

# Reducing Corruption and Improving Institutional Performance with ICT and Digital Media in Developing Democracies

James D. Long  
University of Washington  
jdlong@u.washington.edu

Danielle F. Jung  
Emory University  
danielle.jung@emory.edu

Aaron Erlich  
McGill University  
aaron.erlich@mcgill.ca

October 2, 2017

## Abstract

In response to corruption at the central level, policymakers have increasingly advocated for the creation and strengthening of devolved government structures as a policy tool to confront many challenges linking institutional performance to public services and citizen welfare in developing countries. Devolution reassigns many government functions (including taxation, licensing, and service delivery) from the national level to autonomous local agencies in the hope that they improve institutional economies of scale that work to enhance information, transparency, and accountability to build a more responsive face of government for local residents. Our study employs and evaluates methods to improve the performance of newly devolved government institutions (county councils) in Kenya. First empanelled in 2013, anecdotal evidence suggests wide variation in the degree to which county councils in Kenya perform their duties, generating differences in county operations and the quality of output. We will leverage a scalable and replicable ICT platform on mobile phones and digital media to measure and improve agency functioning and related outcomes. The platform's functionality will send aggregate information from citizens to (elected and non-elected) county officials, as well as reports back to citizens about the performance of their local counties, allowing bureaucratic managers and citizens to audit, monitor, and report on government outputs. The platform will allow randomized treatments on reducing transaction costs, improving monitoring, and heightening accountability through community monitoring, and provide real-time feedback to institutional actors to action responses.

# 1 Corruption and Public Sector Performance in the Developing World

Corruption and poor bureaucratic performance have pernicious effects within and across levels of government in developing countries (Finan, Olken and Pande forthcoming). Accordingly, corruption negatively correlates with economic growth (Mauro 1995); indicators of democracy and the rule of law (Mungiu-Pippidi 2015); and service delivery outcomes in sectors including security (Condra et al. 2016), education (Duflo, Hanna and Ryan 2012), and health (Chaudhury et al. 2006). As a result, “good governance” now forms a core goal of development assistance, highlighted by the World Bank devoting its 2017 World Development Report to the study of public sector governance. Graft is also increasingly a salient political issue shaping citizen mobilization in transitioning societies. Government malfeasance sparked waves of protests against undemocratic regimes during the Arab Spring and Color Revolutions (Breuer, Landman and Farquhar 2015, Tufekci and Wilson 2012), mirroring ongoing anti-government demonstrations in Brazil, Venezuela, Russia, and South Africa. Calls for many of these mass actions have been publicized via text messages and social media.

Despite a statistical association between the quality of governance with economic indicators at the national level (Easterly and Levine 1997) and within country ethnographic accounts confirming this pattern (Gupta 2012), academics, policymakers, and citizens lack important knowledge about numerous dynamics regarding the micro-foundational causes, processes, and outcomes related to corruption and institutional governance in developing democracies.<sup>1</sup> Prior studies note that corruption may arise from the behavioral characteristics and intrinsic motivations of individual bureaucrats (Callen et al. 2015, Dal Bó, Finan and Rossi 2013), as well as institutional pathologies arising from agency management (Fisman and Golden forthcoming). Attempts to mitigate shirking therefore require supervisors to provide bureaucrats the right mix of external incentives with increased monitoring and sanctioning from within government institutions (Becker and Stigler 1974, Shleifer and Vishny 1994).<sup>2</sup> Problematically in developing countries, managers and political leaders frequently lack the information, incentives, or ability to oversee agent behavior. Obtaining credible information about performance is hard because actors committing illicit acts hide them, supervisors are frequently trying to oversee agencies and offices from far away, and bosses may receive unreliable or conflicting (“jammed”) information about performance and outcomes from actors misrepresenting actions. Moreover, improving information and transparency does not automatically transpose to heightened accountability and ser-

---

<sup>1</sup>Standard political economy conceptions of corruption by civil servants define it as the use of public office for personal gain (Rose-Ackerman 1999). This definition includes the propensity for bureaucrats to take bribes, produce waste or inefficiencies in the performance of their duties, and/or failure to meet performance indicators.

<sup>2</sup>Along these lines, numerous policy evaluations in development economics use randomized control trials (RCTs) to treat individual bureaucrats with positive and/or negative incentives to improve public agencies (Finan, Olken and Pande forthcoming). But the field so far has not built a set of knowledge across institutions, sectors, and context about “what works” and why to improve government functioning (Dal Bó and Finan 2016). For example, while teacher or health worker absenteeism may form a critical problem in the education and health sectors (Chaudhury et al. 2006), bribe-taking may be a challenge with the police (Condra et al. 2016). How those sectors’ institutions are organized and where possible pressure points for agent change and reform reside (including methods needed to audit, monitor, and sanction) could be different.

vice delivery (Dal Bó and Finan 2016). As a result, reform attempts to measure and combat corruption has not always proven successful.

To confront these challenges of poor government performance, policymakers have advocated for the creation and strengthening of devolved government structures as a tool to promote better institutional governance and reform of public agencies in developing democracies. Devolution transfers many government functions (including financing, decision-making, and management) from the national (centralized) level to local autonomous government bodies in the hope that achieving institutional economies of scale works to enhance information, transparency, and accountability. Devolution also leverages citizen participation and consultation in development planning and project oversight to build a more responsive government presence (Björkman and Svensson 2009).

Kenya's recent constitutional reform implements devolved government with the creation of 47 county councils, first empaneled in 2013 and replacing a previous system of (poorly financed and functioning) municipal councils. By reallocating many executive responsibilities to counties, devolved government in its design empowers counties with new and important roles in revenue collection, licensing and contracting, and service provision across 11 sectors (e.g., health, transportation, markets, education, water, and housing). Counties consist of a legislative assembly (with individual members elected from single-member wards) led by a popularly elected governor, who forms the executive. Counties also contain a non-elected executive secretariat (County Service Board), with members hired and managed by the governor. Both elected and non-elected members of the county government are responsible for functions and outcomes collectively at the county level, with councilors and secretariat members responsible for operations within their wards. Importantly, these elected councilors and non-elected secretariat members have overlapping responsibilities and mirror each others work. Counties receive only partial (and minimal) financing from the central government, but otherwise they are autonomous units who raise and spend funds locally. Counties also support national ministries locally, and by law must involve citizen participation in budgeting, planning, and oversight of county functions and services.

