The Challenges of Using Citizen Reporting to Improve Public Services: A Field Experiment and Framework

Mark T. Buntaine[1], Patrick Hunnicutt[2], and Polycarp Komakech[3]

[1]University of California, Santa Barbara, buntaine@bren.ucsb.edu

[2]University of California, Santa Barbara, phunnicutt@bren.ucsb.edu

[3]University of California, Santa Barbara, k.polycarp@gmail.com

October 10, 2018

Abstract

Governments around the world are investing in technologies that allow massive, frequent, and localized contact with citizens, though there is little evidence about how these technologies impact the delivery of public services. We report a large-scale randomized controlled trial that involved recruiting 50 citizens in each of 100 neighborhoods across Kampala, Uganda to provide weekly reports on the delivery of solid waste services via an SMS-messaging platform to a municipal government, resulting in 23,862 reports during the study period. Citizen reporting did not reduce waste accumulation. More positively, reporting reduced the amount of burning and unmanaged waste piles for a time, but this positive result did not persist after an unexpected staff restructuring in the unit responsible for waste management. Waste collection did not improve in zones with more reports or more dissatisfied reporters. Using our observations as participants in development and deployment of the platform and interviews with key staff at the government agency receiving citizen reports, we show how the adoption of new technologies to collect data from citizens requires new capacities and data consistent enough to reduce uncertainty about the allocation of effort. We provide a formal framework for analyzing the challenge of utilizing citizen-sourced data for the management of public services.

1 Introduction

Governments around the world are building or adopting platforms to collect and process feedback from citizens about public services. New information and communication technologies (ICTs) enable governments to collect dispersed observations from citizens, opinions about service quality, and ideas for improvements to services, all at lower costs than traditional methods. In addition, they allow governments to communicate with the public about responses to concerns. These tools also increase the potential for public agencies to track trends in service quality, manage follow-up actions, and engage a broader set of citizens in collaborative management (Bertot, Jaeger and Grimes, 2010; Rotberg and Aker, 2013; Grossman, Humphreys and Sacramone-Lutz, 2014).

We provide a field experimental test of the impact of citizen reporting on public services and first-hand qualitative analysis about the challenges of adopting and operating a citizen-reporting platform. Based on the results, we present a general framework for understanding when ICT tools that help governments interact with citizens have the greatest potential to improve the provision of public services.

This kind of test and framework is needed because while data from citizens can be collected more broadly and in a more timely manner than many top-down approaches to monitoring public services, the adoption and operation of such platforms can be difficult. Processing new flows of data and turning them into information that can be used for decision-making requires new skills and capacities, potentially implying significant costs. Being responsive to new information may require a realignment of work effort, which can be costly, politically contentious, or limited by existing procedure.

Besides practical considerations about acting on data, the quality and consistency of data from citizens about public services may be significantly lower than what can be collected using top-down monitoring. Citizen-sourced data is often unstructured, noisy, and inconsistent, which creates challenges for using it to improve public services. For instance, citizens who want governments to exert more effort in improving public services may have incentives to send information that will attract attention regardless of actual service quality. If this is the case, trust in citizen-sourced data might not be warranted. Furthermore, citizen-sourced data may be inconsistent enough that it does not reduce uncertainty about decisions, eroding its usefulness.

In partnership with the Kampala Capital City Authority (KCCA) in Uganda, we study the adoption of a new SMS text-messaging platform to collect, process, and aggregate citizen feedback about waste collection services, which resulted in 23,862 reports over the approximately nine-month study period. First, we present results from a large-scale randomized field experiment that involved recruiting reporters from randomly-assigned neighborhoods to send reports about waste collection services to the KCCA waste management unit. Because the accumulation of solid waste in informal piles is visible and results from low-quality collection services, we are able to independently audit whether citizen reporting decreases informal dumping and burning of waste by residents.

Second, we use participant observations to interpret the opportunities and challenges of adopting new technologies and processing data to improve public services. Our research team was embedded at the KCCA, the agency that adopted the SMS platform, to assist with its development and use. We recorded our participant observations systematically. We also conducted in-depth interviews with all of the KCCA staff who interacted with the platform to understand the opportunities and barriers of using citizen-sourced data to improve public services. We thus provide first-hand evidence about the challenges that governments will face when adopting these tools and the capacities that they should have in place to leverage citizen-sourced data for the management of public services.

To preview our results, we find that ICT-enabled citizen reporting did not significantly impact waste accumulation in Kampala neighborhoods. In the nine-month study period reported here, we do not find evidence that the amount of waste accumulation decreased in neighborhoods assigned to citizen reporting, as compared to neighborhoods without citizen reporting. We find some promising results in terms of the proportion of piles with burning and the amount of non-organic waste at the first post-treatment audit of informal waste piles five months after the baseline, but these results do not persist to the second post-treatment audit of informal waste piles four months later.

We gained additional insights about the process of using citizen-sourced data from an unexpected and disruptive reorganization at the KCCA, which included staff and management rotations in the waste management unit that co-developed and operated the SMS-reporting platform. This reorganization happened immediately after our first post-treatment data collection wave. The new director and team were much less committed to using the platform, considered the data that it produced to be unreliable and inconsistent, and preferred to develop their own systems for monitoring the quality of waste services. We show how differences in perceptions of consistency and cost of citizen-sourced data had major implications for the potential of citizen reporting.

Citizen-sourced data, even though it can be massive, timely, and localized, is no panacea for the problems facing public sector managers with limited budgets and costly options for monitoring the delivery of public services. While citizen reporting can save costs and allow for a greater proportion of available public resources to be spent on improving services, it is also likely to create a more inconsistent and unreliable data stream across many settings, requiring significant effort for processing and interpretation. We conclude with a formal framework that illustrates how optimism about citizen-sourced data is likely misplaced across a range of realistic circumstances.

2 Monitoring the Delivery of Public Services

Public agencies must monitor public services to ensure that they are delivered appropriately to citizens. Without information about the ultimate delivery of services, public agencies cannot hold frontline civil servants or contractors accountable for carrying out their mandates. They also cannot discover patterns of problems in delivering services and optimize provision to overcome these problems. Without monitoring, public agencies have little information on whether services meet the demands of citizens.

Because of these challenges, public agencies invest in a variety of information systems to help manage public services. These monitoring systems include field audits, surveys with citizens, reporting requirements for frontline providers, automatic sensors, GPS tracking of providers, and reporting hotlines. While different techniques are fitted to different kinds of situations, they all require significant investment of time and effort by public agencies to collect and process data, particularly if data is needed frequently and at a broad scale.

Alternatively, public agencies can rely on complaints from citizens about poor services (McCubbins and Schwartz, 1984), but this kind of information is often unstructured and difficult to respond to systematically, and may often be actively biased (de Figueiredo Jr, Spiller and Urbiztondo, 1999).

The idea that information technologies can facilitate citizen monitoring and improve public service delivery has sparked cautious enthusiasm because citizens experience the delivery of services first-hand (Oates, 2003; Grossman, Humphreys and Sacramone-Lutz, 2014; Charalabidis et al., 2012; Linders, 2012; Zurovac, Talisuna and Snow, 2012; Rotberg and Aker, 2013). If it were possible to gain information from many citizens repeatedly, consistently, and accurately, it might be possible to both significantly expand the amount of information available to managers of public services, but also to decrease the costs of monitoring as compared to top-down monitoring systems. Sourcing information from citizens might also make government agencies more responsive to citizen demands and reduce the principle-agent problems that hamper effective service provision. Yet, there are significant political, operational, and data-processing challenges to employing spatial information sourced from non-representative groups of citizens into the delivery of public services (Ntaliani, Costopoulou and Karetsos, 2008; Mossberger, Wu and Crawford, 2013; Evans and Campos, 2013).

While several prominent platforms generate citizen monitoring of public services in developed countries (e.g., SeeClickFix, FixMyStreet, NoiseTube), these platforms are not designed to facilitate research about foundational questions of citizen-sourced data provision, quality, and impact. To this point, the majority of research on citizen-sourced information deals with either disaster responses (e.g., Zook et al., 2010) or environmental monitoring (e.g., Connors, Lei and Kelly, 2012). Neither issue deals with eliciting long-term improvements to public services. Scholars from fields as varied as information science, technology studies, political science, and public administration recognize that research about ICT-enabled citizen feedback requires more focused empirical research approaches (Linders, 2012; Charalabidis et al., 2012; Saxton, Oh and Kishore, 2013; Seltzer and Mahmoudi, 2013). Recent research has begun this work related to participation in reporting on public services (Sjoberg, Mellon and Peixoto, 2017; Buntaine, Nielson and Skaggs, 2017). In addition, some results are emerging about the lack of impact when citizens report deficiencies to politicians, rather than government agencies (Grossman, Platas and Rodden, 2017).

Results about the impacts of citizen monitoring of governments outside of ICT-platforms are likewise mixed. Some studies indicate that monitoring alone is insufficient to generate substantial impact (Olken, 2007; Banerjee, Deaton and Duflo, 2004; Banerjee et al., 2010). Monitoring only seems to have an impact on the delivery of public services when it is tied to credible enforcement mechanisms over providers (Björkman and Svensson, 2009) or when managers are predisposed to working with citizens, based on an interest in securing information, trust, or political advantage (Bryer, 2009).

Yet, the examples of citizen monitoring and reporting about public services that exist have not approached the scale and complexity of the kinds of ICT-enabled platforms that are being adopted by government agencies to interface with large numbers of citizens systematically. Such platforms remove the personal and social connections that characterize many of the smaller-scale efforts to improve service provision through citizen monitoring. They also significantly increase the burden of data processing and the scope of responsiveness. Yet, they offer broader reach and timeliness than smaller-scale or top-down methods of monitoring the delivery of public services. Given considerable interest, we aim to offer both rigorous experimental evidence about impacts and first-hand qualitative analysis about opportunities and challenges of government agencies adopting new ways to interact with citizens.

3 Research Design

In light of the mixed expectations for citizen monitoring and the limited testing at the scale new technologies enable, we test whether the large-scale solicitation of feedback about the provision of public services from citizens will enable more effective delivery of solid waste services in Kampala. We hypothesized that citizen monitoring could play an important role in improving waste services because it both provides information that is hard for our government partner to collect broadly and because citizen monitoring reveals the location of need for improved services.

3.1 Solid Waste Management in Kampala

The city of Kampala is transforming how it provides solid waste services to citizens. Previously, the collection and disposal of solid waste was the sole responsibility of the KCCA, which managed every aspect of solid waste collection, transportation, and disposal. The KCCA also bore the entire cost of providing solid waste services. The only role residents played was delivering their solid waste to collection locations.

In the last few years, the KCCA has gradually adopted a public-private partnership (PPP) approach to providing solid waste services. Under this approach, the KCCA contracts out the management of solid waste services to private concessionaires that are given the responsibility to collect, transport, and dispose of solid waste from particular areas of the city. This places city managers in a challenging position, because they cannot easily monitor the activities of private contractors, which have incentives to provide services where they can induce (sometimes illegally) payment from citizens. In other settings, private companies contracted to remove solid waste provide services of lower quality to groups of people that are unable to share information about their performance with governments (Oteng-Ababio, 2010; Katusiimeh, Mol and Burger, 2012). The KCCA can use information on where services are and are not being delivered to allocate oversight and clean-up efforts.

Under the PPP, the private concessionaires are allowed to charge the residents a specified amount of money in return for collecting their solid waste on a door-to-door basis. At the same time, they are contractually required to provide common collection points available to all residents regardless of ability to pay. The incentive to maximize revenue from citizens through door-to-door collection is at odds with requirements to make collection widely accessible, so contractors have mostly failed to establish common collection points.

The deterioration of solid waste services under this model has led to a spike in interest about waste management among citizens. Yet, to this point, the KCCA has not had a way to collect information on such demands or observations about where concessionaires are not fulfilling their contractual obligations to establish common collection points or even offering collection for households that are willing to pay for door-to-door collection. Under these conditions, the KCCA needs information about how to allocate oversight and resources to supplementary clean-up efforts. Its small professional team at the city headquarters cannot easily monitor the situation around Kampala, which has almost four million residents.

3.2 The Platform

The rapid proliferation of mobile phones in Kampala offers an opportunity to engage a much broader range of citizens in timely ways. The latest statistics in Kampala indicate that more than 90% of adults own a mobile phone (of Statistics, N.d.), creating the potential for significant interaction between the KCCA and citizens in ways that solve information problems related to the allocation of oversight and supplementary clean-ups.

Indeed, the KCCA faces similar problems of monitoring and accountability for solid waste management as many other parts of the world (Bhuiyan, 2010; Okot-Okumu and Nyenje, 2011). With Kampala growing rapidly like many developing cities (Vermeiren et al., 2012), the need to improve the quality and scale of services is pressing. A majority of solid waste in Kampala is disposed of in informal dumps or openly burned in streets and alleys. A large majority of residents are personally concerned and dissatisfied with solid waste services (see Buntaine, Nielson and Skaggs, 2017).

Beginning in 2014, our research team approached the KCCA to investigate whether they would be interested in adopting and testing a platform that would enable them to collect information from citizens about the quality of waste collection services in real-time and at the scale of neighborhoods. The idea was met with enthusiasm from key leadership, overcoming a key constraint on these kinds of efforts (Hansen and Norup, 2017). The platform was co-developed over time and was based on toll-free SMS-messaging from residents in randomly-assigned neighborhoods, who would be invited to sign up as reporters. Because we recruited these citizen-reporters in the field, all of the reports can be tagged to individual "zones" throughout Kampala, which are the lowest-level administrative unit (LC1) in both the city and throughout Uganda.

