Pre-Analysis Plan

An Impact Assessment of Stakeholder Engagement Interventions in Ugandan Oil Extractives

Eric A. Coleman
Associate Professor
Florida State University
ecoleman@fsu.edu

Jacob Manyindo
Coordinator at Maendeleo ya Jamii
JacobManyindo@maendeleoyajamii.org

A. Rani Parker, Ph.D.
President at Business-Community Synergies
RParker@BCSynergies.com

1

1   Introduction

1.1   Abstract

This study examines whether multi-stakeholder engagement improves transparency and accountability within the oil and gas sector of the Albertine Graben region in western Uganda. The study design is a randomized intervention of multi-stakeholder forums (with participants from communities, government and oil companies) among 108 communities in the region. The intervention is directly relevant to the national, regional and international policies on transparency and accountability in the extractive industries. The study also provides rigorous experimental evidence to the large academic literature on corporate social responsibility.

1.2   Motivation

1.2.1 Background

Uganda's Albertine Graben generally comprises large expanses of traditional, subsistence agricultural lands, forest reserves, and prolific wildlife resources all contributing to the region's geographic diversity, ecological importance and socio-economic opportunities. The broad ethnic diversity of the area also adds to the uniqueness of the area. Recent discoveries of oil and gas have propelled this area into the national spotlight with high expectations that it will lead Uganda's economic growth over the coming decades. However, communities associated with the petroleum exploration and early production phases have a number of concerns. These include restricted access to information, limited opportunities to communicate and interact with key public and private sector decision-makers, displacement and inadequate or untimely compensation, environmental pollution, elite capture of communal resources and new economic opportunities, and the decline in cultural and moral standards associated with itinerant labor, amongst others. Communities in the Albertine Graben have very little information on government and oil company plans for the construction phase of Uganda's oil and gas industry cycle and how those plans will affect them. This has led to unrealistic expectations of benefits and misinformed concerns. In the absence of transparency from government and the companies, communities are unable to demand accountability from either.

The Government of Uganda and the three licensed oil companies are preparing to start the production phase that will involve construction of central processing facilities, a refinery, and pipelines through the region. While there has been some effort to improve the availability of information to these communities, active stakeholder engagement on the impacts of the construction phase is virtually non-existent.

1.2.2 Random Assignment of Intervention

In a small section of the Albertine Graben, a local civil society organization, Maendeleo ya Jamii (MYJ,) has been supplementing information provided by the companies themselves by creating opportunities for stakeholder engagement between oil companies, government actors, and local communities. Their work involves community training (about how to set priority areas and interact with decision-makers) and then facilitating multi-stakeholder forums wherein communities advocate on their own behalf. This project evaluates the effectiveness of this stakeholder engagement to address the community concerns.

MYJ is now expanding their work throughout the Albertine Graben region. They have agreed to target

new villages via random assignment for this study. MYJ has identified 108 new villages in the area– half of these villages will be randomly assigned to treatment group and half to a control group. The intervention in these new villages will take place between the baseline and endline data collection in fall 2016.

In both treatment and control villages, MYJ will hold workshops explaining to community members the key junctures in the oil company planning cycle, informing them about their rights and how to exercise them, encouraging them to formulate discussion priorities when engaging with oil companies and government, and helping them understand reasonable expectations from this process. The implementers will develop a standard information package delivered during workshops in all villages. We expose both treatment and control groups to information because there is some information dissemination already happening in the region and we want to ensure that both groups have access to the same information.

Only villages assigned to the treatment groups will participate in cooperative planning through stakeholder engagement processes. MYJ sets up a Multi-Stakeholder Forum that facilitates interaction between communities, government, and oil companies wherein all parties discuss their concerns, coordinate their plans, and commit to actions.

1.2.2 Contributions

This project is funded through the International Initiative for Impact Evaluation (3ie) through the Thematic Window 8, Transparency and Accountability (T/A) in initiatives in governance of extractives. The project fits within a burgeoning literature using experimental methods to evaluate T/A reforms (see Kosack and Fung 2014). Most research to date examines how transparency interventions targeting government agents (e.g. civil servants) improves service delivery. This project extends this literature to examine T/A reforms at the nexus of complex private, public, civil society interactions.