While devolved government should provide new avenues for counties to increase the level and quality of services they provide, after five years Kenyan councils demonstrate wide variation within and across counties in performance and outcomes (Burbidge 2015). Anecdotal evidence shows that the roles and responsibilities for institutional actors remains poorly understood; managers lack robust information and monitoring systems to oversee county functions; and citizens lack information on devolved structures and have not consistently participated in county functions. Although they are now the face of government to most Kenyans and councilors face their first elections as incumbents in August 2017, it is unclear whether devolution has succeeded in improving governance and services in Kenya (Cheeseman, Lynch and Willis 2016).

To evaluate county functions and to work towards improving their performance, our project pursues the following lines of investigation: i) have counties achieved their intended effects? ii) do leaders, county managers, and citizens obtain the information and capacity to monitor the workings and outcomes of councils? iii) can citizen feedback promote political accountability and improve service delivery? iv) do randomized treatments reducing transaction costs through improved information and monitoring technology work to improve outcomes? v) how does top-down monitoring from county

managers work compared to community monitoring? vi) have counties created new sites for corruption and inefficiency? and vii) what heterogeneous characteristics of counties, county sub-units, and institutional actors predict variation in performance and responsiveness to interventions, including with respect to elected versus non-elected members?

To address these questions, our project leverages partnerships with Kenyan researchers and county councils to evaluate and improve components of Kenya’s devolved county government.<sup>3</sup> We will utilize randomized treatments within and across sub-county administrative units (across a selection of counties), using a new multi-channel ICT/digital media platform on mobile phones to better provide leaders, county managers, and citizens information and monitoring technology on the administration and outcomes of county performance. The platform will incorporate information solicited from citizens, managers, and independent auditors to engage stakeholder users within the county government and communities. The platform’s functionality will reduce the transaction costs of gaining information on bureaucratic performance for county officials, further allow county managers to oversee and monitor bureaucrats’ actions, and incorporate community monitoring across service delivery and other outcomes related to county operations.

## 2 Theoretical Approach

A number of important factors related to governance in developing democracies result in corruption, poor service delivery, low government responsiveness, and a lack of citizen engagement. Such societies frequently lack the institutional strength necessary to fulfill the Weberian tenets of effectively managing public institutions (Levi 1988), or those institutions and actors within them lack capacity (Besley and Persson 2014). In transitional democracies with weak institutions, politicians’ comparative advantage is to compete for office based on clientelism (Keefer and Vlaicu 2008). Public managers are therefore likely to use hiring for positions in bureaucracies to support those for whom they and their allies are in patronage relationships (Chabal and Daloz 1999, Chandra 2004). Clientelist systems tend to favor certain regions, groups, or communities at the expense of others, leading to policy-making and outputs that are not inclusive. Reform attempts to reduce graft and efficiency are like to be ineffective when spoilers have incentives to block them since those who object to better policies stand to lose rents and thus political power. Frequently, the best strategy to avoid conflict and violence involves distributing rents through powerful brokers, such that graft and inefficiency is the price for stability (North, Wallis and Weingast 2009). Given these governance factors resulting from the distribution of power among political actors, expecting a reduction in corruption and improvements in services from endogenous institutional change on its own is unlikely to promote reforms in developing countries’ public sector; in these societies, oversight and accountability may need to arise from elsewhere.<sup>4</sup>

---

<sup>3</sup>We have ongoing policy dialogue with local implementing partners who are managing relationships, buy-in, and input from county governments. Given space constraints, we omit an extended discussion of these relationships here, but are happy to discuss further at the workshop.

<sup>4</sup>This insight aligns with an established literature focusing on the positive relationship between inclusive institutions with restraints and long-run economic performance, compared to countries with weak and closed institutions subject to elite capture and poor economic growth (Acemoglu, Johnson and Robinson 2001,

Our approach aligns with political economy insights that monitoring and sanctioning capacity proves essential to study and regulate bureaucratic processes. However, we depart from extant studies that focus exclusively either at the individual bureaucrat or agency (institutional) level. Instead, we harness the organizational design of counties and locate embedded and external sites of contact between principals and agents to study the interaction between individuals and institutional dynamics and how variation in transaction costs, monitoring capacity, and accountability for institutional actors and the public (in)directly affects behavior change and outcomes for counties.

To understand the dynamics of public management in devolved structures and associated processes and outcomes, we advance an analytic framework based on principal-agent theory (Kiewiet and McCubbins 1991). We identify two principal-agent relationships in the public sector that are critical to institutional governance: between managers and bureaucrats within agencies; and between the public and officials who staff and oversee those agencies. To reduce agents' propensity to shirk (e.g., absenteeism, poor output, corruption, wasting resources), both sets of principals require information on performance to provide monitoring capacity to detect and deter misbehavior. However, principals face numerous constraints to obtaining the quantity and quality of information required for robust oversight: agents have incentives to misrepresent themselves, hide behavior, or jam signals to cheat; managers and citizens are constrained by the time, budget, and other resources to monitor agents, or may lack incentives to supervise. Increasing transparency should make it less likely that bureaucrats can shirk and deviate from priorities, but managers and/or citizens may not always prove effective points of accountability within or outside of institutional structures.<sup>5</sup>

First, we argue that reducing the transaction costs of oversight and monitoring with improved reporting tools for managers to oversee bureaucrats should help those managers to reward good performance and enact sanctions in situations where managers themselves have incentives and capacity to influence agents. If transaction costs to monitoring are low, managers can improve their oversight of agent behavior and consequently, shirking should be reduced. Our first experimental treatment therefore reduces transaction costs to test

***Hypothesis 1: reducing transaction costs for managers to oversee bureaucratic behavior will improve county performance.***

However, reducing transaction costs realistically only improves agency outcomes when managers simply lack information; if those managers have competing incentives to shirk, better information on its own will not necessarily force them to improve monitoring. If managers do not find it in their interest to use information to monitor on their own, bureaucrats will likely to continue to shirk as well. Therefore, our second

---

Engerman and Sokoloff 1997, Fukuyama 2014, Robinson, Acemoglu and Johnson 2005). Societies grow more equal, democratic, and wealthy when citizens representing diverse social and economic groups actively participate in political life.