In phases from November 2015 to August 2017, we prompted citizens to send reports about various aspects of solid waste management to a single, toll-free SMS shortcode established for the project. To process citizen reports, we employ a customized application of *SMSOne* procured by the KCCA. This platform offers a tested and convenient way to manage messages from mobile phones and is currently being expanded by the KCCA to manage all types of incoming communication from citizens. The prompts involved questions co-designed by our research team and the KCCA waste management unit about various aspects of waste management, along with pre-defined response categories for most prompts. For example, we used the following prompt at various points throughout the study and implementation period:

> When did the rubbish truck last collect your rubbish? A) never B) more than two weeks ago C) last week D) this week

3.2.1 Baseline Monitoring Strategies

Prior to adopting the ICT-platform, citizens in Kampala communicated with KCCA staff about waste management through toll-free phone lines, an SMS shortcode, and social media websites. KCCA frontline staff or Client Care Officers (CCOs) were responsible for processing information received through these channels. Once processed, input from citizens was relayed to the appropriate Supervisor within the KCCA, who then decided on how to address the problem at hand. Following her evaluation of the incoming information, the Supervisor assigns staff from within the WMU – typically the Solid Waste Officer in charge of the area from which the comment originated – to resolve any issues detailed in comments received from the public. After investigating and resolving the issue at hand, the Solid Waste Officer reports back to the Supervisor on any actions carried out.

The KCCA has additional sources of information for all zones around Kampala. First, the KCCA uses information from local leaders, such as parish councillors and zone chairpersons. These are political representatives elected by the residents in local areas of the city. These local politicians regularly visit City Hall and bring complaints from the residents in their areas. Second, the KCCA had 200 informal "scouts" deployed across the city. Scouts operate in specific areas of the city and report to the Solid Waste Officer in charge of the area, including directed monitoring of areas of interest to managers.

3.2.2 New Information through Citizen Monitoring

The introduction of the ICT-platform altered the flow of information from citizen to the KCCA. First, our research team – not the KCCA Client Care Officers – was responsible for cleaning raw reports from citizen monitors.[1] Once cleaned, our research team would then relay all data from citizen monitors to the KCCA in a spreadsheet format, as well as summary documents conveying additional information on the status of waste service delivery within each zone.[2] The goal was to experiment with formatting and delivery in a way that work for the KCCA, before automating the process. During the longer period when the platform operated, recruited citizen monitors submitted a total of 24,720 verified and on-topic reports, 17,520 of which were sent to the KCCA during the study period reported here (the earlier phases of the project focused on motivating reporters to send reports, see Buntaine, Nielson and Skaggs (2017)).

Second, interactions between staff within the waste management unit became less-linear under the citizen monitoring SMS-platform. Instead of directly evaluating the incoming information, the Supervisor relied on a staff member to process the information from citizen monitors relayed in the zone-wise reports. Then, the staff member would pass on these data or reports to the appropriate Solid Waste Officer in the five Kampala divisions, who would subsequently draft Action Plans to address any pertinent problems.[3] Once complete, Solid Waste Officers would submit their Action Plans to the Supervisor for review. Given the Supervisor's approval, Solid Waste Officers would execute the activities detailed in the Action Plan and report back to the Supervisor with the results. The launch of the platform thus entailed a significant expansion of scope and effort in responding to citizens.

[1]Our research focused specifically on cleaning responses that did not conform to the structured response categories specified in the prompts. For example, our research team would recode a response of "never" to "A" for the prompt specified above.

[2]Our research team processed and relayed summary reports and data based on the KCCA's preferences and did not decide the format of or frequency with which the KCCA would receive reports compiled using data from the citizen monitoring platform.

[3]Note that the KCCA ultimately suspended the practice of drafting Action Plans after an internal staff reorganization, and in part due to the high volume of incoming reports.

3.2.3 Using Citizen Monitoring to Improve Waste Services

We predicted that structured, localized, consistent, and specific information from citizen monitors would improve KCCA waste services throughout Kampala in three specific ways. First, we expected that the ICT platform would help minimize the basic information problems preventing the KCCA from widely distributing waste services throughout Kampala.[4] Prior to the platform's launch, the KCCA lacked a method to collect data from a broad base of citizens, and relied on information from informally employed "Scouts" and administrative records on waste collection.[5] Information on waste conditions, therefore, was limited to a subset easily-accessible areas. By sourcing structured data from citizen monitors throughout Kampala, we expected that citizen-monitoring via the SMS-platform would improve the reach of KCCA waste services.

Second, KCCA staff also believed that citizen monitoring would improve citizen satisfaction with KCCA services by offering citizens a consistent and centralized mode of engagement. Upon launching the program, KCCA staff indicated that the citizen monitoring program was designed to ensure citizens that their concerns mattered and were being addressed. When the KCCA was nationalized in 2008, it was given significant resources and a mandate to improve the satisfaction of Kampala residents with government, given that the city is a stronghold of opposition to the ruling party. Our partners considered building formal channels to respond to citizens at a broad scale an important step toward increasing resident satisfaction in line with their mandate.

Finally, KCCA staff anticipated that consistent information from citizen monitors would improve the KCCA's ability to fulfill its PPP mandate to monitor and regulate private contractors. Before launching the program, information asymmetries existed between KCCA staff and private contractors. Neither administrative records nor sporadic reports from Scouts sufficiently informed KCCA staff about the frequency, quality, and mode of waste services private contractors delivered. KCCA staff expected that reporting from citizens actively

[4] As one Waste Management Officer reported: "My area of supervision contains 23 parishes and over 200 zones. It is impossible for me to be in all those places at the same time. The citizen monitors enable me to keep tabs in those areas by keeping me up-to-date with what is going on" (see interview I).

[5] For a full description of how the KCCA uses multiple sources of information to design its waste services, please refer to [section of SI].

experiencing waste services would expand their ability to detect and punish shirking contractors, thereby improving waste conditions throughout Kampala.

3.3 Hypotheses

Our field experiment thus tests the following main hypothesis:

H1. Zones assigned to citizen monitoring will experience a larger decrease in solid waste accumulation in the piles measured at than zones assigned to control.

Pre-registered measures (from photographs and field measurements)

- Area of total waste accumulation (primary outcome)
- Area of unmanaged waste accumulation
- Amount of burning
- Amount of non-organic waste

Reductions in waste pile sizes would point to some form of intervention by KCCA, for two reasons. First, the clearance of a waste pile does not automatically result in the disappearance of the waste pile. In the absence of consistent truck visits, people still continue dumping waste in the location. Second, because rubbish trucks typically make stopovers at numerous waste pile locations, sometimes by the time it reaches a specific location, there is not enough space in the truck to accommodate the garbage in the waste pile. In such a scenario, the workers usually pick up part of the waste pile and leave the rest for another time. The bottom line is that even marginal reductions in waste pile size are indicators of improved waste services. Changes in treated zones that are additional to control zones would indicate impact of the citizen monitoring in particular.

3.4 Sample and Random Assignment

We randomly selected 200 zones (out of 755) in Kampala to form our experimental sample. We randomly selected an additional 50 zones to use as replacements for zones that were inaccessible to our enumerators, demolished at the time of enumeration, or for which at least two problematic waste piles could not be identified by residents of the zone at baseline. We assigned half of the experimental sample of zones to the citizen monitoring treatment using complete randomization, as indicated in Figure 1.

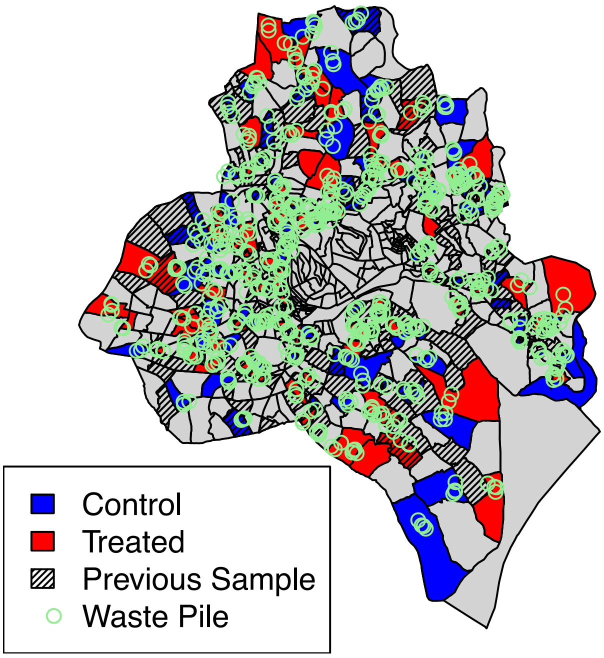


Figure 1: Experimental sample, including continuing reporting from previous phases.

We intended to select a sample for this field experiment that included entirely new zones without any previous reporting. Due to an indexing error, we selected a sample that overlapped with the samples from earlier phases of the project. This error was not caught until after baseline data had been collected. The resulting treatment still adds 50 new reporters to each of these zones, on average boosting the number of reports considerably. Additionally, our baseline measure takes into account any treatment effects that emerged as a function of citizen monitoring in earlier phases of the project.

The KCCA was blinded as to which zones with citizen reporting were being measured, since our design tests their ability to provide better oversight on the basis of citizen monitoring. The KCCA might re-direct attention to zones assigned to treatment apart from the information contained in reports if they were not blinded to the experimental sample. Thus, we continued to collect and pass along reports from hundreds of zones in previous phases.

Figure F1 displays balance and descriptive statistics for pre-treatment covariates, none of which are inconsistent with successful random assignment.

3.5 Treatment

The treatment is the delivery of zone-level reports about waste management to the KCCA. Each week, a prompt was sent out from among a list of questions that the KCCA waste management unit identified as important for management. Our research team then compiled the responses by zone and delivered a spreadsheet containing that information to staff at the waste management unit, as requested. While we observed several of the plans that the waste management unit made with these data, our research team was not involved in planning or delivering any responses to the information. Figure 2 tracks the study design.

3.6 Compliance with treatment

We choose our sample size to ensure that almost all zones would be covered by reports each week. Overall, we observed expected rates of on-topic and usable reports from citizens, averaging around a 10 percent response rate during the reporting period. This rate matches what was observed in previous phases of this project that investigated how citizens could be motivated to provide monitoring (Buntaine, Nielson and Skaggs, 2017). With this response rates, there was an average of 3-7 reports per week, from among 50 recruited reporters. In total, the KCCA received 17,520 verified and usable reports prior to the final waste audit. At one point our counterparts at the KCCA asked that we decrease the frequency of data deliveries, because they were overwhelmed by the volume and speed of data needing to be processed.

3.7 Internal consistency of reporting

The internal consistency of the reports sent in by citizens within zones varied, but was generally reasonable. Figure 4 displays the consistency of responses within zones on categorical

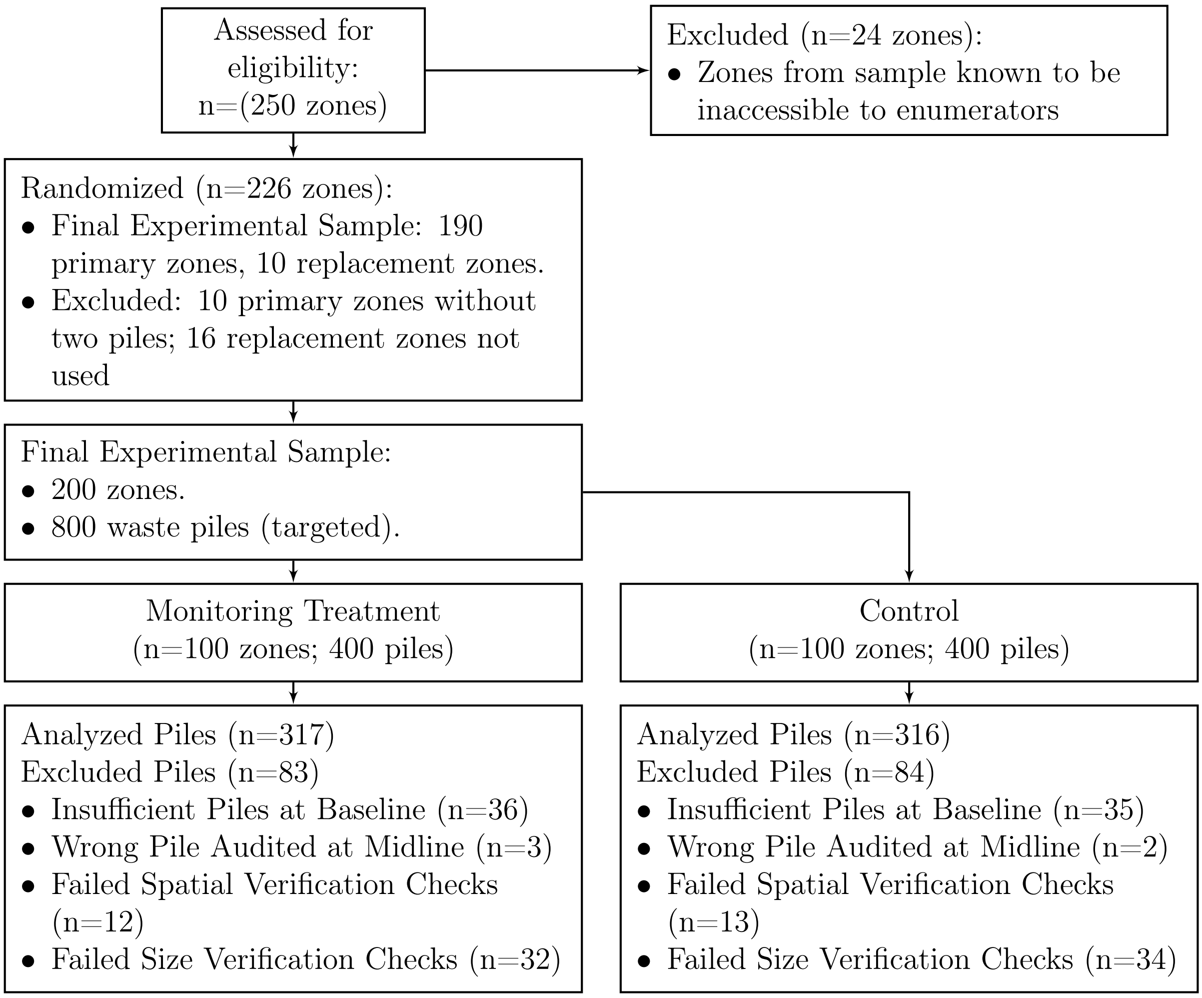


Figure 2: CONSORT diagram tracking study design.

measures that would indicate poor service quality to the KCCA.[6] On average, twenty-eight percent of citizen monitors recorded survey responses that deviated from the zone-level modal response – i.e., monitors indicating that service quality was poor when a majority of respondents in the zone indicated that service quality was acceptable, or vice versa.