The project also touches on a more mature literature in Corporate Social Responsibility (CSR). Many different international organizations insist that companies in the extractive sector engage with local stakeholders and issue "best practices" for how to do so: among others, the International Finance Corporation (IFC) issues a handbook for stakeholder engagement, the United Nations has developed Guiding Principles on Business and Human Rights, and the European Commission has issued a guide on how to implement the UN Guiding Principles in the oil and gas sector. While some oil companies have evaluated the efficacy of these programs internally, there is little public knowledge about the impact of local stakeholder engagement. Furthermore, we do not know of any rigorous experimental investigation— public or private—about the effects of such programs. The principles of stakeholder engagement espoused by MYJ closely match the best practices found in these guidelines, and are particularly close to those found in the IFC handbook.

1.3   Research Questions

Broadly, this study evaluates the effectiveness of a transparency intervention to improve accountability in the oil and gas sector. The implementers provide information and facilitate direct stakeholder engagement among communities, oil companies, and government officials. We hypothesize that such an intervention improves accountability to local communities. Conceptually, we conceive accountability as what Behn (2001) calls "performance accountability," where we emphasize results rather than rules. Stated

differently, then, our main research question is to examine whether stakeholder engagement improves outcomes from the point of view of local communities.

We also have a number of more specific research questions, including:
- How are communities impacted by participation in Multi-Stakeholder Forums?
- Does stakeholder engagement improve political participation of communities? At local (district and below) levels? At regional levels? At national level?
- Does the capacity for community level collective action improve?
- What types of action do we see in intervention/treatment communities that we do not see in others? Who is leading these actions? Are there unique circumstances surrounding these actions?
- What are the unexpected impacts from multi-stakeholder engagements?

1.4 Impact Theory

The intervention occurs in the form of access to information (both treatment and control group) and stakeholder engagement (treatment group only). Conceptually, the control communities thus represent a transparency-only group, while the treatment communities represent an accountability + stakeholder engagement group. Thus, the design is meant to assess the value-added of stakeholder engagement above transparency.

But why should stakeholder engagement improve performance at all? One explanation is that decision-makers respond to "naming and shaming." However, Fox (2007) notes that shame may not motivate change in the shameless. A far more compelling reason comes from research on CSR. Companies in the extractive sector often attempt to engage with local stakeholders not out of a sense of altruism or even shame, but because conflict with and resistance from communities may delay activity and decrease the the profitability of the venture (Franks et al. 2014). In a recent publication in *PNAS* Franks et al. (2014) report that most companies believe that relations with local communities can be amicable by using best practices of stakeholder engagement. This argument justifies stakeholder engagement from the perspective of the oil companies; however, we have much less evidence about whether these efforts simply "greenwash" irresponsible corporate behavior or whether they improve the plight of affected communities.

We have identified three reasons why stakeholder engagement can improve performance from the point of view of the communities. Each of these explanations explain how communities respond to stakeholder engagement by channeling their efforts in ways that decision-makers are unlikely to ignore. If poor relations can compromise the profitability of a venture, decision-makers are likely to respond to communities to which they believe they should respond in order to cultivate beneficial relations.

First, stakeholder engagement clarifies the roles, responsibilities, and duties of different decision-makers. A necessary condition for accountability is that citizens attribute blame or credit to the actors responsible for the state of their concerns (Manin, Przeworski, and Stokes 1999; Gomez and Wilson 2006, 2008). For complex policy problems such attributions are difficult to make. The process of stakeholder engagement simplifies this process: if stakeholders make commitments to perform certain actions, then others can observe whether they executed those actions and seek remediation from the appropriate actor if those actions are not followed. With greater clarity of responsibility, communities are more able and willing to hold the appropriate decision-makers accountable for their performance. If decision-makers anticipate that

people can more clearly identify poor performance, then they have greater incentive to improve.

Second, the treatment improves the capacity of the communities to act collectively. The intervention helps communities work together to both identify and then advocate for their priorities, learn about the actions of other communities, and engage with policymakers within the oil and gas sector. Each of these factors (prioritization, advocacy, learning from others, and practice) have been linked to greater capacity for collective action (Ostrom 1990). Collective action capacity, on the other hand, may improve the performance of actors within the oil and gas sector if they either anticipate or experience resistance to poor performance.