<sup>5</sup>Another intervention type to reduce corruption and improve bureaucratic functioning uses performance-based pay or performance incentives (Banerjee, Duflo and Glennerster 2008, Basinga et al. 2011). We do not believe performance pay forms a realistic or sustainable solution for devolved government in Kenya because a large injection of cash would be needed, and consistently so, which is unlikely. We believe our model is therefore more cost-effective, scalable, and sustainable in the absence of a large external budget to finance performance pay.

treatment directly imposes monitoring from managers to bureaucrats by announcing audits and detection of agent behavior to test

***Hypothesis 2: announcing monitoring by managers of bureaucrats to managers and bureaucrats will improve county performance.***

Additionally, because of manifold institutional pathologies, institutional actors may not be sensitive to increased monitoring if credible threats of sanctioning do not reside internally. Rather, (heightened) accountability could arise from outside threats of sanctioning. Because community members experience corruption and suffer poor performance acutely and locally with devolved structures, community monitoring may generate accountability beyond institutional accountability. Further, citizen participation in monitoring government performance is critical for reform and development within devolution, therefore the public has reasons to impose sanctions on county officials if they are perceived to have performed poorly. Our third treatment accordingly introduces community monitoring to test

***Hypothesis 3: increasing community monitoring will improve county performance.***

To the extent that it exists, accountability generated from community monitoring could result from two distinct mechanisms. First, citizen oversight may result in a general form of “social accountability” and the desire of agency employees to appear good in the eyes of their local communities and fellow citizens. Second, community monitoring may replicate “electoral accountability,” or the ability of citizens to “throw the rascals out” if they do not like how elected representatives (or other members in their offices) perform, in which case elected members would lose their jobs and associated benefits. While we do not have strong theoretical priors on whether community monitoring works through a social or electoral mechanism, we are able to test whether heterogeneous effects of citizen oversight strengthens or attenuates for elected council members versus non-elected bureaucrats.<sup>6</sup>

### 3 Institutional Setting

A critical component to the study of organizational governance requires careful attention to the institutional setting and unit of analysis. Traditionally, many quantitative and large-n cross-national studies of corruption examine country level outcomes (e.g., [Easterly and Levine 1997](#), [Mauro 1995](#), [Rose-Ackerman and Palifka 2016](#)). More recently, studies have focused on better identifying mechanisms and institutional pathologies of corruption, including the industrial organization or “culture” of certain bureaucracies or sectors. Identifying points of delegation, monitoring, and accountability within institutions is difficult, creating ambiguities in theory and differences in practice across institutions and context. Moreover, focusing on how individuals operate and are nested within institutions should also take into account general versus partial equilibrium effects: whereas interventions to encourage behavior change may improve

---

<sup>6</sup>We note below how county governance in Kenya has elected and non-elected officials responsible for the same functions, so that differences between the two should not arise from different roles or responsibilities of county functions. Because we will not pilot our interventions until after Kenya’s 2017 elections, we can also further test hypotheses related to the heterogeneous effects of incumbent victory, vote margin, party/coalition, and ethnicity of winning councilors who retain or gain seats in the county assemblies (depending on power and the administration of the pilot and rollout).

individual action along one margin, institutional mapping (and saturation designs) should inform whether and how information spillovers occur within an organization and whether agents re-adjust to commit malfeasance in other margins.

Kenya provides a propitious case to investigate institutional performance of devolved government. Bureaucratic inefficiency runs through central government agencies and new institutional reforms with devolved government provide important windows into the potential for improving institutional governance. Despite an explication of roles and responsibilities in the new constitution and subsequent legislation, the precise duties, lines of delegation within and between county offices (including managers, bureaucrats, and elected representatives) and across sub-county units, level of community engagement, and coordination of work with ministries, remains poorly understood in practice (Burbidge 2016). Counties are required to involve citizen participation and consultation on budgets and programming. Community involvement could prove effective as counties now form most Kenyans interactions with government and citizens remain acutely sensitive to government performance, corruption, and service delivery (Ferree, Long and Gibson 2014),<sup>7</sup> but citizens lack important knowledge on county structures and points of contact. These dynamics create variation in the functioning and performance of counties and uncertainty of delivery outcomes (Cheeseman, Lynch and Willis 2016).

In their structure, counties consist of an executive branch headed by a popularly elected governor who manages and oversees a legislative body (County Assembly) of elected assembly members and a more permanent non-elected secretariat (Public Service Board). These bodies have separate origin and survival. Assembly members are elected in single member wards (with the first county elections occurring in 2013), while the secretariat is appointed. Therefore, within the county system, certain bureaucrats are more directly accountable to citizens and voters, while others are more insulated, at least in design.

Despite different means of serving in the county, the Assembly and Board are responsible for overlapping jobs. The Assembly and Governor pass the county budget and decide how to spend money on services. The Board competitively appoints Sub-County Administrators. The Sub-County Administrators manage and oversee the general administrative functions in the sub-county unit, including (but not limited to): development of policies and plans; service delivery; developmental activities to empower the community, and exercising any functions and powers delegated by the Board. Within sub-county units are wards, equivalent to the elected unit represented in the County Assembly. The Board works to implement these programs and services, and competitively appoints Ward Administrators overseen by Sub-County Administrators. Alongside the efforts of Ward Administrators, individual councilors from the Assembly are also responsible for implementation and oversight at the ward level.

Working with partners, we have conducted an institutional mapping of the county governance structure. While we omit an extended discussion here, our results demonstrate a lack of clarity regarding the division of roles and responsibilities between the Assembly and the Board, particularly at the ward level. Rather, the institutional

---

<sup>7</sup>Like other developing countries, significant citizen activism in Kenya arises from calls to action on SMS and social media. Kenya has a robust ICT infrastructure with SMS and digital media frequently used for citizen mobilization. As of 2014, Kenya had 74 phone connections per 100 citizens (UN 2016) and skyrocketing usage of smartphones.

design of the county system appears to encourage replication of duties regarding the formation, implementation, and oversight of projects and services. For our purposes, it is however important to note that while assembly members at the ward level are theoretically directly accountable to voters in the ward and to county governors, Ward Administrators are accountable to bureaucratic managers at higher levels of the Board. Anecdotal evidence also suggests that Board managers exert oversight on assembly members and try to coordinate efforts between the Assembly and Board. Therefore, our theoretical prior regarding hypotheses 1 and 2 is that affecting managers' transaction costs and directly imposing monitoring could improve county performance either through the Ward Administrators, the council member, or both. Similarly, hypothesis 3 on community monitoring could also generate accountability and improved performance either via Board or Assembly members, although the effects could be stronger for elected members.