[6]The categorical measure of poor service provision combined the following indicators: the frequency and accessibility of service provision, reported waste collector treatment of citizens outlined, and the amount of waste burning or litter.

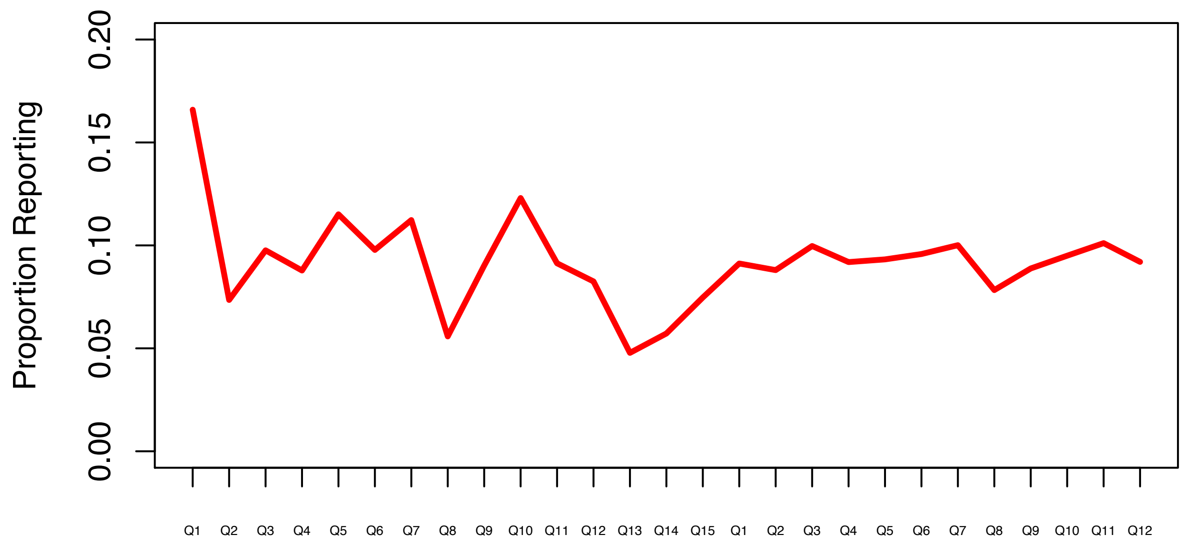


Figure 3: Experimental sample, including continuing reporting from previous phases.

3.8 Outcomes

The core of our measurement strategy involved a field-based audit of waste piles, since the presence of informal dumping is a direct outcome of waste collection services (See SI Appendix A for details). We went to each zone in the experimental sample and asked residents to show us up to four informal waste piles that were of greatest concern to them. We measured the spatial extent of and photographed these waste piles, recorded their locations by GPS, and mapped the easiest way to return to them for re-measurement. We also measured evidence of burning and the composition of waste piles. The core outcome of our field experiment is whether waste piles in treatment zones change more positively than those in control zones, comparing baseline pile sizes to re-measurements at 5 and 9 months post-treatment. Figure 5, Figure 6, and Figure 7 are representative examples of small, medium, and large waste piles, respectively.

3.9 Estimation

We test our hypotheses about pile size by regressing the size of piles on treatment status in the current phase, treatment status in previous phases, the pre-treatment measures of the pile size collected at baseline, and the following zone-level covariates: zone-level population,

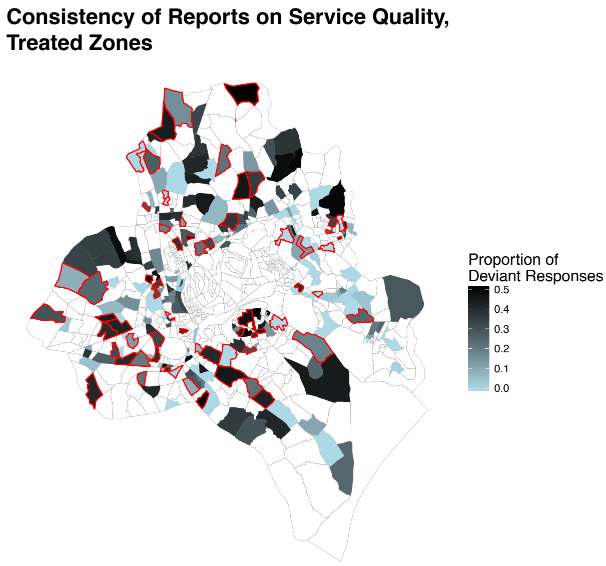


Figure 4: Consistency of reports from treated zones in the experimental sample. *Along a standardized measure of poor service provision, zones in red, indicated that KCCA service provision was poor on average. Zones with darker fills represent zones that inconsistently reported the quality of KCCA services, relative to the zone-level modal response.*

Figure 5: Small pile

Figure 6: Medium pile

Figure 7: Large pile

density of improved roads, and luminosity. We compute sharp null standard errors from the sampling distribution of the relevant parameter estimate, derived from randomization inference simulating 10,000 permutations of the complete randomization procedure that we used to assign treatment. The core estimating equation for measures with both baseline and endline values is Equation 1.

$$y_{ij,t=b+n} = \alpha + \tau M_j^+ + \gamma y_{ij,t=b} + \boldsymbol{\beta}\mathbf{X_j} + \nu_h + \epsilon_h \tag{1}$$

where y is the relevant size measure for pile i in zone j at time b baseline plus some follow-up period n, τ is the treatment effect of interest, M_j^+ is a binary indicator of treatment assigned at the zone-level j, γ is the parameter estimating the relationship of baseline size measure $y_{ij,t=b}$ to the follow-up outcome measure, $\boldsymbol{\beta}\mathbf{X_j}$ is the estimated adjustment for pre-treatment, zone-level covariates including the treatment status of zones during previous phases, ν_h is a fixed effect for division, and ϵ_h is an error term clustered at district, irrelevant in our case because all standard errors reported are sharp null standard errors. This estimation deviates from our pre-registered strategy in that it takes the pile, rather than the zone as the unit of analysis, which increases precision.

4 Results

4.1 Pile Sizes

We find no evidence that citizen monitoring reduces the number of existing waste piles in sampled zones (Figures 8, 9). Even when changing the definition of a cleaned pile to include sites for which all waste was collected into a single, transportable container, we find no significant difference in the proportion of waste piles cleaned up among treatment and control groups. Speaking directly to our primary hypothesis (H1), we find no evidence indicating that treated zones experienced greater reductions in waste accumulation than did control zones. Table 1 shows that we cannot rule out no effect of citizen monitoring on pile size and pile status with any confidence.

We observe similar results across other measures of pile characteristics. Citizen monitoring does not reduce the total area of unmanaged waste – estimated both as the level of waste storage and the level of waste organization – in treated zones as compared to control zones (Table 2). This effect is robust to various specifications of unmanaged waste area (see caption, Table 2).

Table 1: RI Results, Primary Dependent Variables (Cleaned)

	Pile Size	Pile Size	Pile Cleaned	Pile Cleaned	Pile Rank	Pile Rank
Audit	M1	M2	M1	M2	M1	M2
Treatment Effect	-2.54	1.14	-0.04	-0.02	-9.77	-7.35
Standard Error	2.67	3.46	0.03	0.04	15.89	14.71
p-value	0.18	0.6	0.12	0.31	0.26	0.3
N	621	621	621	621	621	621

Note: results calculated using cleaned waste pile size measurements.

Description of Dependent Variables

1. ***Pile Size:*** *change in waste pile area between the baseline audit and specified midline audit (m^2). Enumerators recorded waste pile dimensions at each audit. These dimensions were used to estimate waste pile area for each site in the sample. Waste pile size at the baseline audit was subtracted from waste pile size at each midline audit to produce an estimate of the change in waste pile size from the baseline audit to each midline audit.*
2. ***Pile Cleaned****: dummy variable indicating whether a waste pile was cleaned at the specified midline audit. Recorded values of 0 indicate that no pile was present at the given midline audit (e.g. the pile had been cleaned); recorded values of 1 indicate the opposite.*
3. ***Pile Rank****: change in waste pile size rank between the baseline audit and specified midline audit. Due to high variance in the recorded waste pile sizes, we perform a rank test comparing the ranked change in pile size between the baseline audit and each midline audit. Waste piles were ranked at each audit based on their size relative to other waste piles. Then, waste pile ranks at the baseline audit were subtracted from waste pile ranks at each midline audit to create a change in rank value.*

Table 2: RI Results, Secondary Dependent Variables (Cleaned)

	Storage	Storage	Organization	Organization	Burning	Burning
Variable Specification	A	B	A	B	A	B
Treatment Effect	3.48	3.58	4.87	5.09	1.99	1.72
Standard Error	2.15	2.25	2.62	2.62	1.51	1.67
p-value	0.95	0.95	0.97	0.98	0.9	0.83
N	621	621	621	621	621	621

Note: results calculated using cleaned waste pile size measurements.

Description of Dependent Variables

1. ***Storage:*** *change in uncontained waste pile area (m^2). At each midline audit, enumerators recorded how rubbish was stored in each waste pile. Responses ranged from "all of the rubbish is neatly contained with sacks or other containers" to "no rubbish is contained in sacks or containers." Each response was assigned a scalar between 0.0 and 1.0, which was multiplied against the recorded waste pile area to generate an estimate of uncontained waste pile area. Estimates were differenced to calculate the change in uncontained waste pile area.*

2. ***Organization:*** *change in unorganized waste pile area (m^2). At each midline audit, enumerators recorded the dispersion of rubbish in each waste pile. Responses ranged from "all of the rubbish is collected in a single pile" to "rubbish is spread all around [with] no evidence of the rubbish being organized." Each response was assigned a scalar between 0.0 and 1.0, which was multiplied against the recorded waste pile area to generate an estimate of unorganized waste pile area. Estimates from each midline were differenced to calculate the change in unorganized waste pile area.*

3. ***Burning:*** *change in burnt waste pile area. At each midline audit, enumerators recorded any evidence of burning they observed at each waste pile. Responses ranged from "no evidence of burning" to "more than half of the area of the rubbish pile contains evidence of burning." Each response was assigned a scalar between 0.0 and 1.0, which was multiplied against the recorded waste pile area to generate an estimate of burnt waste pile area. Estimates from each midline were differenced to calculate the change in burnt waste pile area.*

Variable specification A is less conservative than variable specification B. Specification A assigns a smaller scalar score to enumerator responses indicating less organization/storage and more burning. As a result, mean estimates of unmanaged, uncontained, or burnt pile area using specification A are smaller than similar estimates using specification B.

Figure 8: Dependent Variables, Midline 1

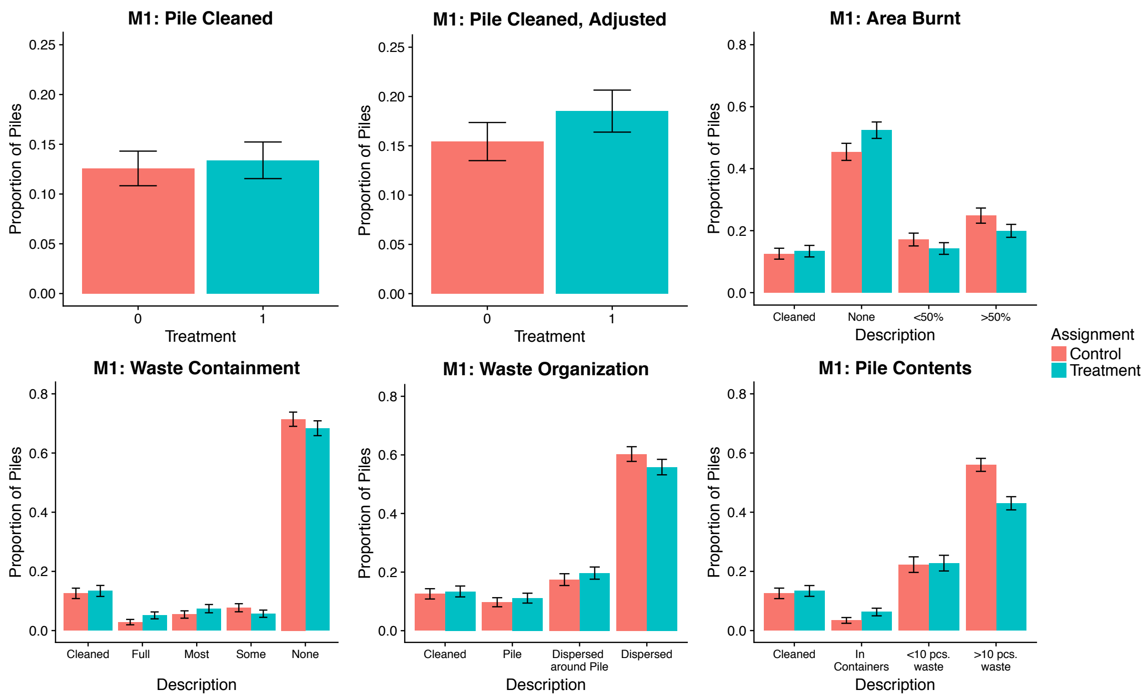
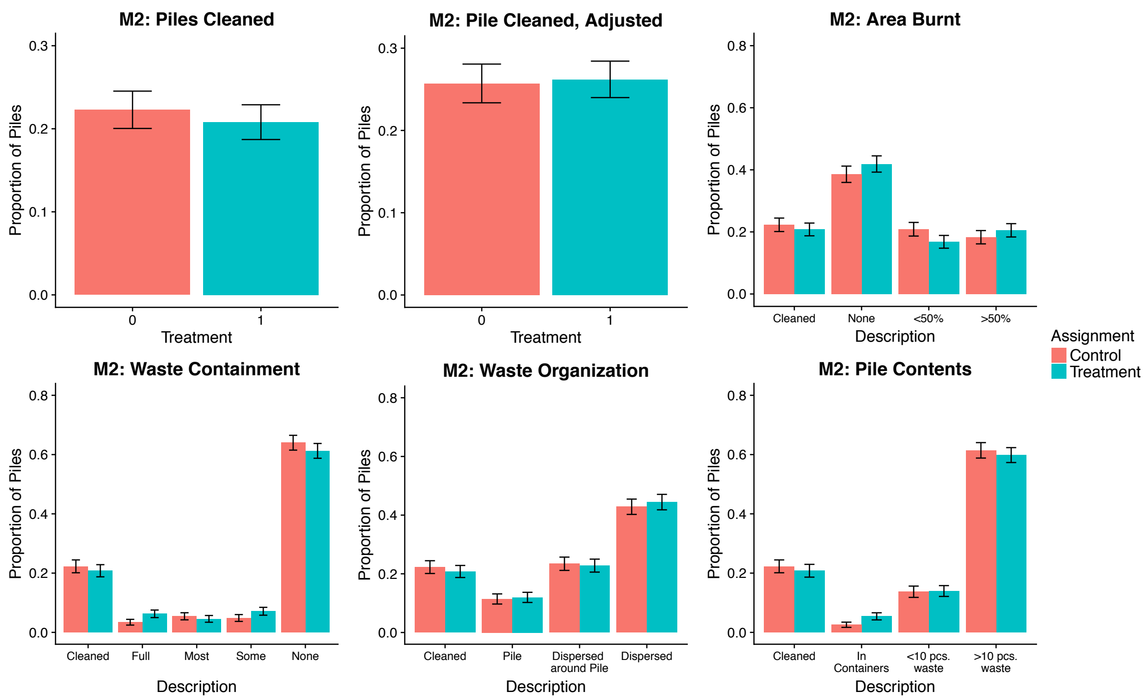


Figure 9: Dependent Variables, Midline 2



4.2 Pile Characteristics

While there is no significant effect of treatment on waste pile size and cleaned status, there is marginal evidence that citizen monitoring may have temporarily improved certain facets of the waste management. This effect is particularly pronounced in when comparing pile characteristics among treated and control zones following the first midline audit[7].