Third, stakeholder engagement increases community political activity in order influence decisions or to seek remediation. Stakeholder engagement helps communities identify the actors responsible for the various aspects surrounding petroleum and then learn from other communities' attempts to influence this process. For example, MYJ has previously worked with a community to write a letter of petition to a government official about a problem they were experiencing. This action successfully enlisted a government official to advocate upon the communities behalf. Other communities have met personally with government officials or with oil company community liaison officers. MYJ believes that stakeholder engagement activities have increased community awareness of the potential for policy influence and remediation as well as clarify the appropriate political actions to take for specific concerns. This differs from collective action in that political behavior may be either individual or collective. Even non-collective political activity may motivate decision-makers to improve performance. Performance improves if decision-makers either experience or anticipate such political activity.

2 Research Strategy
2.1 Sampling
2.1.1 Sampling Frame

The sampling strategy for the study started by identifying all communities in the area that are affected by oil operations and with populations between 200 and 2,000 according to the most recent census. We sampled from the communities in each of the following districts of western Uganda: Hoima, Buliisa, Moyo, Yumbe, Arua, Nebbi, Nwoya Ntoroko.[1] Thus, the sampling frame is all communities within those regions with a population within the constrained bounds.

To address the spillover problem, we randomly sampled 108 communities within these districts under the constraint that all sampled communities be at least 4 kilometers away from any other sampled community. Figure 1 shows all the communities identified by the population constraint (blue dots) as well as the final communities that were randomly selected for the study under the constraint that there is at least 4 kilometers between each community (red crosses). To get a better idea of how close the communities are and the effects of sampling a single community among those that are clustered, Figure 2 shows a zoomed in map where each study community (red cross) is at least 4000 meters away from other sampled communities.

---

[1] This is a slight deviation from our original proposal to 3ie where we limited the study to Hoima and Buliisa. We expanded the study to more districts because we were concerned about spillovers between communities.
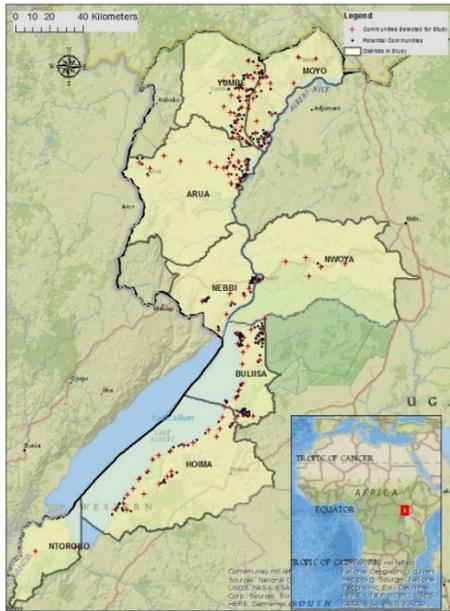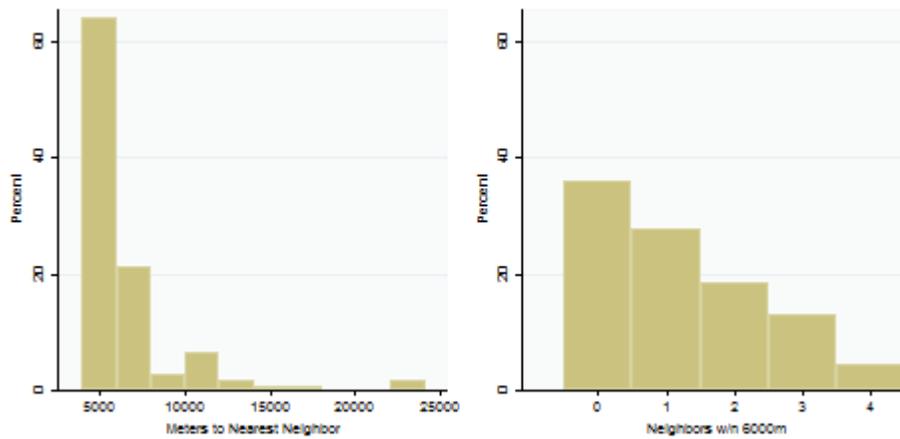
Figure 1. Sampled communities throughout the region.
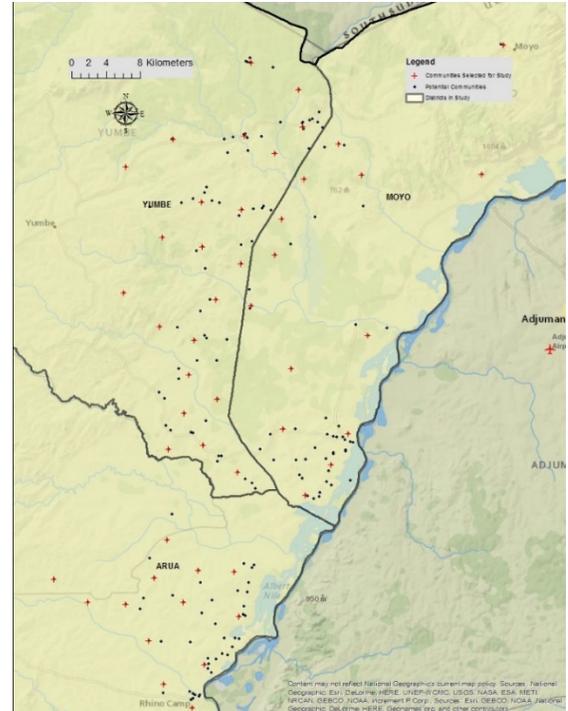
Figure 2. Sampled communities within the North.