## 4 Research Design and Evaluation Strategy

To test our hypotheses, we will construct and launch a multi-channel ICT platform that works alongside of experimental variations regarding transaction costs, monitoring, and sanctioning. Crucially, to serve an oversight function, actors must be able to detect and report on (mis)behavior.<sup>8</sup> For our project, this will include a single platform with county government and citizen facing interfaces across ICT and digital media channels, allowing access via standard mobile phones, and web-enabled feature and smart phones. Working with partners, we will recruit a broad, scalable user-base of institutional actors (county/sub-county managers, ward administrators, and council members) and citizens to push and pull information. The platform will provide a venue to obtain and publish data about county functioning and performance from citizens' opinions and needs of their county government, user experience with the county and sub-county officials, perceptual and administrative data on levels and quality of services, reporting functionality for users, ratings of outcomes, and data on whether stakeholders provided responses when given reports. Information will be aggregated and relayed to managers, bureaucrats, and citizens at the relevant level, including reports of response action (Callen et al. N.d.). This will create a news-feed for officials to see representative messages from constituents and vice versa.

### 4.1 Experimental Treatments

**Stage 1, Treatment 1: *Reducing Transaction Costs*** To reduce transaction costs of obtaining information about how agents are behaving, we will treat county/sub-county managers with the introduction of the ICT platform. Institutional actors (including elected officials and managers) will receive a dashboard with information relevant to their function and area (described below). This will allow us to test hypothesis 1 to see whether simply reducing transaction costs by providing information to managers in the platform improves performance of the frontline bureaucrats who they manage. We will not experimentally vary the introduction of the platform itself,

---

<sup>8</sup>Audits require technology that is readily available and easy to use for detection of illicit acts. Conditional on detection, sources must report it and form a plausible source of sanctioning and/or communicate results to those who do (Callen et al. 2016, Callen and Long 2015).

but will give all wards within each county selected to participate access to the platform and compare this to baseline (this helps with institutional buy-in and limiting spillovers). We will measure trends in outcomes before the introduction of the platform itself and treat the before and after comparison (controlling for any trends) as the effect of the platform in reduction transaction costs. Non-experimentally, we can compare outcomes in these counties against other counties that do not receive the treatment using matching techniques.

**Stage 2, Treatment 2, *Increasing Bureaucratic Monitoring*** In the second stage, we will select a random sample of wards for agents to receive direct information that they are being audited and monitored by county managers. We will enhance this treatment by pushing reminders to managers to check their agents' work and respond with the platform information. Beyond reducing transaction costs from Treatment 1, this treatment may further improve performance by strictly encouraging monitoring to test hypothesis 2. This treatment will involve messaging to the ward administrators through our platform telling them that their behavior is being monitored through the platform by their managers (sub-county administrators and elected officials). We will also push messages with information on ward administrator's performance (e.g., absenteeism) in an easily digestible format that has been shown to have effects elsewhere (Callen et al. N.d.) (see Outcomes below). As specified by hypothesis 2, this treatment should increase the performance of ward officials above and beyond just the reduction in transaction costs.

**Stage 3, Treatment 3, *Increasing Community Monitoring & Sanctioning*** In a sample of wards who already received the second treatment, we will send information to county officials that citizens using the platform are reporting and monitoring in the platform. We will follow this with regular pushes to citizens through the platform with reports about the performance of county officials. These pushes will also include easily digestible reports about the performance of both bureaucrats and elected officials. This component tests hypothesis 3 to see if citizens participation and feedback generates heightened accountability through possible sanctioning by the community.

## 4.2 ICT Platform

Since poor government performance persists because of lack of principal oversight, managers and citizens require technological tools to provide information on the behavior of bureaucrats to detect and deter corruption. Beyond principals' monitoring, agents also operate more effectively if they have accessible information that they need to do their jobs well. We therefore focus our treatments on leveraging components to an ICT platform, and indeed, many citizens and government officials in developing democracies already enjoy access to potential ICT and digital media technologies through standard and web-enabled cell phones. We will build a single online platform with government and citizen facing interfaces and layers. We will recruit a broad, scalable user-base and engage institutional (county) and citizen actors to push information on performance and outcomes, and elicit feedback, across diverse channels regardless of phone type to allow for the broadest possible adoption as well as use of users' own devices. These channels will include SMS (text), USSD (text sessions), interactive voice recognition (IVR), and social media.

The platform will provide a venue to elicit information about citizens' opinions and needs to their county government, user experience with county officials and offices, and