At the first post-treatment audit, there is an overall reduction in the amount of non-organic waste found per pile in treated zones. We see that the proportion of treated piles with greater than ten pieces of non-organic waste is significantly lower than the proportion of similar control piles (Figure 8; $te = -0.11$, $p<0.001$, B-Y corrected $p=0.019$). However, the amount of non-organic waste in treated piles increases by the second post-treatment audit (Figure 9, $te = -0.006$, $p=0.46$).

Data from the first post-treatment audit on waste burning indicates a similar temporary improvement. In the first midline audit, treated zones contained a larger proportion of piles with no evidence of burning (Figure 8; $te=0.07$, $p=0.04$). There was also a smaller proportion of piles with evidence of widespread burning in treated zones than in control zones ($te = -0.07$, $p=0.07$). However, both treatment effects attenuate for the second post-treatment audit (Figure 9, $p=0.17$ and 0.11, respectively).

Waste containment and pile organization saw a gradual but enduring improvement throughout both midline audits. While the difference in proportion of fully-contained piles – where all rubbish is stored in transportable sacks or containers – among treated and control zones was not significantly different from zero following midline one, treated zones following midline two contained a greater proportion of fully-contained piles than did control zones($te=0.03$, $p=0.058$). Reports on general pile descriptions from the second midline audit corroborate this effect. Following the second post-treatment audit, treated zones contained a larger proportion of piles with waste stored for transport than did non-treated zones (Figure 9; $te = -0.03$, $p=0.001$).

[7]However, table 2 shows that the hypothesized effect of treatment is undetectable when comparing the change in uncontained, unorganized, or burnt waste pile area from the first midline audit to the second midline audit.

5 Mechanisms and Heterogeneous Treatment Effects

5.1 Political Targeting

Several recent papers show that politicians often use public goods and services as a way to reward supporters in elections (Jablonski, 2014; Drazen and Eslava, 2010; Baldwin, 2013; Briggs, 2012). Using public goods in this way is often an effective strategy to build political support. In the setting of our study, the National Resistance Movement (NRM) is the ruling party nationally, but faces generally low levels of political support within Kampala. In 2011, aiming to reverse the trend of entrenched opposition within the capital city, the municipal government was nationalized and responsibility for services transfered from the elected city council to the KCCA. Thus, the KCCA to use there discretion is targeted to reward areas of the city that vote for NRM candidates, as compared to opposition or independent candidates.

We test for this possibility by examining both the baseline amount of waste accumulation and whether the reporting treatment was more effective where the winning candidate in 2016 division elections for the parish constituency was a member of the NRM ruling party. As displayed in Table F3, we fail to find evidence that either the status or change in waste pile sizes is conditional on the party of the Division councillor, ruling out the possibility that political targeting is driving the allocation of effort.

5.2 Reporting Rates and Message Content

Vocal stakeholders often receive the most attention from public service providers. Under public pressure, KCCA officials might respond disproportionately to zones that either frequently or consistently report shortfalls in waste service provision. Alternatively, KCCA officials might respond disproportionately to a zone where some citizen-monitors express severe dissatisfaction with KCCA services.

Using the content of reports collected prior to the first post-treatment audit, we test these hypotheses. We take a zone-level count of responses to examine if frequent reporting improves waste service provision. We use the content of reports to create zone-level measures of reporting consistency, dissatisfaction services, and waste problem severity. On all measures, we do not find that the amount or content of reports affects waste pile size at the conventional level of statistical significance (Tables F4, F5, F6, and F7).[8]

These results largely rule out the possibility that the weak main effect of treatment is a consequence of a heterogeneous treatment effects. Zone-level reporting frequency and consistency, dissatisfaction with KCCA services, and waste problem severity each fail to moderate the effect of treatment.

6 Participant Observations and Staff Interviews

Our team was embedded in the KCCA waste management unit for close to one year. During this time, we interacted with a variety of KCCA staff members, from managers to frontline staff providing waste services. We had access to and reviewed KCCA documents, participated in KCCA meetings, and regularly observed interactions between the KCCA and its stakeholders. Following the second midline audit, we conducted in-depth interviews with all individuals who interacted with the SMS-reporting platform.

This section summarizes what we learned from those experiences. We first detail how KCCA staff used the platform to act on citizen-sourced information regarding service shortfalls throughout Kampala. We then use information surfaced in our interviews to shed light on the shortcomings of citizen-sourced information in this context; namely, the inconsistency of incoming citizen reports and the platform's perceived cost relative to other monitoring mechanisms. Finally, we show how an unexpected organizational shift within the KCCA culminated in the platform's abandonment following the second midline audit.

[8] We operationalize "severity of waste management problems" in Table F5 using reports from an item asking citizen-monitors to report if a rubbish-collection truck visited their neighborhood. Possible responses include: yes, no, don't know. The latter two responses were coded as indicative of severe waste management problems. Following the logic outlined above, one would expect the KCCA to deploy trucks disproportionately to zones reporting that they had not received pick-up services recently. We additionally operationalize waste management problem severity using citizen-monitor reports commenting on rubbish burning, litter and illegal piles, rubbish spilling from KCCA trucks, and mistreatment by KCCA waste collectors.

6.1 Complications in Using Citizen Monitoring

As our experimental results confirm, the expected effects of citizen monitoring on waste service delivery in Kampala failed to materialize. Zones with citizen monitors showed only marginal improvements in certain pile characteristics, and citizen monitoring did not detectably increase waste pile clearing in treated zones.

From our participant observations and staff interviews, we have identified two reasons that may explain the observed null effect and the KCCA's transition away from citizen monitoring: (1) the inconsistency of information from citizen monitor reports and (2) the high cost of citizen monitoring. Below, we describe each reason in detail and make comparisons between citizen monitoring and alternative monitoring mechanisms the KCCA used during the study period. We then show how both the perceived high cost and low consistency of citizen monitoring influenced the KCCA's decision to abandon the SMS-platform following an unexpected staff restructuring between the first and second midline audit.

6.1.1 Inconsistent Reporting and Verification

One component that managers considered when assessing citizen monitoring was its ability to produce consistent and reliable information on zone-level waste conditions. While citizen monitoring did increase the flow of information to KCCA staff, this information could only improve KCCA service delivery to the extent that it signaled a persistent waste problems in a verifiable manner. In practice, though, information from citizen monitors proved to be both inconsistent and unverifiable.

Among the pool of citizen monitors in a given zone, a quarter would report waste conditions that contradicted the responses of other monitors (see Figure 4 for a visual representation). For example, some citizen monitors would report that a garbage truck visited the area within the past week, while other citizen monitors from the same area would report that the garbage truck had not appeared in more than a month. There were also instances in which the same citizen monitor would send consecutive responses to the same prompt that were contradictory.

As a result, responses from citizen monitors often confused the KCCA staff responsible for planning interventions. Resource- and time-constraints frequently prevented KCCA staff from verifying inconsistent reporting with citizen monitors.[9] Unverifiable reports were simply discarded. In the words of the acting WMU Supervisor who came on after the platform had been operating for several months: "The data is not useful because its authenticity or accuracy cannot be verified" (see interview J).

Resolving inconsistencies in citizen monitor reports additionally slowed the KCCA's delivery of waste services. The time required to sift through, verify, and apply incoming data delayed KCCA responses to citizen concerns. Some KCCA staff reported that it took between one and two working days to transform the data received in a spreadsheet information into actionable information.

Conversely, information from KCCA Scouts was consistent and required no verification. Scouts are KCCA employees with no overt motivations to misrepresent waste conditions. Scouts also follow strict reporting parameters when monitoring waste conditions, making incoming information more consistent and, therefore, actionable. When Scouts do file contradictory reports, resolving these inconsistencies is far less-time consuming. KCCA managers reach out to Scouts via WhatsApp, where Scouts can easily upload pictures of waste conditions to corroborate their reports. Given the overall consistency of Scout-based information and the ease of report verification, KCCA staff felt more comfortable using Scout reports to inform the design of interventions than they did using information from citizen monitors.

6.1.2 High Costs

Another reason the KCCA abandoned the citizen monitoring program was its unexpectedly high operating costs. In the initial phases of our partnership, our research team covered the costs of the platform (see Buntaine, Nielson and Skaggs (2017)). Once our research team handed the platform over to the KCCA in the period reported here, it became responsible

[9] The managers initially employed a number of strategies to cope with the inconsistency of incoming citizen reports. These strategies included (1) personally contacting individual monitors who sent consecutive contradictory reports; (2) following up with citizens where reporting inconsistency was high to get additional input on local waste conditions; (3) utilizing staff knowledge of those areas to interpret the information from the citizens; and (4) following up with other stake-holders in the service provision process (e.g. speaking with private contractors operating in a zone with inconsistent reporting). The verification process proved costly and time-consuming from a management perspective, given the small staff size at the WMU and the extremely demanding mandate to deliver services to millions of residents.

for shouldering the program's costs. Immediately, KCCA staff began questioning the cost-effectiveness of citizen monitoring.

Many concerns about the cost-effectiveness of citizen monitoring stemmed from the program's low response rates. On most occasions, the KCCA received reports from no more than 12 percent of citizen monitors enrolled in the program (see Figure 3). Given that the KCCA was billed for every SMS it sent to citizen monitors – unconditional on the monitor's response – many KCCA staff felt that an overwhelming portion of the program's budget was being wasted. This sentiment was not lost on high-ranking officials in the KCCA. During a presentation of the Phase I and Phase II results, the Deputy Executive Director of the KCCA criticized the citizen monitoring program for producing limited information on waste conditions at such a high cost.

Comparing the monthly cost of citizen monitoring to monthly cost of other available monitoring strategies helps illustrate staff concerns about costs. One engagement cycle of the citizen monitoring platform cost the KCCA UGX 915,000 ($254 USD), and on average yielded 750 responses from citizen monitors. Over the course of a month, the KCCA would go through at least four engagement cycles. Without accounting for the cost of processing incoming information, the monthly cost of citizen monitoring was approximately 3,660,000 UGX ($1016 USD).

Complete funding for the 72-person team of KCCA Solid Waste Scouts for a month similarly cost 3,660,000 UGX ($1016 USD). However, incoming information from Scouts seldom required additional verification or processing, shielding the KCCA from the downstream costs it incurred in processing the citizen monitoring.

Thus, while citizen monitoring gave the KCCA access to a broader base of information, the opportunity cost of its use was too high. Low response rates coupled with the need to process downstream data rendered citizen monitoring a costly alternative relative to using KCCA scouts, both practically and economically. A quote from the current WMU Supervisor summarizes the comparison bluntly: "For me, these messages are very expensive for nothing. That is why I was saying, 'Why don't we buy the scouts airtime and communicate on WhatsApp?" (see interview J).

6.1.3 KCCA Restructuring: Transitioning from Citizen Monitoring

Between May and July 2017, the KCCA underwent a staff restructuring process to, in the words of the KCCA's Executive Director, "improv[e] service delivery." In total, around 120 people were transferred, fired, or newly-hired across all branches of the KCCA. The Waste Management Unit (WMU) was not exempt from this process. A number of staff were moved into and out of the unit, and those that remained in unit were often reassigned to different roles. The most radical change occurred in the unit's leadership with the introduction of a new Supervisor.

The transition in leadership of the WMU surfaced many of the flaws of the citizen monitoring program. Prior to her departure, the outgoing Supervisor championed the ICT and emphasized that citizen engagement in and of itself was a means to improve waste conditions throughout Kampala. In her words: "As KCCA, one of the core values of the institution is client care. And client care cannot be actualized if the client is not satisfied. Public service delivery is directed towards the clients. If the clients send any feedback, it is incumbent upon KCCA to respond to this feedback (and if need be re-align its priorities/operations/services in line with the client feedback)" (see interview G). She embraced the challenges associated with using citizen-sourced information and perceived the payoffs of citizen engagement to be greater than the additional costs the WMU incurred from using the platform.

Conversely, the incoming Supervisor played no role in launching the citizen monitoring program and did not trust citizen monitors. The inconsistency of citizen reports, overt criticism of the KCCA in reports, and high relative costs of citizen monitoring were interpreted as barriers to effective service provision. He felt as if citizen monitors were maliciously misrepresenting local waste conditions: "I find it difficult to act on such messages when some even which are insulting. I cannot tell whether the message which is sent is genuine. Where somebody is not being paid, even if they give you wrong information, how do you track?" (see interview J).