The left panel of Figure 3 shows a histogram of the nearest neighbor (in meters) for each community in the study and the right panel shows a histogram of the number of communities within a 6000 meter radius of each community in the study.

Figure 3. Histogram of distance of nearest neighboring communites for each study community.

Within each community we use a quasi-random sampling technique (random walk) to survey 30 households. In total we have 108 communities within which the household survey is clustered. Therefore, the total household sample size is 3240 (=108 x 30), clustered within the 108 communities.

Random sampling under the minimum distance constraint ensures that those communities selected into the study are not systematically different than other communities of the minimum distance. While the sampled communities may be more isolated than a typical community, we do not believe this poses a serious bias in the design and has the important advantage of mitigating spillover effects. The intervention may be more effective in communities with populations in the required range. However, MYJ, the implementing partner, works in these types of communities.
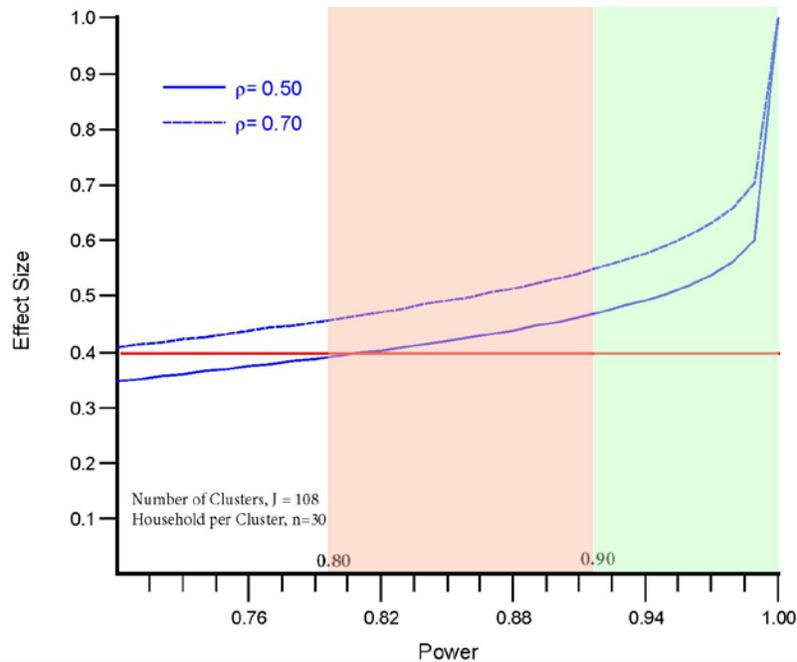
2.1.2 Statistical Power

There are no direct studies from which to benchmark an effect size. However, we have identified a few broadly relevant studies. First, Daley (2007) uses a Heckman treatment selection model to estimation the effect of stakeholder engagement in the U.S. Environmental Protection Agency's response for remedial action in Superfund sites. After controlling for selection of sites where community groups participate, she finds that community group participation increases the stringency in EPA decisions by 1.2 ($p<0.01$) standard deviations. Second, Bjorkman and Svensson (2009) report on an experimental intervention of community-based transparency and monitoring intervention for health services in Uganda. They report a number of treatment effects and a few analyses of difference-in-differences (DID) estimates at the village level that we use here. They find an average standardized effect of about 2 ($p<0.05$) for the Difference-In-Differences utilization of health services. That study looks at a number of outcomes, and some of their outcomes have a smaller effect size, but most are above 0.5 standard deviations.