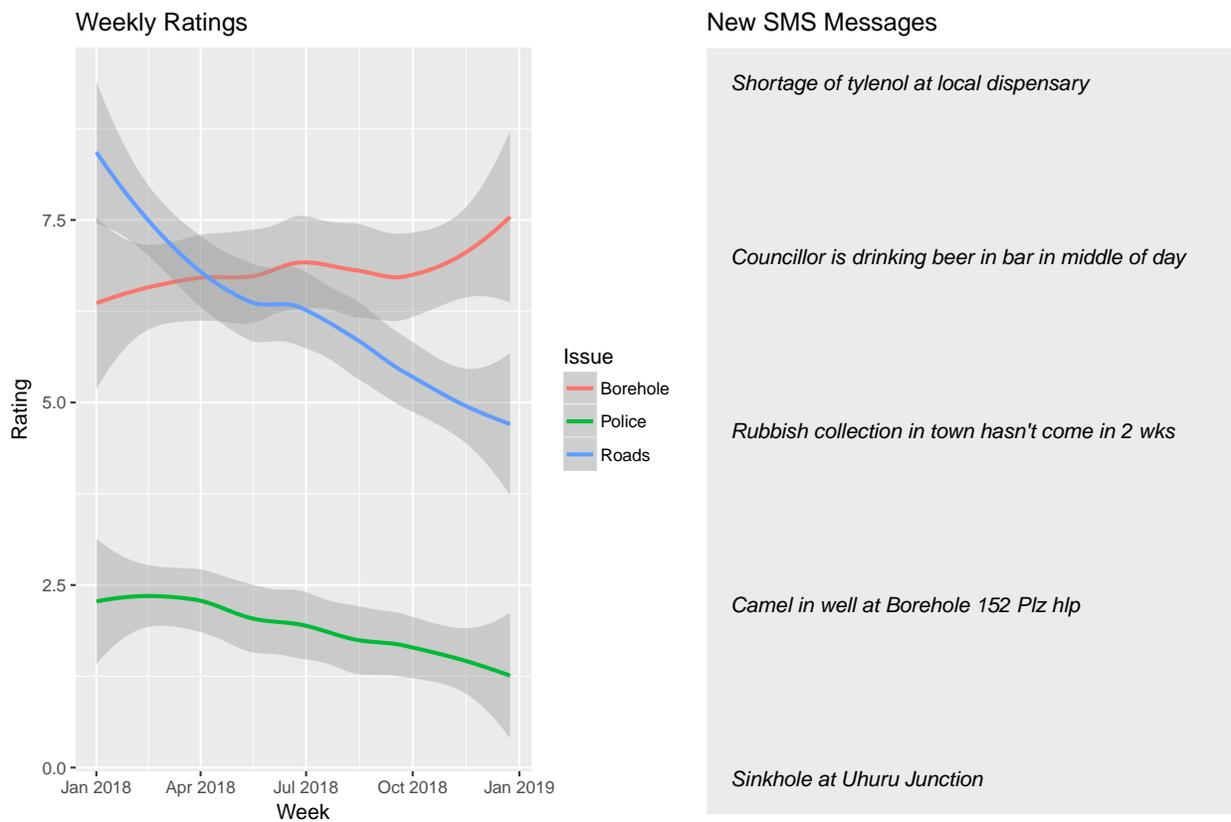


Figure 1: county official dashboard view

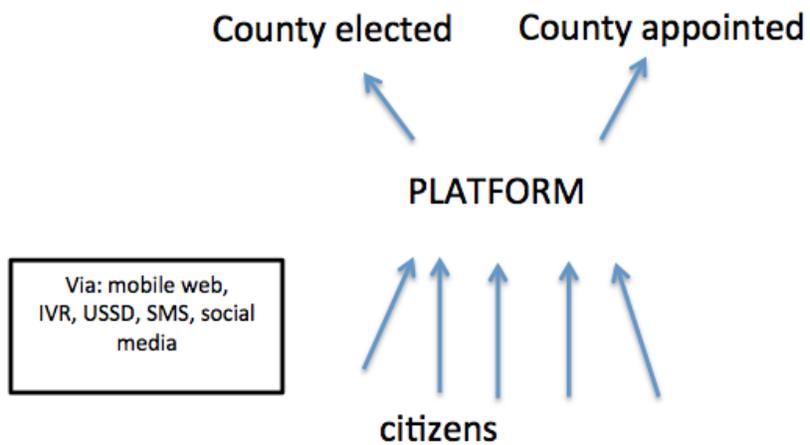


Figure 2: Platform schematic

data on levels and quality of services. The platform will have reporting functionality for users, ratings of outcomes, and data on whether stakeholders provided responses when given reports. The interface will be accessible regardless of channel, and the platform will allow for a variety of kinds of monitoring and data capture, including alert messages, photo capture, feeling thermometers, and user experience surveys (see Appendix A). The platform will relay ratings and allow feedback directly to county officials. Specifically, we will provide access to the results aggregation platform to Governors, Sub-County Administrators, Ward Councilors, and Ward Administrators. Each level of government will only see the data relevant to their governing task, including data on performance, outcomes, and ratings. In line with our research design below, Governors will see all of the information, and will be able to drill down in the data on a ward by ward basis to understand what is occurring within each ward so that the effect of the governor receiving information should be theoretically homogeneous across wards within their county.

### 4.3 Experimental Deployment

Each county in Kenya is divided into electoral wards (with an average of 30 wards per county), and wards will serve as our primary unit of randomization. We will pilot our program in a selection of counties. After the pilot, we are planning a phase 1 roll-out to 7-11 counties.<sup>9</sup> The selection includes a mix of rural counties (where we expect more homogeneity within wards) and urban counties (with more diversity within and across wards.), and counties designated as historically “marginalized.” The target population of the intervention will include all residents within a county, and will incorporate actors in the county governance structure (e.g., county managers, governors, councilors, secretariat members). Baseline survey data (described below) will provide data for treatment assignment and power calculations.

#### Design

We will implement a Stepped Wedge (Waitlist, Phase-in) design. Unlike typical cluster randomized trials (CRT), we will pursue a crossover design to increase stakeholder buy-in, while also addressing ethical concerns about withholding the program from some communities. Appendix B shows the randomization protocol for a stepped-wedge design compared to a parallel or simple crossover design as visualized by [Hussey and Hughes \(2007\)](#).

Our design is slightly more complicated than this stepped-wedge design given our multiple treatment conditions (shown in Table 1). Table 1 illustrates that we will first implement **Treatment 1** — our ICT platform (reducing transaction costs)— for all wards (the difference between columns 1 and 2, which is a before and after comparison). Subsequently, we will implement a randomized incomplete stepped-wedge design where each treatment is additive (the difference between columns 2, the baseline, and 3-5), as in [Miguel and Kremer \(2004\)](#).

---

<sup>9</sup>Per ongoing conversations with partners and policy dialogue with counties, we are not able to list them currently but can discuss at the workshop.

## Estimation

Following pre-program measurement, we will roll out Treatments 2 and 3 over time after every unit receives Treatment 1 at the same time.<sup>10</sup> More specifically, wards within counties will be randomly assigned to phase-in at different times to allow for measurement of changes in outcomes (including service provision, county priority responsiveness, and citizen reporting and mobilization) both within and across county-ward units.

We can measure impact in two ways, reflecting disciplinary discussions about how to estimate treatment effects from stepped wedge design models. The medical literature tends to use model-based inference for stepped wedge designs (Hussey and Hughes 2007), using a random effects models to estimate

$$Y_{itj} = \mu + \alpha_i + \beta_t + \theta X_{ij} + \varepsilon_{itj}$$

where  $\mu$  is the cluster (ward) average for the control group,  $\alpha$  is a normally distributed cluster random intercept,  $\beta$  is a fixed effect for each period of measurement, and  $X_{ij}$  is an indicator of the treatment status and  $\theta$  the treatment effect. In our case,  $\theta$  should be a vector ( $\theta'$ ), since we have two treatments to estimate beyond the provision of the platform.

The economics literature uses least squares estimation with clustered standard errors. For example, Miguel and Kremer (2004) employ an incomplete stepped-wedge design (the control group never received treatment in the last time period) to estimate the impacts of de-worming in Kenya. They also estimate spillover effects, which has elicited controversy (e.g., Aiken et al. 2015).