Following persistent technical glitches and in light of these criticisms, the new Supervisor decided to discontinue the citizen monitoring program in favor of expanding the KCCA Scout program. He nearly doubled the size of Scout program at the beginning of his tenure

– increasing the size of the program to 200 employees – and modified Scouts' roles in the WMU. In addition to investigating illegal waste management practices, the new Supervisor insisted that Scouts begin monitoring general waste conditions and service delivery, effectively subsuming the role of citizen monitors.

7 Framework

In light of these experiences, we offer a framework for considering when citizen monitoring can help to improve the provision of public services. We aim to generalize the problems encountered by our partners in a way that can offer guidance to other public managers.

7.1 Theoretical Preliminaries

Consider a public manager that is under a budget constraint C, such that her spending on improvements to public services can be divided between any number of tasks j such that $C = \sum c_j$. Her goal is to allocate this budget so that she maximizes improvements to public services. Her main problem is uncertainty about how to best allocate her resources to maximize payoffs, which we denote as $\sum \Theta_j$, where Θ_j is the payoff of task j that is stochastic in each period from an underlying probability.

For each task j, we assume the public manager has a belief about the underlying distribution of payoffs $p(\theta_j)$ that will be achieved when allocating budget toward that task. Without any additional monitoring, the manager chooses in order the tasks that have the highest payoff relative to the budget outlays that are required to accomplish them, until her budget is used up.

The manager can also spend resources on monitoring m_j to acquire better information about the realized values Θ_j drawn from $p(\theta_j)$ in a period of effort. But the manager is also uncertain about the value of new information relative to using the underlying beliefs about the cost-effectiveness of actions. We assume the manager spends her entire budget constraint C in any scenario, such that her objective function is, with tasks $(1')$ through (k') being those chosen after monitoring:

$$U_m = \sum_{(1')}^{(k')} \Theta_j - \sum^{j} m_j \tag{2}$$

There are two primary problems that the manager must confront when attempting to maximize this objective function. She might pick the wrong tasks j for a given period of effort. And she might spend part of her budget on monitoring m_j that does not help her make better decisions, leaving less budget to actually deliver public services.

Monitoring technologies differ both in their costs and in their ability to reveal information about the realization of Θ_j such that they help make better decisions. It only makes sense to do any monitoring if the cost of acquiring information about the realization of Θ_j is less than the expected increase in payoffs attained from choosing tasks more effectively.

7.2 Allocating Effort and Monitoring

To make the problem more tractable, consider a manager with a set of tasks j that all have the same binary payoff structure $\Theta_j \in [0, 1]$, which in each instance is drawn from a Bernoulli process where there is an underlying true probability $p(\theta_j)$, which is known to the manager.

7.2.1 No monitoring

Consider the manager who does not have any access to monitoring technology, but must allocate effort toward tasks j when there exists a realized $\Theta_j \in [0, 1]$ drawn from the true distribution of $p(\theta_j)$. The costs of carrying out the tasks are fixed whether or not the payoff for each task is realized. For simplicity of exposition, we also assume that the cost of tasks j are constant.

The goal of the manager is to allocate tasks j such that $\sum_{(1)}^{(k)} \Theta_j$ is maximized in each period. When monitoring is not available that reveals information about the realizations of Θ_j, the manager will simply be guided by their prior beliefs and allocate effort to the tasks that have the highest payoffs in expectation. Specifically, she will order her beliefs about each task such that:

$$p(\theta_j)_{(1)} > p(\theta_j)_{(2)} > ... > p(\theta_j)_{(k)} > ... > p(\theta_j)_{(z)} \tag{3}$$

The manager will then choose tasks (1) through (k), where $U_m = \sum_{(1)}^{(k)} \Theta_j$. In this case, the payoff she expects is simply the sum of all probabilities through task (k):

$$E\left[\sum_{(1)}^{(k)} \Theta_j\right] = \sum_{(1)}^{(k)} p(\theta_j) \tag{4}$$

Without monitoring, the manager is likely to allocate effort to optimize between payoffs of effort in a particular period and information that helps to make decisions more effectively in future periods. For example, the manager might use *randomized probability matching*, where effort is allocated to each task according to the probability that it is among the set of highest payoff tasks given the prior beliefs at each time period (Scott, 2010). Yet, even this strategy can be improved under certain conditions by collecting information on specific instances of θ_j through monitoring.

7.2.2 Perfect monitoring

If monitoring can be added that reveals information about the particular realization of Θ_j drawn from θ_j prior to the allocation of effort, then it becomes possible to make better choices. Consider first that the manager can pay some monitoring cost m_j that will reveal Θ_j, subject to the budget constraint such that $\sum m_j < C$. We assume that m_j is strictly less than c_j, otherwise it would never make sense to pay for monitoring.

The problem for the manager who has the option of perfect monitoring is when and where to spend resources on monitoring, leaving fewer resources for carrying out the public service tasks. If there are many tasks j, the costs of monitoring can be large and quickly consume her budget. The severity of the monitoring versus effort trade-off will depend on the relative costs of monitoring and action related to the public service; when monitoring is inexpensive the trade-off between searching for new information and acting is small, but when monitoring is expensive the trade-off is large.

Monitoring leads to increased payoffs only when it changes the allocation of effort. In particular, the payoff to monitoring will be exactly equal to the number of $\Theta_{(k)} = 0$ tasks avoided and replaced with tasks for which $\Theta_{(k')} = 1$ from Eq. 2. From changes in the allocation of effort between the baseline where no monitoring information is available ((1)

through (k)), to the allocation of effort following monitoring ($(1')$ through (k')), the value of monitoring can be conceptually described. The manager will choose to pay for monitoring whenever there exists an additional unit m_j for which the value of monitoring $V(m_j)$ is positive:

$$V(m_j) = \sum_{(1')}^{(k')} \Theta_j - \sum m_j - \sum_{(1)}^{(k)} p(\Theta_j) \tag{5}$$

The problem for the manager is that this value cannot be solved because the payoffs to monitoring are unknown prior to the application of monitoring. This is an extremely complex problem that defies simple analytical solutions, because the payoffs to m_j are not independent of other decisions m_{-j}. A complete set of monitoring decisions chosen will determine the ordering of posterior beliefs that will drive the actual allocation decision. Monitoring only benefits managers when it changes decisions about the allocation of effort to tasks, otherwise it is pure cost.

This problem can be simplified based on the idea of regret, as applied to individual parts of the ordering of the prior beliefs given by Eq. 3. For each $p(\theta_j)$ where $(j) \leq (k)$, the expected regret of allocating effort is equal to the expected probability that effort will lead to zero payoff:

$$E[R_j] = c_j * (1 - p(\theta_j)) \tag{6}$$

It makes sense it engage in perfect monitoring in sequence for each task j whenever the cost of monitoring is less than the expected value of the regret, that is where $m_j < R_j$, until there exists an ordering of beliefs given by Eq. 3 such that it no longer makes sense to pay for monitoring.[10] The expected payoff of monitoring within this dynamic search process will depend on a number of factors, including the sequence of underlying probabilities $p(\theta)$, the cost of monitoring m_j, the benefits that can be attained through effort Θ_j, and the budget constraint C. Overall, however, evaluating monitoring costs in terms of expected regret helps make sense of when it is advantageous to invest in perfect monitoring. When monitoring

[10] To avoid discontinuities in effort based on the budget constraint, we assume that partial effort can be applied to task (k'), with the payoff equal to zero or the proportion of full effort exerted.

costs are high, more regret will be tolerated and vice versa.

7.2.3 Imperfect monitoring

Technologies for citizen monitoring might drive down the costs of monitoring, broaden the number of tasks that are monitored, and reduce the trade-off between monitoring and effort. Indeed, the push towards more bottom-up and citizen-driven monitoring systems is largely premised on the idea of preserving more budget for the actual delivery of public services. Additionally, there may be political or reputational benefits for being responsive to citizen concerns (McCubbins and Schwartz, 1984).

Monitoring done by citizens is imperfect, however, because it comes in the form of a noisy signal about a particular draw of Θ_j. In the context of citizen monitoring, people who report on the value of Θ_j might disagree, have faulty observations, or provide misinformation purposefully, all of which will harm the ability of a manager to draw clear inferences about the true realized value of Θ_j, which would help to make decisions about effort.

With imperfect monitoring, the manager has to extract signal from noisy, citizen-sourced data. Imperfect monitoring is time-bound and only provides information on a single instance of Θ_j. The value of citizen monitoring for decision-making is directly related to the consistency and amount of the information provided by citizens. Consider that the manager must calculate $p(\Theta_j = 1|y_j)$ for each task j, where y_j is a mix of reports containing binary information about whether $\Theta_j = 1$. By Bayes rule, this belief about the particular realization of Θ given y_j can be computed as:

$$p(\Theta_j = 1|y_j) = \frac{p(y_j|\Theta_j = 1)p(\Theta_j = 1)}{p(y_j|\Theta_j = 1)p(\Theta_j = 1) + p(y_j|\Theta_j = 0)p(\Theta_j = 0)} \tag{7}$$

One additional assumption is needed to compute a posterior probability $p(\Theta_j = 1|y_j)$: the proportion of reports that are incorrect. We assume that the manager can estimate the proportion of incorrect reports $w \in [0, 1]$, by examining globally the proportion of reports that deviate from the modal value. An important assumption at this point is that the manager will be able to determine that the reports tend toward being correct or incorrect on average, that if 30 percent of reports deviate from the modal response, this indicates that

30 percent of reports are incorrect, rather than 70 percent of reports being incorrect. With this assumption, the proportion of reports that deviate from the modal value will transform $p(\Theta_j)$ into a probability model for realized reports.

$$p(\Theta_j = 1|y_j) = \frac{\binom{n}{k}(\Theta - w)^k w^{n-k} p(\theta)}{\binom{n}{k}(\Theta - w)^k w^{n-k} p(\theta) + \binom{n}{k} w^k (\Theta - w)^{n-k}(1 - p(\theta))} \tag{8}$$

From this posterior belief, the value of imperfect monitoring (I) alone can be expressed by the expected amount of regret that is avoided as compared to acting only on prior beliefs:

$$V(I) = \sum_{(1')}^{(k')} p(\Theta_j = 1|y_j) - c_I - \sum_{(1)}^{(k)} p(\theta_j) \tag{9}$$

Since regret is a directly a function of the probability of making the wrong decision when allocating effort (see Eq. 6), the value of imperfect monitoring will be realized when it increases the confidence in allocating effort among the chosen tasks, particularly such that further perfect monitoring can be avoided. This is equivalent to decreasing the amount of expected regret among the actions that are chosen.

7.3 Illustrations and Predictions

Using this framework, it is possible to illustrate the conditions under which imperfect monitoring is predicted to lead to improved public services. We conduct a small simulation loosely fitted to the conditions in our field study. In particular, we assume that 28% of reports deviate from the zone modal value and assume also that the number of reports per task is a random variable in each period with 10 percent of reports active from among 50 recruited reporters. We assume that the manager has the budget to implement 20 tasks out of a total of 100 possible tasks. We assume that the manager has beliefs for each period of effort $p(\theta_j)$ drawn from a random uniform distribution [0.2,0.8].

Under these parameters, Figure 10 shows that imperfect monitoring can avoid regret and improve the number of tasks successfully completed, as long as the cost of imperfect monitoring (e.g., data collection, processing, planning) is less than 25 percent of the total implementation budget each period. While imperfect monitoring does not help make all decisions more certain, it has the potential to make enough decisions more certain to improve the use of limited budgets.

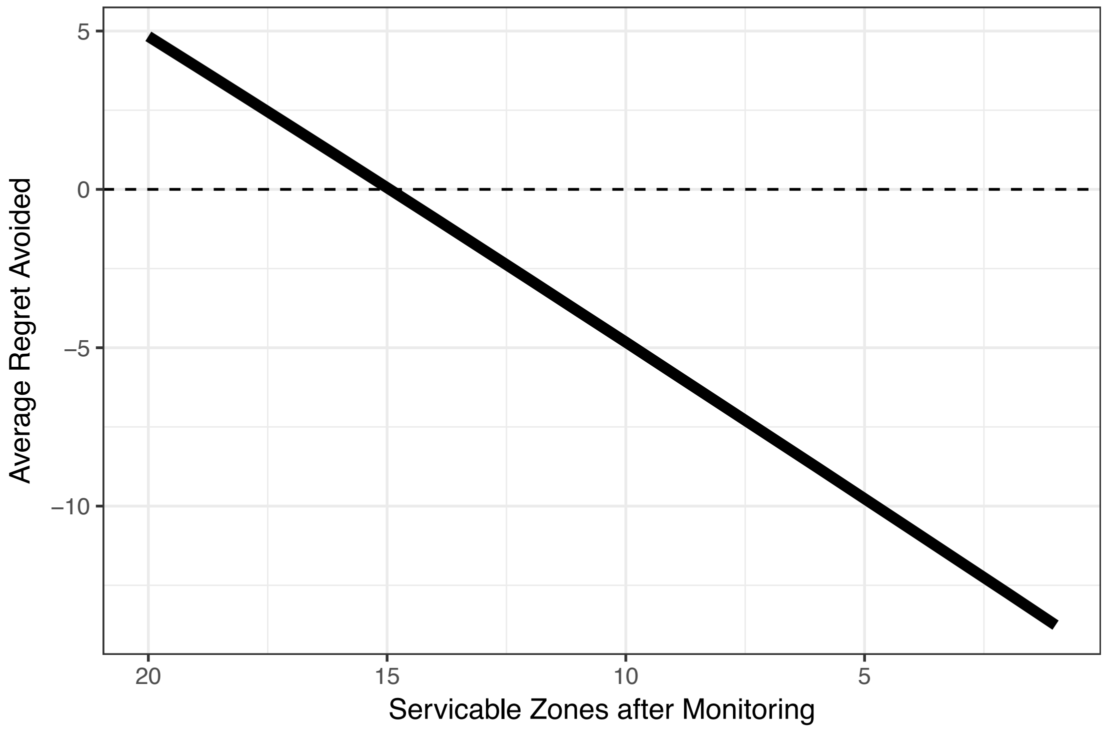


Figure 10: Regret avoided by the number of zones serviceable after monitoring.