Statistical power calculations are based on the quantitative component of the study and use a level of significance of $\alpha=0.05$. We sample 30 households per community to receive adequate coverage within each of the 108 communities. We test the sensitivity of the power calculations by calculating power based upon two values of the interclass correlation coefficient, ρ, which represents the proportion of variation in the outcome between communities. We calculate power assuming values of $\rho=0.5$ and $\rho=0.7$. Both estimates are conservative parameters of the power calculation, and they imply that 50 and 70 percent of the variation in total giving in the matched donation game is accounted for between communities rather than within communities.

We use the software Optimal Design Plus for the power calculations, a useful application for cluster randomization.[2] Figure 4 reports the results. Even when using a very conservative estimate of the interclass correlation coefficient ($\rho=0.7$) the design has a power of about 0.9 to detect an effect of 0.6. Using a more reasonable assumption of $\rho=0.5$, the design has a power of above 0.95. For effect sizes above 0.6 the design has sufficient power.

---

[2] See http://hlmsoft.net/od/

Figure 4. Relationship between statistical Power (x-axis) and Minimum Detectable Effect Size (y-axis) give the parameter assumptions discussed, with 108 communities and 30 household surveys per community. Pink shaded area represents power above 0.8 and the green shaded area represents power above 0.9. The solid red line represents a minimum effect size of 0.4, the lowest effect size in the studies reviewed. The solid blue curve shows the relationship between power and effect size assuming the interclass correlation coefficient is 0.5, while the dashed blue line shows the relationship assuming the interclass correlation coefficient is 0.7.

2.1.3   Assignment to Treatment

We randomly assign treatment at the community level. We use block randomization and will ensure balance of pre-treatment covariates between treatment and control communities within each of the eight districts of the study.

2.1.4   Attrition from the Sample

We do not expect attrition at the community level. MYJ has been working in this area for some time and has not had an incidence where a community did not want them to return.

2.2   Fieldwork
2.2.1   Instruments

The data collection has both quantitative and qualitative components. The enumerators will gather household data using a smart phone app and upload them to a secure server. These surveys were collaboratively developed among participants at a workshop including staff from MYJ, Prof. Coleman, and Dr. Parker. The questions on the survey, by and large, have been used before, although some questions are measured in a novel way. For example, questions measuring performance accountability are adopted from participatory ranking methods.

8

Qualitative data collection extends and complements the data gathered in the quantitative survey. In particular, the qualitative data is geared to understand the mechanisms about why the intervention did or did not work in some communities. We also hope to gain a richer understanding of community concerns not possible from the quantitative data. To this end, we have collaboratively developed a series of open ended questions and a protocol for identifying subgroups of people, specifically groups of leaders separately from groups of women, to discuss these issues. We used the methods developed over the extensive experience from Dr. Parker.

2.2.2   Data Collection

We will collect both pre- and post-intervention data. We anticipate that each phase of data collection, among the 108 sites, will take approximately 2 months. The post-intervention data collection happens approximately one year after the pre-intervention data.

The quantitative data collection entails each household completing the household survey with each enumerator. We anticipate this will take approximately one hour. Data is entered into a smartphone app at the field site and then uploaded to a secure server in the evening that is password protected and where only the PIs have access. Enumerators only have access to the limited data they themselves gather during a given day and this data is then erased after data validation following each wave of the survey.

2.2.3   Data Processing

Data processing takes place after each wave for both the quantitative and qualitative data. We anticipate that data processing will take approximately 1 month after each wave. Data processing entails cleaning and clarifying data where discrepancies exist. Again, data is stored on a secure, password protected server that only the PIs can access. The PIs jointly own the data. After the study is complete, the data will be removed from a shared server and each PI has the responsibility to ensure their copy of the data is secured long term.