Following Miguel and Kremer, we can estimate our treatment effects for our two interventions beyond the provision of the platform itself using the following equation:

$$Y_{ijkt} = \alpha + \sum_{t=1}^K \beta_i T_{it} + \sum_{t=1}^{K-w} \gamma_i T_{it} + X'_{ijkt} + D_k$$

Following Miguel and Kremer's notation, we estimate  $Y_{ijkt}$  where  $i$  indexes ward;  $j$  indexes one of {survey respondent, administrator, elected official};  $k$  are dummy variables for counties. Here  $t$  (in  $1 \dots K$ ), unlike the medical literature, is not a time period fixed effect but represent represents a dummy variable for the  $t^{th}$  time period in treatment. Therefore, each  $\beta_i$  or  $\gamma_i$  is the coefficient on ward assignment to treatment with our program for each additional roll-out period of the program.<sup>11</sup> We obtain effects based on within person variation over time, based on ward treatment status. We can also include covariates as represented by the vector  $X$ 's. Our specification differs from Miguel and Kremer in that we also include fixed effects for counties ( $D_k$ ) and have to estimate the effect of multiple treatments.

---

<sup>10</sup>It may be possible to phase the roll out of Treatment 1, but we worry that not enough time will have passed to measure impact.

<sup>11</sup>Note, that there will be fewer  $\gamma$  coefficients as compared to  $\beta$  ( $K - w$ ) coefficients because **Treatment 3** is rolled out later. Pooling of multiple time periods is also an issue for which we seek advice.

## Power, Sample Size, and Reporting

We will reach a final decision with our partners about the number of periods in which to roll out the study. This will in part depend on power calculations, as the formula for such calculations is complicated and the subject of a quickly developing literature. Power (and sample size) in CRTs depends not only on within and between cluster variation, but also the number of clusters randomized at each step, the number of steps in the design, and assumptions about the role time plays (Hemming et al. 2011, 2015, Hughes, Granston and Heagerty 2015, Hussey and Hughes 2007).<sup>12</sup>

## 4.4 Measurement and Outcomes

Broadly, our outcomes of interest include reducing corruption and waste in counties, improving services and government responsiveness to citizen’s needs, and generating citizen participation and community mobilization. We will track data that measure government spending and responsiveness, citizen attitudinal measures about priorities and ratings of outcomes, and county level actions on whether counties are “working” (i.e., successful at bringing government activities and outcomes closer in line with citizen preferences). Specifically, we will obtain and display relevant outcome measurements in the ICT platform, including citizen ratings and feedback (elicited through household surveys and deputized “citizen monitor” in village units), administrative data collected by county and reform agencies (on meetings, budgets, project spending, service delivery), and independently by our partners.

**Baseline Survey** Within our system, we need to recruit two types of citizen users that provide measures on county performance, operations, and outcomes. The first type of citizen user will be a “typical” citizen user of the system obtained from a representative baseline survey (either household or via CATI). The second type of user will be a “citizen monitor,” who will be a highly engaged user withing the system that can carry out specific monitoring tasks. Generally, we will attempt to recruit with the same level of *per capita* intensity across all wards. Most likely, we will engage in a multimedia campaign to enroll users in the platforms. One option to recruit our “citizen monitors” would be to identify users with high volume and high quality usage of the platform to carry out monitoring tasks. A second option to recruit citizen monitors would be to recruit a sample of individuals from the baseline survey. Using both types of citizen models allows us to gain dispersed but also consistent measures of citizens’ views and ratings of service provision, county responsiveness, performance of institutional actors, and participation in county, ward, and village unit meetings.<sup>13</sup>

---

<sup>12</sup>We also seek advice on time as a factor both in terms of time trends and power calculations. It is reasonable, for example, for Miguel and Kremer to assume that no other de-worming program occurred and therefore they can assume no time trends in their study. Our program arguably differs.

<sup>13</sup>Specifically, a typical citizen user may not consistently provide quality measurements across citizen users across wards. However, this type of user is a more realistic method of viral adoption and citizen-scaling. A deputized citizen monitor will provide higher quality and more robust measures consistently across wards, but this is a less scalable model (in the absence of highly engaged citizens who always provide quality data.) For the purposes of evaluation, we believe we need deputized citizen monitors, but for scaling, we are interested to build out the robustness of citizen users. In Ferree et al. (2017) and Erlich et al. (2017), we document the opportunities and challenges of this trade-off using an ICT platform we deployed during South Africa’s 2014 election.

**Ward Outcomes** Ward level outcomes are measured at the ward representative and ward administrator level, aggregate ward population averages, and individual citizens of wards. Measurement of both Ward Representatives and Ward Administrators requires audits from outside the government themselves, as well as measures derived from government documentation. Since under the new constitution each wards elects one ward representative to the county assembly, we can measure the effect of the ward treatment on these ward representatives. At the level of ward officials we expect to measure i) the number of public meetings the ward official attends ii) the presence of the ward official at his proscribed constituency hours, as measured by “citizen monitors” and the independent auditors, and iii) the number of sessions and committees where the ward representative votes from the constituency assembly record. To do so, we will employ two types of audit. First, we will employ a group of outside trained individuals (from our research partnerships) to take independent measurements (akin to an external and independent auditor). Second, “citizen monitors” (typical and/or deputized) we identify through recruitment during the baseline survey can serve the same type of auditing. On the assumption that external professional auditors will supply robust reporting data, using both independent and citizen auditing methods will also allow us to evaluate the quality of measurements provided by citizen monitors.

**Electoral data** Kenya does not release individual turnout data, be we can estimate participation rates across treated and untreated wards and measure differences in the proportion of incumbent ward representatives removed from office per our examination of heterogeneous treatment effects discussed above.

**Administrative data** To understand difference in government performance, we measure differences in outcomes across the devolved services in the county system with objective data provided by counties.

**Survey data** Since those who remain outside of the platform will never be measured by the platform we will use our baseline and endline survey data collected with our partners to measure pre and post-treatment differences in both the different arms of our study at the ward level for all citizens regardless of whether the actually engaged in the platform.<sup>14</sup> We can also measure within ward differences between those who used the app and those who did not on the individual level.

**Platform data** We will use indicators collected within the platform to measure differences in wards across our treatments. Outcomes include engagement in the platform as a proxy for increased monitoring of the principal. These proxy measures include: time spent on the platform for those with smartphones or tablets (including all politicians and administrators in the study), the number of messages sent for those users that have feature phones, and the total number of activities in which users engaged, per Erlich et al. (2017).