This result highlights that even noisy and inconsistent data has value, if the authority in question has an ability to process and respond to the incoming data in ways that identify at least several problem areas that need services with a high degree of certainty, when uncertainty prior to monitoring is significant and the number of potential tasks is large. Of course, it may be particularly challenging to arrive at this kind of conclusion when the volume of information is very large and the cost of processing increases with volume (Hiltz and Plotnick, 2013). If the cost of processing bottom-up information larger than the reduction in regret, it does not make sense to use imperfect monitoring in the management of public services.

There are a number of assumptions underlying this framework that if relaxed reveal the potential for other functions of citizen monitoring. Bottom-up, citizen monitoring is a form of public pressure and reveals how much the public is tracking the performance of the government. This monitoring might cause the government unit receiving reports to work more efficiently (lowering the cost of completing tasks) or even to allocate more effort to tasks within the larger context of budgeting decisions (see Grossman and Michelitch, 2016).

8 Conclusion

Communication technologies create new spaces for governments and citizens to come together to improve the delivery of public services and they offer the potential for governance and public management in the decades ahead. Yet, we lack solid evidence that the tools offered by emerging communication technologies can translated into the improved management of public resources. On the one hand, these tools have the direct potential to solve information problems for public agencies that deliver frontline services. After all, citizens directly experience these services or the lack of these services as part of their daily lives and have information on their experience. On the other hand, integrating high-frequency, high-volume, and hyper-local data streams into the active management of public services requires considerable commitment and capacity on the part of public managers. It also requires the information to be relatively inexpensive compared to alternatives and to be of a quality and consistency that is useful for decision-making.

We fail to find improvements in the amount of waste accumulation in zones assigned to citizen reporting. We find some indications of promising results that later disappear, but at most the evidence is marginal that any improvements to solid waste services were made based on citizen monitoring.

The results of this study point out the many challenges of moving from citizen reporting to improved public services, like waste management. Citizen monitoring of public services is noisy, inconsistent, and costly to process. It can be frustrating for managers to follow-up on information when clarifications are needed prior to acting, since volunteer reporters are not at the disposal of managers. Additionally, the volume of data can be overwhelming, with managers scarcely having enough time to process one period of data before more data comes in requiring processing and action. Indeed, the waste management team even stopped producing weekly action plans in response to the data, because they felt they did not have enough time to act on each one and were spending more effort processing data, as compared to actually responding to the information that they received.

Overall, citizen-sourced data is promising because of its potential to expand the scope of monitoring, while at the same time offering localized and timely data. We find that this promise is likely to be overstated because of the complexities involved in processing citizen-source data and the inconsistencies that are inherent to citizen reporting. We frame the conditions under which citizen reporting will be helpful, that is when the data is produces is easy to process, consistent, low-cost relative to alternatives, and is brought to bear on decisions with high degrees of uncertainty. These conditions are unlikely to exist across a range of realistic circumstances.

References

Baldwin, Kate. 2013. "Why vote with the chief? Political connections and public goods provision in Zambia." *American Journal of Political Science* 57(4):794–809.

Banerjee, Abhijit, Angus Deaton and Esther Duflo. 2004. "Health, health care, and economic development: Wealth, health, and health services in rural Rajasthan." *The American economic review* 94(2):326–330.

Banerjee, Abhijit V, Rukmini Banerji, Esther Duflo, Rachel Glennerster and Stuti Khemani. 2010. "Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India." *American Economic Journal: Economic Policy* 2(1):1–30.

Bertot, John C, Paul T Jaeger and Justin M Grimes. 2010. "Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies." *Government information quarterly* 27(3):264–271.

Bhuiyan, Shahjahan H. 2010. "A crisis in governance: Urban solid waste management in Bangladesh." *Habitat International* 34(1):125–133.

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.

Briggs, Ryan C. 2012. "Electrifying the base? Aid and incumbent advantage in Ghana." *The Journal of Modern African Studies* 50(4):603–624.

Bryer, Thomas A. 2009. "Explaining Responsiveness in Collaboration: Administrators and Citizens Role Perception." *Public Administration Review* pp. 271–283.

Buntaine, Mark T, Daniel L Nielson and Jacob T Skaggs. 2017. Escaping the Valley of Disengagement: Two Field Experiments on Motivating Citizens to Monitor Public Goods. Technical Report 41 AidData.

Charalabidis, Yannis, Anna Triantafillou, Vangelis Karkaletsis and Euripidis Loukis. 2012. "Public policy formulation through non moderated crowdsourcing in social media." *Electronic participation* pp. 156–169.

Connors, John Patrick, Shufei Lei and Maggi Kelly. 2012. "Citizen science in the age of neogeography: Utilizing volunteered geographic information for environmental monitoring." *Annals of the Association of American Geographers* 102(6):1267–1289.

de Figueiredo Jr, Rui JP, Pablo T Spiller and Santiago Urbiztondo. 1999. "An informational perspective on administrative procedures." *Journal of Law, Economics, and Organization* 15(1):283–305.

Drazen, Allan and Marcela Eslava. 2010. "Electoral manipulation via voter-friendly spending: Theory and evidence." *Journal of development economics* 92(1):39–52.

Evans, Angela M and Adriana Campos. 2013. "Open government initiatives: Challenges of citizen participation." *Journal of Policy Analysis and Management* 32(1):172–185.

Grossman, Guy and Kristin Michelitch. 2016. Information Dissemination and Politician Performance between Elections. A Field Experiment in Uganda. Technical report Working Paper.

Grossman, Guy, Macartan Humphreys and Gabriella Sacramone-Lutz. 2014. ""I wld like u WMP to extend electricity 2 our village": On Information Technology and Interest Articulation." *American Political Science Review* 108(3):688–705.

Grossman, Guy, Melina Platas and Jonathan Rodden. 2017. Crowdsourcing Accountability: ICT for Service Delivery. Technical report.
URL: *https://ssrn.com/abstract=3027923*

Hansen, Morten Balle and Iben Norup. 2017. "Leading the Implementation of ICT Innovations." *Public Administration Review* 77(6):851–860.

Hiltz, Starr Roxanne and Linda Plotnick. 2013. Dealing with information overload when using social media for emergency management: Emerging solutions. In *ISCRAM*.

Jablonski, Ryan S. 2014. "How aid targets votes: the impact of electoral incentives on foreign aid distribution." *World Politics* 66(2):293–330.

Katusiimeh, Mesharch W, Arthur PJ Mol and Kees Burger. 2012. "The operations and effectiveness of public and private provision of solid waste collection services in Kampala." *Habitat International* 36(2):247–252.

Linders, Dennis. 2012. "From e-government to we-government: Defining a typology for citizen coproduction in the age of social media." *Government Information Quarterly* 29(4):446–454.

McCubbins, Mathew D and Thomas Schwartz. 1984. "Congressional oversight overlooked: Police patrols versus fire alarms." *American Journal of Political Science* pp. 165–179.

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

Mossberger, Karen, Yonghong Wu and Jared Crawford. 2013. "Connecting citizens and local governments? Social media and interactivity in major US cities." *Government Information Quarterly* 30(4):351–358.

Ntaliani, Maria, Constantina Costopoulou and Sotiris Karetsos. 2008. "Mobile government: A challenge for agriculture." *Government Information Quarterly* 25(4):699–716.

Oates, Briony J. 2003. "The potential contribution of ICTs to the political process." *Electronic Journal of e-Government* .

of Statistics, Uganda Bureau. N.d. National Population and Housing Census 2014 Area Specific Profiles: Kampala Capital City Authority. Technical report.
URL: *http://www.ubos.org/onlinefiles/uploads/ubos/2014CensusProfiles/KAMPALA-KCCA.pdf*

Okot-Okumu, James and Richard Nyenje. 2011. "Municipal solid waste management under decentralization in Uganda." *Habitat International* 35(4):537–543.

Olken, Benjamin A. 2007. "Monitoring corruption: evidence from a field experiment in Indonesia." *Journal of political Economy* 115(2):200–249.

Oteng-Ababio, Martin. 2010. "Private sector involvement in solid waste management in the Greater Accra Metropolitan Area in Ghana." *Waste Management & Research* 28(4):322–329.

Rotberg, Robert I and Jenny C Aker. 2013. "Mobile phones: uplifting weak and failed states." *The Washington Quarterly* 36(1):111–125.

Saxton, Gregory D, Onook Oh and Rajiv Kishore. 2013. "Rules of crowdsourcing: Models, issues, and systems of control." *Information Systems Management* 30(1):2–20.

Scott, Steven L. 2010. "A modern Bayesian look at the multi-armed bandit." *Applied Stochastic Models in Business and Industry* 26(6):639–658.

Seltzer, Ethan and Dillon Mahmoudi. 2013. "Citizen participation, open innovation, and crowdsourcing: Challenges and opportunities for planning." *CPL bibliography* 28(1):3–18.

Sjoberg, Fredrik M, Jonathan Mellon and Tiago Peixoto. 2017. "The effect of bureaucratic responsiveness on citizen participation." *Public Administration Review* 77(3):340–351.

Vermeiren, Karolien, Anton Van Rompaey, Maarten Loopmans, Eria Serwajja and Paul Mukwaya. 2012. "Urban growth of Kampala, Uganda: Pattern analysis and scenario development." *Landscape and Urban Planning* 106(2):199–206.

Zook, Matthew, Mark Graham, Taylor Shelton and Sean Gorman. 2010. "Volunteered geographic information and crowdsourcing disaster relief: a case study of the Haitian earthquake." *World Medical & Health Policy* 2(2):7–33.

Zurovac, Dejan, Ambrose O Talisuna and Robert W Snow. 2012. "Mobile phone text messaging: tool for malaria control in Africa." *PLoS medicine* 9(2):e1001176.

Supporting Information

Contents

A Waste Pile Measurement Protocol

A.1 Background

The measurement of waste piles was carried out during the Baseline and two the Midlines of the study. Each waste pile was measured once during the Baseline and once during the subsequent Midlines. The following subsections contain the protocols and instruments used to implement the measurements. In the Baseline subsection, there is a prototol for locating the most problematic waste pile locations, as well as the waste-pile audit survey - used to record the particulars of each waste pile. In the Midline subsection, the waste pile audit survey used in the Baseline was modified to cater for instances where the wastepile was cleared and is thus non-existent.

A.2 Baseline

Locating the Waste Piles

This is the protocol given to the enumerators to guide them in locating the most problematic waste pile locations in each zone visited. The language is in second-person because the enumerators were supposed to read and follow these instructions.

Dividing Up The Zone

Upon arriving at the zone, the first activity is to divide it up into 4 sections. These sections should be as equal as possible. You will work with the LC1 to determine the boundaries of the zone, and divide it up into four sections. You will assign each of these sections a letter from A to D. For instance, here is an example of how a divided up zone will look like. Dividing up the zone is essential for collecting data which is representative of the entire zone. Each of the sections will form the basis of where the different activities will be carried out.

Locating A Solid Waste Pile

Our target is to find the most problematic solid waste location (or site) in each of the four sections of a zone. The definition of "problematic" is contextual. It depends on the particular section of the zone. So, you will get that information from locals. Ask the LC or the people you find in the particular section of the zone. Ideally, the site should be a location where the locals feel that the obligation to remove the rubbish lies with KCCA or a contractor. Examples include unofficial dumping sites, designated dumping areas, rubbish collection points, and drainage channels.

However, the ultimate people to decide are the locals. If they feel that a specific location is problematic in terms of garbage, then that is the location we shall take. Once the solid waste pile has been located, the next step is to conduct the waste pile audit.

Measuring the Waste Pile (Waste Pile Audit Survey) Name of Staff Member______
Name of Zone______
Name of Division *(SELECT)*

1. Central
2. Kawempe
3. Makindye
4. Nakawa
5. Rubaga

B Data Cleaning

The results reported in the following section are those arrived at after an extensive effort to clean the data. We had two kinds of cleaning that are particularly important to mention as part of the present analysis.

First, the data collection reported in this paper depended on the ability to visit the same area repeatedly to assess the area of waste accumulation. Because the amount of waste that people would add to unmanaged piles is directly related to the availability and use of formal pick ups, understanding changes in pile sizes that community members identified as most important should be a strong measure of waste services. We cross-checked the GPS locations of all piles in baseline and both midline waves and excluded from the data any pile location that was more than 100m from the baseline location, based on the field-tested accuracy of the tablets that we used for enumeration.

Second, there appears to be unit errors in each of audit files, with pile sizes recorded that are implausible given the associated photographs. For the reported analyses, we have completed a double-review of all piles. Certain piles with implausible and unverifiable baseline measurement sizes were excluded from our analyses altogether. The review process for the remaining piles was as follows:

1. Two reviewers examined the pile size measurements at each audit and corresponding photo.

2. Each reviewer recorded a score of 1-4 indicating their confidence in the reported pile measurement given the corresponding photograph. Lower scores indicate more confidence in the enumerator's measurement.

3. When applicable, each reviewer offered an alternative measurement based on their interpretation of the provided photograph.

4. Reports from each reviewer were relayed to a final reviewer for adjudication.

    a If either reviewer agreed with the enumerator's initial pile measurement, the original value entered by the enumerator was kept.

b If both reviewers disagree with the enumerator, but agree with each other closely, then agreed value is automatically kept. To the extent that the suggested pile sizes were reasonably close, the suggested size closer to the original enumerator value was kept without further checking.

c If both reviewers disagree with the enumerator and they disagree with each other, then the final reviewer checked and assigned a final measurement based on the pile's corresponding photograph.

B.1 Robustness to Exclusion Criteria

Using our recoded estimates of waste pile size does not change the nature of our results. While using our estimates of pile size does reduce the standard errors associated with treatment in the randomization inference, we still observe a null effect of treatment on changes in pile size and characteristics (see Tables F1, 1, F2, 2).

