Headquarters, residents from which villages are most likely to attend?

C. Computation of Economic Index

The covariates used for both the predictors regression shown in Table 2 include an economic index variable that proxies for household economic status. I construct the economic index using principal component analysis (PCA) of twenty-one variables. PCA finds the orthogonal linear combinations of a set of variables that best capture the overlapping information contained by the variables; the first principal component is the linear combination with maximum variance and is used as the economic index. See Filmer and Pritchett (2001) for more discussion of the method and for a demonstration of its validity in constructing such an index.

The variables in the index include the following: ratings of (1) the quality of house external wall construction material and (2) the quality of house floor construction material; indicators for ownership of each of the following by any household member: (3) bicycle, (4) motorcycle, and (5) car or truck; indicators for possession/ownership of each of the following: (6) non-solar electricity, (7) solar panel, (8) radio, (9) television, (10) mobile telephone, (11) refrigerator, (12) clock or watch, (13) bed/mattress, (14) chair/stool, (15) table, and (16) mosquito bednet; (17) number of acres of agricultural land owned by household members; and number of each of the following livestock animals owned by the household: (18) chickens, (19) goats, (20) pigs, and (21) cows.

For the variable measuring the quality of house external wall construction material, the rating is given a value of 1 for mud and pole; 2 for unburnt bricks with mud or unburnt bricks with cement; 3 for burnt/stabilized bricks; and 4 for cement blocks, concrete/stone, or wood. For the variable measuring the quality of house floor construction material, the rating is given a value of 1 for rammed earth; 2 for bricks or stone; 3 for cement screed, concrete, or wood; and 4 for tiles. The household survey also collected information on the construction material used for the roof, but this measure exhibited virtually zero variability across households and therefore was not included in the economic index.

D. The Ugandan Political Environment

This section provides a more detailed assessment of the likely impact of the meetings intervention in the context of the Ugandan political environment. In non-democratic settings, a fundamental question is whether the meetings intervention truly has the potential to impact health service provision. Indeed, it is possible that the success of the intervention may hinge on the functioning of local elections. If local electoral processes do not function well, the incentives of political leaders to adopt new behaviors in the health sector may be undercut. For instance, a couple recent studies in Uganda have observed effects that depend on underlying features of the local political environment (Grossman and Michelitch 2018; Raffler 2020). Such conditioning interactions introduce another intellectually interesting dimension and underlie the stratified experiment design of this study.

In regard to Uganda specifically, several considerations suggest that the country's political environment provides an appropriate setting for the study. First, the Ugandan government has been described as a hybrid, semi-authoritarian regime (Tripp 2010). While Uganda does not exhibit democratic principles at the national executive level, citizens exercise substantial freedom in other areas of governance, and the country's local political systems operate according to different norms. Although evidence on local electoral outcomes is limited, lo-

cal elections have been reported to be relatively competitive, with high turnover of local government councilors each cycle (Awortwi 2010; Kiyaga-Nsubuga and Olum 2009).

Still, there is no doubt that local governance in Uganda, as in many low- and middle-income countries, faces a range of systemic weaknesses, which may introduce additional constraints on political processes and service delivery. However, many of these weaknesses constitute the very motivation for the citizen-politician meetings intervention. The intervention is directly aimed at strengthening democratic practices and may therefore serve as a possible counterbalancing strategy to address shortfalls in democratic functioning. While the underlying causes of status-quo behaviors such as clientelism and electoral malpractice are complex, one contributing factor is likely limited citizen participation in political processes. Without sufficient information on politician performance, for example, citizens may be constrained in making informed voting decisions and therefore may maximize utility merely based on the clientelism of politicians.

In fact, such a dynamic has been observed in Benin by Fujiwara and Wantchekon (2013), who show that information-based campaigns — through candidate-endorsed town hall meetings discussing specific policy platforms — successfully reduce clientelism in presidential politics. Thus, the provision of performance information — implemented herein through the citizen-politician meetings — may serve to shift the status-quo equilibrium toward one in which citizens make decisions more informed by the realities of service delivery. Providing further evidence of this possibility in the setting of Uganda, Larreguy et al. (2017) describe focus group discussions that reveal that citizens have little information to assess candidates, with elected officials arguing that they are not responsible for service delivery and voters being uninformed about the appropriate parties to hold accountable for service quality problems. Results from the anti-vote buying campaign that the authors evaluate also confirm the plausibility of Ugandan citizens changing their voting behaviors (to vote for their preferred candidates) in response to information.

Lastly, as several studies in Uganda have pointed out, policy-focused citizen participation and consultation during non-electoral time periods is a persistent gap (Kakumba 2010; Devas and Grant 2003). Such non-electoral participation would complement the role of elections, which tend to be fairly blunt accountability mechanisms. Pointing to the promise of the meetings intervention, Grossman and Michelitch (2018) observe positive effects on district politician performance of a transparency-scorecard intervention implemented in the middle of the electoral term. Notably, politicians may be driven by other forms of incentives in addition to electoral accountability. For instance, politicians are likely to be motivated by an aversion to public or professional humiliation. Even local political leaders who are confident in their electoral prospects are likely to prioritize positive relations with their constituents. Should the performance of political leaders fall short of citizen expectations, it is possible that leaders at higher levels of government will learn of the circumstances, providing an additional source of professional pressure on local politicians.