## 5 Conclusion

Our project will provide important theoretical, methodological, and empirical insights to the study of corruption and institutional performance of devolved government agencies in the developing world, demonstrating evidence of impact for policy action. Whereas institutions like public agencies form the locus of change for economic growth

---

<sup>14</sup>Importantly, we will also use the baseline survey for pre-treatment ward stratification.

and service delivery, institutional variation in performance has thus far prevented a consistent body of evidence of what works to improve bureaucratic processes and outcomes (Finan, Olken and Pande forthcoming). Despite continued push for devolved government as a means to improve institutional governance and achieve institutional economies of scale in Kenya specifically, this policy and new institutional intervention has not been rigorously or systematically evaluated, with evidence showing significant variation in outcomes.

Theoretically, we will leverage an institutional mapping of county governance to inform research and platform designs that harness the interaction of institutional and individual actors, including how external actors (citizens) may improve institutional performance. Acknowledging the principal-agent problems for agency and citizen principals to receive robust information, we examine whether and how information treatments and technology improves that process by reducing transaction costs, increasing monitoring, and comparing this to citizen monitoring. Accordingly, we pursue a more inclusive participatory development approach to monitoring government performance, using cheap, accessible, scalable ICT that can be replicated across contexts, working directly with county governments and citizens. By looking at county government, we leverage institutional economies of scale to examine the full range of /examine impact on a range of agency outcomes across sectors related to economic development and citizen welfare, rather than focus on one sector. This is important since these sectors are linked locally in the county council. We include the participation of agencies so that they are provided a feedback mechanism, and institutional autonomy and platform design can limit spillovers.

The results from our study should also provide a number of important comparative insights. First, our project should provide important evidence to action on improving devolved government structures in Kenya. Kenya is not alone in having local government executive and legislative bodies similar to counties. In fact, most African countries have local municipal councils, and these bodies have grown more important in their operations and robust in design given pushes towards devolution over the last decade, although they have variation in degrees of fiscal and governance autonomy. Second, our platform design will specifically incorporate and reflect the needs of county managers and workers, and citizens reporting on county activities and outcomes. However, the technology itself, its multi-channel capabilities and usage, are agnostic with respect to adding layers and how data is visualized and displayed. Therefore, we think building out this system will lend important technological evidence of impact to other similar platforms. Our results regarding citizen uptake, usage, engagement, and reporting mechanism are not just useful for local government reporting, but with respect to all crowdsourced platforms in developing countries. We believe our project demonstrates significant policy potential by helping policymakers to better understand and address institutional challenges. Our strategy is cost-effective, and our government partners indicate that they believe the platform will assist their decision-making and programming. After the platform is developed and research finalized, there is no real marginal cost increase for scale-up.

## References

- Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–1401.
- Aiken, Alexander M., Calum Davey, James R. Hargreaves and Richard J. Hayes. 2015. "Re-analysis of health and educational impacts of a school-based deworming programme in western Kenya: a pure replication." *International Journal of Epidemiology* 44(5):1572–1580.  
**URL:** <https://academic.oup.com/ije/article/44/5/1572/2594560/Re-analysis-of-health-and-educational-impacts-of-a>
- Banerjee, Abhijit V., Esther Duflo and Rachel Glennerster. 2008. "Putting a Band-Aid on a Corpse: Incentive for Nurses in the Indian Public Health Care System." *Journal of the European Economic Association* 6(2–3):487–500.
- Basinga, Paulin, Paul J. Gertler, Agnes Binagwaho, Agnes Soucat, Jennifer Sturdy and Christel Vermeersch. 2011. "Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation." *Lancet* 377(9975):1421–1428.
- Becker, Gary S. and George J. Stigler. 1974. "Law enforcement, malfeasance, and compensation of enforcers." *The Journal of Legal Studies* 3(1):1–18.
- Besley, Timothy and Torsten Persson. 2014. "Why do Developing Countries Tax so Little?" *Journal of Economic Perspectives* 28(4):91–120.
- Björkman, Martina and Jakob Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *The Quarterly Journal of Economics* 124(2):735–769.
- Breuer, Anita, Todd Landman and Dorothea Farquhar. 2015. "Social media and protest mobilization: Evidence from the Tunisian revolution." *Democratization* 22(4):764–792.
- Burbidge, Dominic. 2015. *The Shadow of Kenyan Democracy: Widespread Expectations of Widespread Corruption*. New York, NY: Routledge.
- Burbidge, Dominic. 2016. The Dawn of Devolved Government in Kenya. Technical report Oxford Human Rights Hub.  
**URL:** <http://ohrh.law.ox.ac.uk/the-dawn-of-devolved-government-in-kenya/>
- Callen, Michael, Clark Gibson, Danielle F. Jung and James D. Long. 2016. "Improving Electoral Integrity with Information and Communications Technology." *Journal of Experimental Political Science* 3(1):4–17.
- Callen, Michael and James D Long. 2015. "Institutional corruption and election fraud: Evidence from a field experiment in Afghanistan." *The American Economic Review* 105(1):354–381.

- Callen, Michael, Saad Gulzar, Syed Ali Hasanain and Yasir Khan. N.d. “The Political Economy of Public Sector Absence: Experimental Evidence from Pakistan.” *NBER Working Paper 22340*. Forthcoming.  
**URL:** <http://www.nber.org/papers/w22340>
- Callen, Michael, Saad Gulzar, Syed Ali Hasanain, Yasir Khan and Arman Rezaee. 2015. “Personalities and Public Sector Performance: Evidence from a Health Experiment in Pakistan.” *NBER Working Paper 21180* .
- Chabal, Patrick and Jean-Pascal Daloz. 1999. *Africa Works: Disorder as Political Instrument*. Bloomington, IN: Indiana University Press.
- Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed*. Cambridge, UK: Cambridge University Press.
- 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.
- Cheeseman, Nic, Gabrielle Lynch and Justin Willis. 2016. “Decentralisation in Kenya: the governance of governors.” *Journal of Modern African Studies* 54(1):1–35.
- Condra, Luke, Michael Callen, Radha K. Iyengar, James D Long and Jacob N. Shapiro. 2016. “Damaging Democracy? Security Provision and Turnout in Afghan Elections.”  
**URL:** [http://scholar.princeton.edu/sites/default/files/jns/files/ccils\\_damaging\\_democracy.pdf](http://scholar.princeton.edu/sites/default/files/jns/files/ccils_damaging_democracy.pdf)
- Dal Bó, Ernesto and Frederico Finan. 2016. “At the Intersection: A Review of Institutions in Economic Development.” *Economic Development and Institutions Working Paper* (16/11.01).
- Dal Bó, Ernesto, Frederico Finan and Martín A Rossi. 2013. “Strengthening state capabilities: The role of financial incentives in the call to public service\*.” *The Quarterly Journal of Economics* 128(3):1169–1218.
- Duflo, Esther, Rema Hanna and Stephen P. Ryan. 2012. “Incentives Work: Getting Teachers to Come to School.” *American Economic Review* 102(4):1241–1278.
- Easterly, William and Ross Levine. 1997. “Africa’s Growth Tragedy: Policies and Ethnic Divisions.” *The Quarterly Journal of Economics* 112(4):1203–1250.
- Engerman, Stanley L. and Kenneth L. Sokoloff. 1997. Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States. In *How Latin America Fell Behind*, ed. Stephen Haber. Stanford: Stanford University Press pp. 260–304.
- Erlich, Aaron, Danielle F. Jung, James D. Long and Craig McIntosh. 2017. “The double-edged sword of cross-channel ICT engineering for development.”
- Ferree, Karen E., Clark C. Gibson, Danielle F. Jung and D. Long, James. 2017. “How Technology Shapes the Crowd: Participation in the 2014 South African Election.” *Working Paper* .