3   Empirical Analysis
3.1   Variables

MYJ's primary concern is that companies and government are accountable to local people. Accountability, however, is difficult to measure directly for a number of reasons. First, different households have difference preferences about issues. Second, different communities have different preferences about issues. Third, the multitude of issues, and the different importance that households and communities place on each issue makes scaling the issues makes comparisons between treatment and control villages difficult. How does one compare a school built by an oil company in one village to a five oil-sector jobs in another village?

We use a ranking and weighting exercise to measure seven types of concerns to each household: employment, environment, land, social services, moral standards, local economic development, and information. First, we ask the household to rank the importance of each issue by laying out 3 x 5 index cards with each issue written on each card. Then, we give each household 20 small stones and ask them to place the stones on the cards in proportion to how important each issue is. (See Figure 5 below)

Read: "We would now like you to take a minute and think about the issues that are most important for you in relation to the oil projects in the region. We have identified seven issues that often come up when we talk with people. We would like you to rank how important you think each of these issues is. To do so, we will give you twenty rocks to place on these cards—each card represents one of the seven issues. Place more rocks on cards that represent the issues you believe are more important. Issues that are not important receive few or no rocks."

IEMPLOY	IENVIRO	ILAND	ISOCSERV
Having members in my household employed by the oil sector.	Maintaining and improving the local environment (e.g. water quality, waste disposal).	Managing land rights (conflict, compensation, registration, grabbing).	Receiving social services (health clinics, schools).

IMORAL	IECONDEV	IINFO	IEMPLOY IENVIRO ILAND ISOCSERV IMORAL IECONDEV IINFO
Maintaining my community's moral and cultural standard.	Improving my household's well-being from local economic development.	Improving my information about the oil activities, management, and opportunities.	

Figure 5. Survey script for issue importance

We then ask each respondent their satisfaction with how each issues has been handled. They now allocate up to 10 rocks on each card representing how well they think the issue has been handled (with more rocks representing better handling of the issue). We then form an index of *Performance Index* calculated as follows:

$$Peformance\ Index = \sum_{k=1}^{7} w_k z_k \tag{1}$$

Here, k indexes an issue, w_k, are the importance weights (the number of rocks on an issue divided by twenty, such that $\sum_{k=1}^{7} w_k = 1$), and $z_k \in [0,10]$ is the ratification with how each issue is handled. This measure has the advantage of being flexible enough to allow different issue areas to have priority for different households and communities. It also provides a scalable way to weight the satisfaction with each issue based upon their importance. The measure does has the disadvantage of being subjective, in that *Performance* depends upon each household's perception of how the issues are handled rather than on the state of the actual issues—the actual number of jobs, the level of economic development, or service provision. However, we feel this approach is justified both because of the flexibility of the measure and because the subjective perceptions of households determines whether the decision-makers are held accountable.

In addition to the *Performance Index* we also measure the intermediate outcome variables mentioned in our Impact Theory (Section 1.4). We briefly mention each of these below.
1. *Causal Attribution*. Index of how the household attributes responsibility for their concerns among stakeholders. This measure somewhat mirrors the *Performance Index*. We have the survey respondent engage in a similar task assigning credit or blame to each of seven actors who might

influence the performance issues. They then place rocks on index cards to indicate how much each actor bears responsibility for this performance.
2. *Collective Action Capacity.* We have each surveyed household anonymously play an incentivized public goods game with the other surveyed households in the village.
3. *Political Activism.* Index of political activities the household engages in to hold stakeholders accountable. We ask each household to indicate whether they have engaged in a variety of political activities in relation to issues in the oil and gas sector

In addition to these outcomes, we will gather data to measure transparency (objective knowledge questions around the oil and gas development), demographic, socioeconomic, and background variables.

3.2   Balancing Checks

We will assess balance of pre-intervention variables between treatment and control communities from the quantitative data after the first wave of household surveys. We will use standard procedures to check balance including difference in means (proportions) as well as examine the distribution of variables. We will try and balance on all pre-treatment variables.