C Deviations from Pre-Analysis Plan

Timing We pre-registered the design and analysis of this study on November 18, 2016 prior to any research activities, including baseline data collection (pre-registration [REDACTED]). That pre-analysis plan describes our research plan for the current study only, exclusive of earlier phases that dealt with promoting citizen reporting. We had originally planned a 7-month study period, but due to the holiday season at the end of 2016 and beginning of 2017, we suspended platform operation for a time. We extended the timing of the first and second post-treatment audits accordingly.

Measurement We had planned to measure depth of waste piles to calculate volume, along with the proximity of the waste heap to residences, active businesses, and public roads. We found these measurements to be infeasible based on pilot measurement activities conducted during enumerator training. In the end, we only measured the area covered by waste, rather than the volume of waste.

Analysis Our pre-registered estimating equation was:

$$\Delta Y_j = \alpha + \tau M_j^+ + \beta \mathbf{X_j} + \nu_h + \epsilon_h \tag{C1}$$

To increase power, we use this modify this analytical strategy and use the baseline pile size as a covariate as outlined in Eq. 1 instead of using it to directly transform the outcome variable, per guidance discussed in McKenzie (2012). Because we fail to find significant treatment effects, we believe boosting power is desirable given relatively low autocorrelation in pile sizes between measurement waves. Additionally, instead of aggregating to the zone level as pre-specified, we use individual piles as the unit of analysis. Since we were not able to find the same number of piles in all zones, this is a better analytical strategy.

Outcomes As noted in our pre-analysis plan, our secondary hypotheses pertained to resident satisfaction with waste services, but our ability to measure satisfaction and test these hypotheses depended on our ability to raise additional funds. We were ultimately unsuccessful in raising additional funds, so we are not able to test any of the H2 hypotheses.

We include an additional primary dependent variable in our analyses: a dummy variable indicating whether or not enumerators found a waste pile at the baseline and subsequent midline audits. This binary indicator of waste pile presence is not among the set of preregistered dependent variables, which include the change in total waste pile area (m^2) and the change in waste pile area along a number of characteristics (e.g. total change in burnt waste pile area). However, we feel justified in its use given the large potential for measurement error in our primary pre-registered dependent variable, change in waste pile size (m^2). As noted previously, we undertook multiple rounds of data cleaning and verification to address seemingly implausible waste pile measurements taken during enumerator audits. When possible, we used enumerators' photos of waste piles to verify or alter the recorded pile size, but frequently we were forced to exclude waste piles from the analyses given the poor quality of enumerator pictures and implausibly large measures (likely made by recording centimeters instead of meters). Therefore, using the binary indicator of pile presence proved a logical choice when conducting our analyses.

At each audit, enumerators recorded whether or not the waste pile had been cleaned up at each location; even poor quality photos can be used to corroborate the accuracy of these reports. Notably, too, we find no major difference in the results when using the preregistered dependent variables and the dummy "pile cleaned" variable. Across all variable specifications, treatment has no statistically significant effect on waste pile size or presence.

D KCCA Administrative Set-Up

There are five key positions within the KCCA Waste Management Unit (WMU) which are responsible for ensuring that all waste produced within Kampala is collected, transported and properly disposed of in landfills. These are the Supervisor – Solid Waste, Solid Waste Officer, Contract Manager, Fleet Supervisor and Solid Waste Scout.

The Supervisor – Solid Waste (hereafter referred as the Supervisor) is the person in charge of the WMU. They are charged with creating an overall strategy, assigning responsibilities and overseeing the operations of the unit. All staff within the unit report directly or indirectly to the Supervisor.

The Solid Waste Officer is responsible for the day-to-day running of operations which deal with the collection and transportation of solid waste. Each Solid Waste Officer is in charge of a division – one of the key administrative units within Kampala. Each division has its own waste trucks, collection equipment, and personnel (including Fleet Supervisors and Solid Waste Scouts). The Solid Waste Officer manages all these, and reports directly to the Supervisor.

The Fleet Supervisor is a casual worker who works within a Division under the Solid Waste Officer. Their main task is to identify the most efficient routes to be taken by the garbage trucks so that there is maximum impact in terms of garbage collected. Trucks generally make stopovers in different locations to collect garbage. It is the job of the Fleet Supervisor to plan the most efficient routes and communicate them to the drivers.

The Solid Waste Scout is a casual worker who works within a specific location within a Division. The scouts are tasked with collecting waste management related information. Examples include the locations of illegal dumps, the main complaints of residents, and the operations of private contractors. They report directly to the Solid Waste Officer and are supposed to be the "eyes and ears" of the officer in different parts of the city.

The Contract Manager is the person responsible for monitoring the compliance of private contractors. Under the PPP arrangement to waste management, Kampala was divided into six waste management service zones. Three contractors were awarded concessions, with each covering two zones. A Contract Manager was appointed to monitor operations in each of the waste management service zones. At the time of implementing the project, each of the Solid Waste Officers was appointed to be the Contract Manager for one of the zones, and the sixth zone has the Supervisor as the Contract Manager.

Information and Allocating Waste Services in Kampala

The KCCA designs its waste services in Kampala based on each zone's unique waste management needs. For instance, the KCCA will send additional clean-up crews to zones where illegal dumping is problematic, or coordinate with private contractors in zones reporting infrequent truck visits. This localized style of service provision relies on the KCCA's ability to uniformly monitor the delivery of waste services throughout Kampala. When determining where and how to allocate services, the KCCA relies on three primary sources of information.

First, the KCCA uses administrative records to inform its delivery of waste services. Using current and projected population data from the Ugandan Bureau of Statistics, the KCCA estimates the amount of waste production in each zone and adjusts its services accordingly. Relatedly, the KCCA has administrative records noting the date, weight, and origin of all waste deposited at its Kiteezi dumpsite. While an imperfect source of data due to non-uniform waste collection efforts across Kampala, the KCCA uses this data to assess the current and future waste service needs of zones across the city.

Second, the KCCA uses its staff located throughout Kampala to informally monitor waste conditions and service delivery. According to the former Solid Waste Unit Supervior, the KCCA employs up to 200 casual workers—known as KCCA Solid Waste Scouts—to report on a variety of problematic waste conditions, such as illegal dump sites, open sewers, or leaking drainage pipes. While never systematized, reporting from Scouts allows the KCCA to internally monitor waste management needs across Kampala and accurately deploy services when required. In the words of the current Supervisor: "I have a problem of illegal dumping. And I have my scouts. When they find a suspect, they use WhatsApp to send a message, I send a car to pick the suspect and take them to court" (see interview J).

Finally, the KCCA engages with citizens to collect information on waste management throughout the city. Using a number of channels—e.g. Twitter, WhatsApp, a toll-free line, office walk-ins, community outreach events—the KCCA informally seeks feedback from citizens to help identify zone-specific waste management needs. KCCA staff initially viewed the citizen monitoring program as a way to augment this flow of information [11].

E List of KCCA Interviews

A Solid Waste Officer, 02/05/2018

B Solid Waste Officer, 02/05/2018

C IT Support Staff, 02/05/2018

D Solid Waste Officer, 02/06/2018

E Solid Waste Officer, 02/06/2018

F Solid Waste Officer, 02/06/2018

G Supervisor - Solid Waste Department (Former), 02/06/2018

H Solid Waste Officer, 02/07/2018

I Solid Waste Officer, 02/07/2018

J Supervisor - Solid Waste Department (Current), 02/08/2018

F Additional Figures and Tables

This section contains additional tables and figures referenced in the main text but omitted because of space constraints.

[11] Citizen monitor reports contained responses to a series of prompts asking monitors about waste conditions and services in their neighborhood. For instance, one prompt asked “Does a rubbish truck come into your neighborhood?" with response option "A) No", "B) Yes", and "C) I don't know.” For this question, responses A and C indicate responses indicate a shortfall in the delivery of waste services: the citizen monitor has not seen or cannot recall if she has seen a collection truck recently. Citizen monitor responses to each prompt were aggregated by zone and delivered to the KCCA as indicators of zone-level waste conditions and service quality.

Figure F1: Balance on pre-treatment covariates

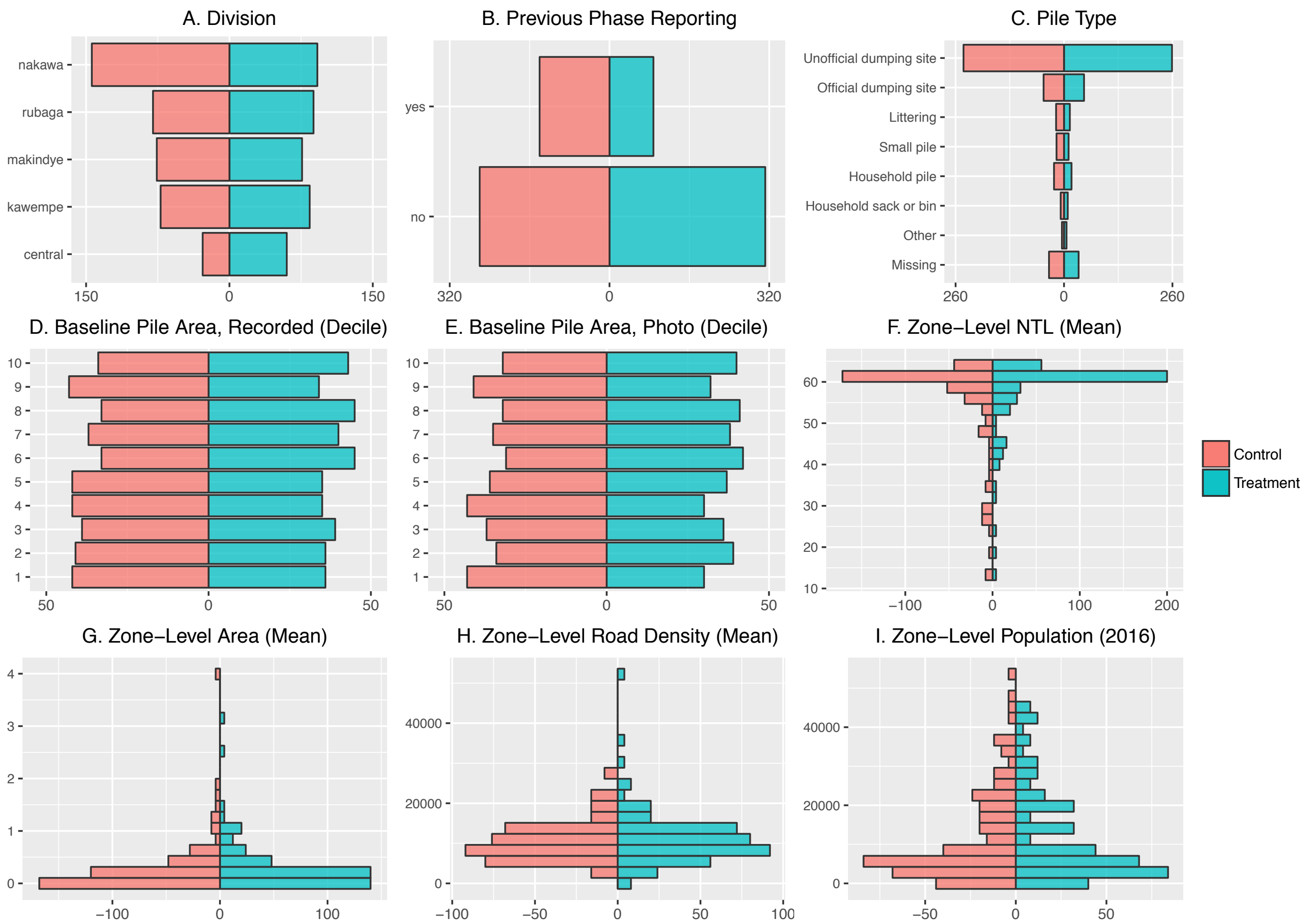


Figure Notes:

1. ***For Plots D and E:*** *to better visualize the data on pile sizes—which include significant outliers in both treatment and control groups—we placed piles into deciles based on their pile size. The first decile contained piles between 0 and 2* m^2*, the fifth decile contained piles between 7.5 and 9* m^2*, and the tenth decile contained piles between 50 and 3000* m^2*.*
2. ***For Plots F, G, H, and I:*** *figures use default binwidth specified by ggplot2.*

Table F1: RI Results, Primary Dependent Variables (Raw)

	Pile Size	Pile Size	Pile Cleaned	Pile Cleaned	Pile Rank	Pile Rank
Audit	M1	M2	M1	M2	M1	M2
Treatment Effect	3.52	6	-0.01	0	-11.96	-8.48
Standard Error	5.46	11.19	0.03	0.04	16.76	14.85
p-value	0.68	0.6	0.32	0.44	0.24	0.29
N	621	621	621	621	621	621

Note: results calculated using raw waste pile size measurements.

Description of Dependent Variables

1. ***Pile Size:*** *change in waste pile area between the baseline audit and specified midline audit (m^2). Enumerators recorded waste pile dimensions at each audit. These dimensions were used to estimate waste pile area for each site in the sample. Waste pile size at the baseline audit was subtracted from waste pile size at each midline audit to produce an estimate of the change in waste pile size from the baseline audit to each midline audit.*
2. ***Pile Cleaned****: dummy variable indicating whether a waste pile was cleaned at the specified midline audit. Recorded values of 0 indicate that no pile was present at the given midline audit (e.g. the pile had been cleaned); recorded values of 1 indicate the opposite.*
3. ***Pile Rank****: change in waste pile size rank between the baseline audit and specified midline audit. Due to high variance in the recorded waste pile sizes, we perform a rank test comparing the ranked change in pile size between the baseline audit and each midline audit. Waste piles were ranked at each audit based on their size relative to other waste piles. Then, waste pile ranks at the baseline audit were subtracted from waste pile ranks at each midline audit to create a change in rank value.*

Table F2: RI Results, Alternate Dependent Variables (Raw)

	Storage	Storage	Organization	Organization	Burning	Burning
Variable Specification	A	B	A	B	A	B
Treatment Effect	2.02	2.18	3.55	3.74	2.2	1.6
Standard Error	11.7	11.54	11.7	11.66	1.72	1.93
p-value	0.54	0.55	0.58	0.59	0.9	0.78
N	621	621	621	621	621	621

Note: results calculated using raw waste pile size measurements.