E. Health Service Delivery in Uganda

E.1 Health Sector Performance

This sub-section offers a more in-depth assessment of the state of health service delivery in Uganda around the time of the study. Some of the most recent evidence on the quality

of health service provision comes from the World Bank's (2016) *Service Delivery Indicators* survey, which was carried out in 2013 across 387 health facilities throughout the country. The results of this survey demonstrate the persistence of alarming service delivery deficiencies. For example, health worker absenteeism rates exceed 47 percent in the public health sector, a level that suggests a decline in performance compared to the absenteeism rate of 37 percent reported several years earlier by Chaudhury et al. (2006). Moreover, health centers have on average only 52.3 percent of all essential drugs, 39.3 percent of essential maternal drugs, and 42.3 percent of essential pediatric drugs in stock. Results for vaccine levels are more favorable, with an average of 76.3 percent of all essential vaccines in stock at public health centers. With regard to equipment availability, only 39.0 percent of public health centers meet minimum requirements (World Bank 2016).

Notably, Dizon-Ross, Dupas, and Robinson (2017) present evidence that counters an extended series of studies documenting low health service quality in countries such as Uganda. Focusing on programs that distribute free malaria bednets, the authors find relatively high performance in the delivery of bednets by public sector health workers. However, as the authors acknowledge, the unique features of bednet programs — as publicly visible campaigns distributing a simple, preventative product for which willingness-to-pay tends to be low — may explain the low levels of corruption observed. Such features may not generalize to the broad set of health facility practices and outcomes that this study examines. For example, while the lack of a high-turnover retail market largely precludes diversion of bednets, diversion of inputs such as drugs and supplies is one of the key outcomes of interest in the setting of the present study.

Interestingly, while Dizon-Ross, Dupas, and Robinson (2017) report health worker absenteeism rates in a similar range as those cited above, they present additional evidence showing that health worker performance appears to be high for tasks other than bednet distribution as well. However, the data employed here are restricted to a limited number of antenatal care indicators. Other process quality measures for antenatal care, such as blood pressure measurement and deworming, remain low in the Ankole Region (UBOS and ICF 2018). More generally, while antenatal care quality has been the emphasis of many healthcare improvement programs, other aspects of healthcare — such as gestational age at first antenatal care visit, postnatal care utilization, and vaccination rates — continue to show substantial shortfalls, as described in the following sub-section.

E.2 Healthcare Delivery Gaps, Potential Health Impact Pathways, and Health Outcomes of Interest

This sub-section establishes a more detailed health outcomes framework based on the key healthcare delivery gaps that the interventions are most likely to address. Extending from the substandard health facility performance indicators detailed in the preceding sub-section, Uganda continues to fall short in several aspects of health intervention delivery to patients. For instance, even though antenatal care delivery has improved substantially in recent years, women in Uganda present for their first antenatal care visit at a median gestational age of 4.7 months — implying several months of missed opportunities for health prevention measures. In addition, well over half of women and newborns in the Ankole Region of Uganda go without a single postnatal care visit. Vaccination rates remain relatively low,

with 61.8 percent of children age 12-23 months having all basic vaccinations and only 34.4 percent having all age-appropriate vaccinations (UBOS and ICF 2018). Importantly, these low levels of intervention adoption and delivery reflect not only supply-side service gaps but also perceived low returns to seeking healthcare, which likely arise at least in part from poor service delivery itself. Along similar lines, recent public health literature has placed a strong emphasis on the role of respectful care in influencing utilization of health services (Kyaddondo et al. 2017; Sharma et al. 2015; Abuya et al. 2015)

Following these observations, the interventions in this study have the potential to affect outcomes on both the supply and demand sides. On the supply side, the programs may address two main dimensions of health facility performance: the intensive margin and the extensive margin. The intensive margin includes areas such as health worker absenteeism, drug and supply diversion, facility condition and sanitation, and respectful care, all of which local leaders can help to address. The extensive margin of facility performance includes the total capacity of local health centers, such as available infrastructure and staffing. If, as a result of either program, local politicians can successfully report and advocate upwards for increased resource allocation to their communities, the interventions can plausibly lead to extensive margin changes such as staff increases and infrastructural improvements.[17] More specifically, through the training program, political leaders gain skills to address the full breadth of potential supply-side problems at health facilities. In the meetings program, citizen complaints to politicians tend to relate more to the intensive margin of facility performance but can also relate to the extensive margin.

On the demand side, service delivery improvements due to either program may lead to increased healthcare utilization, helping to drive increased health intervention adoption and delivery. Because the meetings intervention in particular fosters the civic engagement of citizens to improve healthcare in their communities, community members in this arm may be especially likely to exhibit higher levels of healthcare utilization.

Based on the foregoing possible pathways, the study's health services outcomes fall into one of four categories, as highlighted in Section 4.3.1: (1) healthcare utilization, (2) healthcare quality, (3) intervention delivery and adoption, and (4) patient satisfaction. The healthcare utilization category includes measures such as number of visits to the local health center in the past 12 months, visits to other health providers, antenatal care completion by pregnant women, and postnatal care attendance and completion. The healthcare quality category includes measures such as drug and supply availability, health worker absence rate, wait time, use of equipment, facility condition and sanitation, and size of staff. The intervention delivery and adoption category includes childhood immunizations and treatment of common childhood illnesses. Finally, the patient satisfaction category includes measures of patient experience at recent health visits (e.g., respect of health workers, understanding of instructions) and overall satisfaction with the health center services. At the impact level, beyond these intermediary health outcomes, I examine child mortality, weight-for-age, and mid-upper arm circumference.

[17] Should these changes be realized, however, they are likely to occur over an extended period of time and therefore fall within the scope of a longer-term study.