3.3   Treatment Effects

3.3.1   Intent to Treat

We do not anticipate much, if any, non-take-up of the intervention. Again, the collective experience of the research teams suggests that communities are anxious to participate in the multi-stakeholder forums. However, some communities may decline to participate or simply not show up to the forums. At each forum, we ensure that only people invited from the community participate. This way, we can keep track of which communities have participated and which have not. If there is significant non-participation, then we can back out an intent-to-treat effect. We will include the pre-intervention outcome variables in this estimation and some additional covariates. Specifically, we propose to estimate the following equation:

$$y_{ij,post} = \alpha_j + \tau T_j + \theta y_{ij,pre} + \beta x_{ij,pre} + \varepsilon_{ij} \qquad (1)$$

Where *i* indexes individual, *j* indexes community, and *pre/post* index the time of measurement. The variables *y* denotes each outcome outlined above, *T* denotes whether the community was assigned to the treatment group, and *x* denotes other covariates (specifically gender and socioeconomic status). We directly interpret the parameter τ as the intent-to-treat effect.

3.3.2   Spillover Effects

We have adjusted our design to incorporate the largest possible buffers between communities. However, we remain concerned about spillover effects between communities. Note that the treatment itself cannot be obtained by communities assigned to the control group: they are only invited to participate in stakeholder engagement at the conclusion of the endline survey. Thus, they do not have access to policymakers inherent in these forums.

However, control communities may interact with people from treatment communities and learn about different actors and strategies that might improve their outcomes. Further, positive outcomes in treatment community may provide external benefits to control communities. For example, if a school is built in a treatment community it is possible that a control community may send their children to that school.

We have dealt with these concerns within the design in two important way. First, we proposed the largest possible buffer to attain our sample size with at least 4 kilometers between all communities. Second, we collect both baseline and endline data and estimated treatment effects while controlling for a pre-treatment values of each outcome. Thus, we can assess whether outcome improve for both treatment and control communities or just treatment communities.

Other than these design features, we will also model spillover effects statistically. We do this in two ways. First, we propose geographic distance as a proxy for spillovers. This approach seems adequate to capture the external benefits received by control communities. We also ask a series of social network questions (about centers of economic exchange, familial connections, etc.) to control for information dissemination that may happen between communities.

3.4   Heterogeneous Effects

We are somewhat concerned about the distributional effects of the intervention for people within the study communities. In particular, we are concerned about whether women and those of low socioeconomic status benefit as much from the intervention as men and those of relatively higher socioeconomic status. Since communities participate in these forums as a whole, it is possible that the concerns of more marginalized groups may not be represented. While we have tried to address this in the broader design of the study (at least one women must represent each community in the forums, and we randomize whether we speak with a man or a woman before each survey) we still remain somewhat concerned. However, we have not explicitly built examining heterogeneous effects into the study.

3.5   Standard Error Adjustments

We estimate Equation 2 through a hierarchal modeling approach, where individuals are clustered within communities. This mirrors our power calculations.

Many of our outcomes use indexes. The weighting exercises imply index weights for the *Performance Index* and the *Causal Attribution Index.* We will construct a simple additive indexes for *Political Activism*. In addition, we have specifically designed the questions in our indexes to provide variation between respondents. For example, when measuring transparency we ask each respondent 10 True/False questions with verifiable definitive answers that are either right or wrong. Some of these questions are easy and some are a bit more difficult. In this way we hope to increase variation and avoid ceiling/floor effects. However, after collecting the pre-intervention data we will have a better sense of this and can adjust accordingly if necessary.

4   Research Team

The principle investigators of the study are Jacob Manyindo, Coordinator at Maendeleo ya Jamii; A. Rani Parker, Ph.D., President at Business-Community Synergies, and; Eric A. Coleman, Associate Professor of Political Science at Florida State University. Each of the PIs has invested significant time and energy designing the study, and each will participate in data analysis, writing reports, writing scientific articles, and disseminating the results to key stakeholders. Mr. Manyindo leads the day-to-day operations in Uganda for the intervention and oversees data collection efforts. He also leads the stakeholder outreach. Dr. Parker is the lead PI and oversees the budget, leads the qualitative data collection and analysis, and leads policy dissemination activities. Prof. Coleman leads the quantitative data collection and analysis.

Dr. Coleman also has a research assistant who assists with developing the quantitative data collection platform, data processing, and data analysis.