Description of Dependent Variables

1. ***Storage:*** *change in uncontained waste pile area (m^2). At each midline audit, enumerators recorded how rubbish was stored in each waste pile. Responses ranged from "all of the rubbish is neatly contained with sacks or other containers" to "no rubbish is contained in sacks or containers." Each response was assigned a scalar between 0.0 and 1.0, which was multiplied against the recorded waste pile area to generate an estimate of uncontained waste pile area. Estimates were differenced to calculate the change in uncontained waste pile area.*

2. ***Organization:*** *change in unorganized waste pile area (m^2). At each midline audit, enumerators recorded the dispersion of rubbish in each waste pile. Responses ranged from "all of the rubbish is collected in a single pile" to "rubbish is spread all around [with] no evidence of the rubbish being organized." Each response was assigned a scalar between 0.0 and 1.0, which was multiplied against the recorded waste pile area to generate an estimate of unorganized waste pile area. Estimates from each midline were differenced to calculate the change in unorganized waste pile area.*

3. ***Burning:*** *change in burnt waste pile area. At each midline audit, enumerators recorded any evidence of burning they observed at each waste pile. Responses ranged from "no evidence of burning" to "more than half of the area of the rubbish pile contains evidence of burning." Each response was assigned a scalar between 0.0 and 1.0, which was multiplied against the recorded waste pile area to generate an estimate of burnt waste pile area. Estimates from each midline were differenced to calculate the change in burnt waste pile area.*

Variable specification A is less conservative than variable specification B. Specification A assigns a smaller scalar score to enumerator responses indicating less organization/storage and more burning. As a result, mean estimates of unmanaged, uncontained, or burnt pile area using specification A are smaller than similar estimates using specification B.

Table F3: Treatment Effect of Citizen Reporting Conditional on the Party Affiliation of the Division Councillor Elected in 2016.

	DV: Pile Cleaned (0/1) or Change in Waste Pile Size (m^2)			
	M1 Cleaned	M1 Change	M2 Cleaned	M2 Change
	(1)	(2)	(3)	(4)
Treatment	2.964	5.823	9.421	12.691
	(10.945)	(12.260)	(5.736)	(8.515)
Independent	0.876	−4.053	2.041	−3.544
	(13.170)	(14.748)	(6.907)	(10.251)
Opposition	1.455	9.184	−3.181	5.629
	(13.398)	(14.987)	(7.023)	(10.413)
Baseline Pile Area	0.258***		0.156***	
	(0.075)		(0.039)	
Treatment X Independent	−3.953	−4.841	−5.725	−6.752
	(19.767)	(22.150)	(10.374)	(15.407)
Treatment X Opposition	28.514	10.640	−9.198	−29.536**
	(18.968)	(21.158)	(9.957)	(14.720)
Covariates	Yes	Yes	Yes	Yes
Observations	391	391	392	392
R^2	0.063	0.011	0.073	0.031
Adjusted R^2	0.028	−0.023	0.039	−0.002
Residual Std. Error	74.698	83.705	39.223	58.252
F Statistic	1.805**	0.335	2.124**	0.944
*Note:*two-tailed tests			*p<0.1; **p<0.05; ***p<0.01	

Table F4: Estimated Effects of Treatment Conditional on Zone-Level Response Rate

	DV: Pile Cleaned (0/1) or Change in Waste Pile Size (m^2)			
	M1 Cleaned	M1 Change	M2 Cleaned	M2 Change
	(1)	(2)	(3)	(4)
Zone-Level Response Rate	0.126	-35.858	-0.035	114.571
	(0.244)	(58.304)	(0.296)	(126.081)
Baseline Pile Area	0.0002	-0.563^{***}	-0.0003	-0.740^{***}
	(0.001)	(0.121)	(0.001)	(0.261)
P1/P2 Monitoring	0.068	23.004^{*}	-0.033	-5.903
	(0.050)	(11.997)	(0.061)	(25.944)
Covariates	Yes	Yes	Yes	Yes
Observations	313	313	313	313
R^2	0.052	0.112	0.026	0.064
Adjusted R^2	0.017	0.079	-0.010	0.030
Residual Std. Error	0.338	80.844	0.411	174.823
F Statistic	1.491	3.446^{***}	0.723	1.874^{**}

Note: two-tailed tests $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table F5: Estimated Effects of Treatment, Conditional on Baseline Quality of Service Provision

	DV: Pile Cleaned (0/1) or Change in Waste Pile Size (m^2)			
	M1 Cleaned	M1 Change	M2 Cleaned	M2 Change
	(1)	(2)	(3)	(4)
Service Quality	−0.041	2.227	−0.059	−0.059
	(0.042)	(10.166)	(0.051)	(0.051)
Baseline Pile Area	0.0002	−0.563***	−0.0003	−0.0003
	(0.001)	(0.121)	(0.001)	(0.001)
P1/P2 Monitoring	0.073	22.659*	−0.027	−0.027
	(0.050)	(12.050)	(0.061)	(0.061)
Covariates	Yes	Yes	Yes	Yes
Observations	313	313	313	313
R^2	0.054	0.111	0.030	0.030
Adjusted R^2	0.019	0.078	−0.006	−0.006
Residual Std. Error	0.338	80.888	0.410	0.410
F Statistic	1.556	3.412***	0.843	0.843

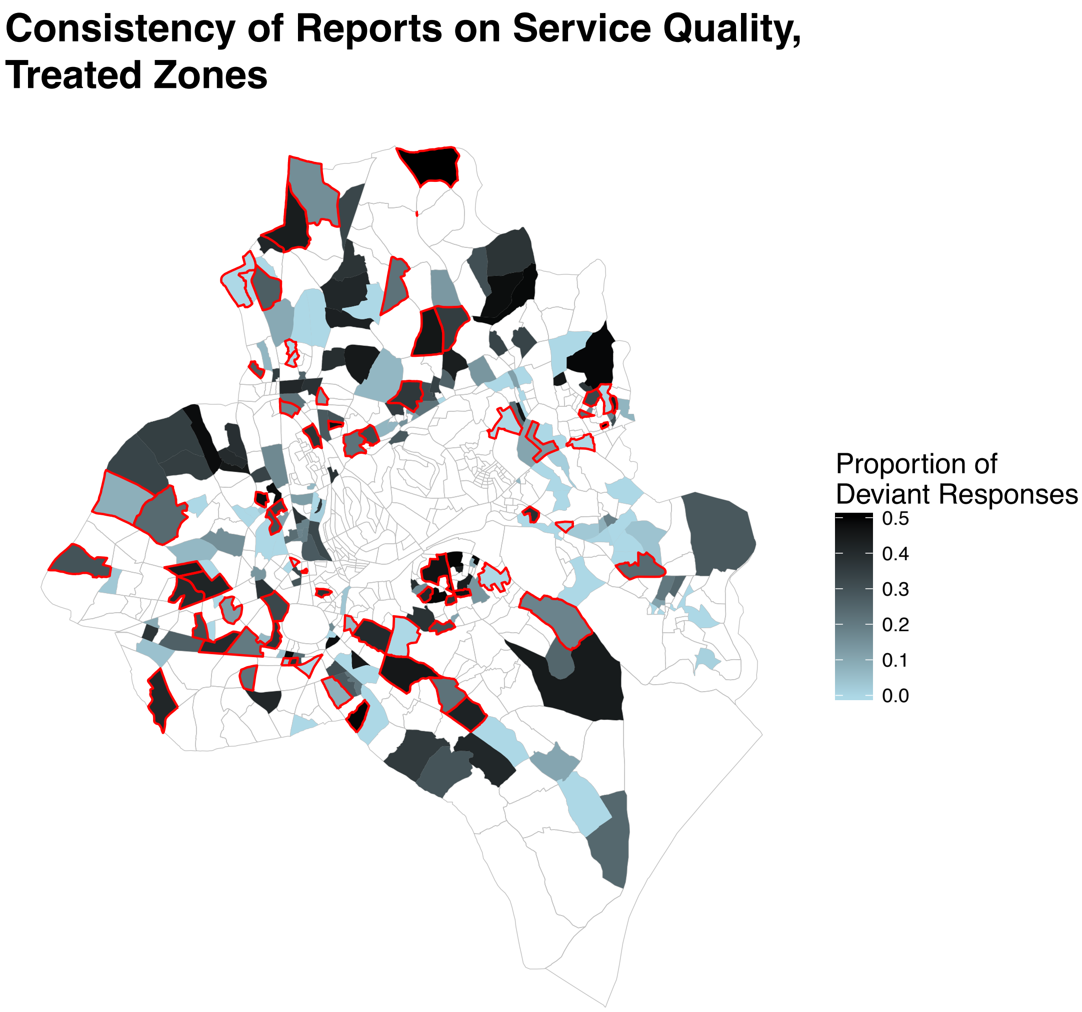
Note: two-tailed tests *p<0.1; **p<0.05; ***p<0.01

Table F6: Estimated Effects of Treatment Conditional on Zone-Level Dissatisfaction

	DV: Pile Cleaned (0/1) or Change in Waste Pile Size (m^2)			
	M1 Cleaned	M1 Change	M2 Cleaned	M2 Change
	(1)	(2)	(3)	(4)
Dissatisfaction	0.068	19.286*	0.036	−0.092
	(0.048)	(11.585)	(0.058)	(25.081)
Baseline Pile Area	0.0003	−0.584***	−0.0003	−0.818***
	(0.001)	(0.124)	(0.001)	(0.268)
P1/P2 Monitoring	0.056	17.863	−0.044	−7.023
	(0.052)	(12.539)	(0.063)	(27.146)
Covariates	Yes	Yes	Yes	Yes
Observations	293	293	293	293
R^2	0.065	0.130	0.029	0.070
Adjusted R^2	0.028	0.096	−0.009	0.033
Residual Std. Error	0.339	81.995	0.411	177.510
F Statistic	1.767*	3.824***	0.752	1.911**

Note: two-tailed tests *p<0.1; **p<0.05; ***p<0.01

Figure F2: Consistency of zone-level reporting on KCCA service quality.



Along a standardized measure of poor service provision, zones in red, on average, indicated that KCCA service provision was poor. The standardized measure of poor service provision combined citizen monitor responses on the following indicators: the frequency and accessibility of service provision, reported waste collector treatment of citizens outlined, and the amount of waste burning or litter.

Table F7: Treatment Effect of Citizen Reporting Conditional on Consistency of Zone-Level Reports on Service Quality

	DV: Pile Cleaned (0/1) or Change in Waste Pile Size (m^2)			
	M1 Cleaned	M1 Change	M2 Cleaned	M2 Change
	(1)	(2)	(3)	(4)
Consistency	−0.053	18.930	0.042	41.544
	(0.125)	(30.883)	(0.152)	(65.384)
Baseline Pile Area	0.0003		−0.0003	
	(0.001)		(0.001)	
P1/P2 Monitoring	0.070	23.115*	−0.034	−5.607
	(0.050)	(12.419)	(0.061)	(26.293)
Covariates	Yes	Yes	Yes	Yes
Observations	313	313	313	313
R^2	0.051	0.047	0.026	0.038
Adjusted R^2	0.017	0.016	−0.010	0.006
Residual Std. Error	0.339	83.585	0.411	176.962
F Statistic	1.483	1.504	0.729	1.188
Note: two-tailed tests			*p<0.1; **p<0.05; ***p<0.01	

We use data from baseline surveys deployed in treated zones to measure the consistency of citizen-monitor reports within zones. Here, consistency is operationalized as the proportion of zone-level responses deviating from the zone-level modal response on an overall indicator of KCCA service quality. Higher proportions of deviant responses indicate that citizen-monitors from a given zone were providing inconsistent information to the KCCA regarding the quality of baseline waste services. We construct our overall measure of service quality using information from the baseline surveys on the following: the frequency and accessibility of service provision, reported waste collector treatment of citizens outlined, and the amount of waste burning or litter.

Table F8: Spillover Results, Primary Dependent Variables (Raw)

	Pile Size	Pile Size	Pile Cleaned	Pile Cleaned	Pile Rank	Pile Rank
Audit	M1	M2	M1	M2	M1	M2
Treatment Effect	1.22	9.42	0	0.01	2.44	3.02
Standard Error	5.98	7.57	0.03	0.05	21.83	24.27
p-value	0.84	0.22	0.91	0.83	0.91	0.9
N	532	532	532	532	532	532

Note: results calculated using raw waste pile size measurements.

Table F9: Spillover Results, Primary Dependent Variables (Cleaned)

	Pile Size	Pile Size	Pile Cleaned	Pile Cleaned	Pile Rank	Pile Rank
Audit	M1	M2	M1	M2	M1	M2
Treatment Effect	-3.36	2.7	0	0.01	5.75	9.77
Standard Error	3.67	3.07	0.03	0.05	20.71	22.8
p-value	0.37	0.39	0.91	0.83	0.78	0.67
N	532	532	532	532	532	532

Note: results calculated using raw waste pile size measurements.

Table F10: No Spillover Results, Primary Dependent Variables (Raw)

	Pile Size	Pile Size	Pile Cleaned	Pile Cleaned	Pile Rank	Pile Rank
Audit	M1	M2	M1	M2	M1	M2
Treatment Effect	-3.36	-1.71	-0.18	-0.04	-194.53	-127.4
Standard Error	5.59	9.51	0.11	0.09	56.59	40.14
p-value	0.56	0.86	0.13	0.64	<0.05	<0.05
N	103	103	103	103	103	103

Note: results calculated using raw waste pile size measurements.

Table F11: No Spillover Results, Primary Dependent Variables (Cleaned)

	Pile Size	Pile Size	Pile Cleaned	Pile Cleaned	Pile Rank	Pile Rank
Audit	M1	M2	M1	M2	M1	M2
Treatment Effect	-0.81	1.95	-0.18	-0.04	-174.08	-138.47
Standard Error	4.76	8.28	0.11	0.09	54.98	37.24
p-value	0.87	0.82	0.13	0.64	<0.05	<0.05
N	103	103	103	103	103	103

Note: results calculated using cleaned waste pile size measurements.