- Ferree, Karen, James D. Long and Clark Gibson. 2014. "Voting Behavior and Electoral Irregularities in Kenya's 2013 Election." *Journal of Eastern African Studies* 8(1):153–172.
- Finan, Frederico, Benjamin Olken and Rohini Pande. forthcoming. The Personnel Economics of the State. In *Handbook of Field Experiments*, ed. Abhijit Banerjee and Esther Duflo. Cambridge, MA: Elsevier.
- Fisman, Raymond and Miriam A. Golden. forthcoming. *Corruption: What everyone needs to know*. New York, NY: Oxford University Press.
- Fukuyama, Francis. 2014. *Political Order and Political Decay*. New York, NY: Farrar, Straus, and Giroux.
- Gupta, Akhil. 2012. *Red Tape: Bureaucracy, Structural Violence, and Poverty in India*. Durham, NC: Duke University Press.
- Hemming, K., A. Girling, A. Sitch, J. Marsh and R. Lilford. 2011. "Sample size calculations for cluster randomized controlled trials with a fixed number of clusters." *BMC Medical Research Methodology* 11.
- Hemming, K., T. P. Haines, P. J. Chilton, A. J. Girling and R. J. Lilford. 2015. "The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting." *BMJ* 350:h391.  
**URL:** <http://www.bmj.com.proxy3.library.mcgill.ca/content/350/bmj.h391>
- Hughes, James P., Tanya S. Granston and Patrick J. Heagerty. 2015. "Current issues in the design and analysis of stepped wedge trials." *Contemporary Clinical Trials* 45, Part A:55–60.  
**URL:** <http://www.sciencedirect.com/science/article/pii/S1551714415300434>
- Hussey, Michael A. and James P. Hughes. 2007. "Design and analysis of stepped wedge cluster randomized trials." *Contemporary Clinical Trials* 28(2):182–191.  
**URL:** <http://www.sciencedirect.com/science/article/pii/S1551714406000632>
- Keefer, Philip and Razvan Vlaicu. 2008. "Democracy, Credibility, and Clientelism." *The Journal of Law, Economics, and Organization* 24(2):371–406.
- Kiewiet, Roderick D. and Mathew D. McCubbins. 1991. *The Logic of Delegation*. Chicago: Chicago University Press.
- Levi, Margaret. 1988. *Of rule and revenue*. Berkeley: University of California Press.
- Mauro, Paolo. 1995. "Corruption and Growth." *The Quarterly Journal of Economics* 110(3):681–712.
- Miguel, Edward and Michael Kremer. 2004. "Worms: identifying impacts on education and health in the presence of treatment externalities." *Econometrica* 72(1):159–217.  
**URL:** <http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0262.2004.00481.x/full>
- Mungiu-Pippidi, Alina. 2015. *The Quest for Good Governance: How Societies Develop Control of Corruption*. Cambridge, UK: Cambridge University Press.

- North, Douglass C., John Joseph Wallis and Barry R. Weingast. 2009. *Violence and Social Orders*. Cambridge, UK: Cambridge University Press.
- Robinson, James A, Daron Acemoglu and Simon Johnson. 2005. "Institutions as a Fundamental Cause of Long-Run Growth." *Handbook of Economic Growth* 1A:386–472.
- Rose-Ackerman, Susan. 1999. "Political Corruption and Democracy." *Connecticut Journal of International Law* 14(2):363–378.
- Rose-Ackerman, Susan and Bonnie J. Palifka. 2016. *Corruption and Government: Causes, Consequences, and Reform*. Cambridge: Cambridge University Press.
- Shleifer, Andrei and Robert W. Vishny. 1994. "Politicians and Firms." *The Quarterly Journal of Economics* 4(109):995–1025.
- Tufekci, Zeynep and Christopher Wilson. 2012. "Social media and the decision to participate in political protest: Observations from Tahrir Square." *Journal of Communication* 62(2):363–379.

	Period 1	Period 2	Period 3	Period 4	Period 5
Cluster 1	0	1	2	2	3
Cluster 2	0	1	2	2	3
Cluster 3	0	1	1	3	3
Cluster 4	0	1	1	3	3
Cluster 5	0	1	1	1	1

Table 1: Visualization of an incomplete stepped wedge design with three treatments. The first treatment only allows for a before and after comparison. All additional treatments are additive so 3 signifies **Treatment 3** in addition to **Treatment 2**.

# Appendix

## A Messages from System

Your government heard you. Thanks for reporting.

We have been working on services in your area.

New borehole in your ward because of citizens like you report.

Figure 3: Text announcement of services provided by county government

Your voice is important. County officials are holding a meeting in your ward.

### **Location**

Makadara

### **Time**

14:06

### **Date**

January 30, 2017

Figure 4: Text announcement of a local meeting by county

## B Stepped Wedge Visualization

<u>Parallel</u>				<u>Crossover</u>				<u>Stepped Wedge</u>						
Time				Time				Time						
		1			1	2			1	2	3	4	5	
Cluster	1	1			1	0			1	0	1	1	1	1
	2	1			2	0			2	0	0	1	1	1
	3	0			3	1			3	0	0	0	1	1
	4	0			4	1			4	0	0	0	0	1

Figure 5: Visualization of stepped-wedge design from Hussey and Hughes (2007)