Information Technology and Political Engagement: Mixed Evidence from Uganda*

Guy Grossman† Macartan Humphreys‡ Gabriella Sacramone-Lutz§

October 20, 2016

Abstract

This study integrates three related field experiments to learn about how Information Communications Technology (ICT) innovations can affect who gets to communicate with politicians. We implemented a nationwide experiment in Uganda following a smaller-scale framed field experiment that showed that ICTs can lead to significant "flattening": a greater share of marginalized populations used SMS-based communication compared to existing political communication channels. We find no evidence for such flattening from the national experiment, however. Instead patterns of participation look like politics as usual: participation rates are low and marginalized populations engage at especially low rates. We examine possible reasons for these differences, and then present the design and analysis of a third mechanism experiment that helps parse rival explanations for these divergent patterns. The evidence suggests that even when citizens have issues they want to raise, technological fixes to communication deficits can be easily undercut by structural weaknesses in political systems.

*We thank the National Science Foundation for its generous support, as well as the National Democratic Institute for a fruitful collaboration. This study benefited tremendously from comments from Michael Findley and Daniel Nielson and from participants at EGAP's workshop meeting at Rice University and seminars at John Hopkins University and Harvard University.

†Assistant Professor, Department of Political Science, University of Pennsylvania. ggros@sas.upenn.edu.

‡Professor, Department of Political Science, Columbia University. mh2245@columbia.edu.

§PhD candidate, Department of Political Science,Columbia University. Contact: gs2580@columbia.edu.

1 Introduction

Weak political communication channels characterize many developing countries. Traditional aggregators of interests, such as political parties, have limited reach, and regular public opinion polls are all but non-existent. Most citizens have only limited opportunities to directly communicate with politicians, usually around elections, and only very few are willing to bear the high costs of reaching out to representatives to articulate interests. In turn, weak political communication channels have important implications for the health of a country's democratic institutions: with poor information on their constituents' preferences and policy priorities, elected representatives have a hard time representing, and political parties cannot differentiate themselves in meaningful ways. The starting point of this study is that strengthening weak political communication channels offers a promising way to begin addressing weak accountability relations.

In this paper we report findings from a multi-year research project (involving three related field experiments) designed to test whether innovations in information communication technologies (ICTs) can be harnessed to improve political communication in low-income countries. Since the existence and costs of new ICT platforms are likely correlated with features of a political system that may independently determine political engagement, assessing the effects of technological innovations on political communication is fraught with difficulties. To overcome this identification challenge we partnered with the Parliament of Uganda and the National Democratic Institute (NDI), an international non-government organization (NGO), to implement one of the largest field experiments involving political elites, to date.

The primary experiment examines a nationwide Parliament-led program that introduced new means to contact elected representatives. In the terminology of Harrison and List (2004) this experiment is a "natural field experiment" (NFE), implemented as part of the political process in Uganda. The intervention established and subsidized a mobile-technology platform for political communication, with the goal of increasing and diversifying citizen 'voice.' Citizens in over 100 treatment constituencies were able to communicate with their Member of Parliament (MP) by sending text-messages at low (or no) cost. MPs, representing treated constituencies, could respond to messages via the platform, and use the system's functionalities to aggregate messages and to learn about usage patterns overtime. The ICT platform was introduced to voters via twice daily short radio ads, over a six-months period. This context is unique in scale—the program involved over 10 million voters—but also in nature: change in access was led by political elites and thus provided a relatively strong invitation to citizens to engage in politics.

The results of the nation-wide field experiment are disappointing: uptake in treatment constituencies was low, and marginalized populations largely refrained from using the ICT platform. In fact, because of the disappointing level of citizen engagement and apparent low interest among Members of Parliament (MPs), the Ugandan Parliament ultimately decided to phase out the SMS service altogether.

Importantly, these disappointing findings differed markedly from findings from a more controlled experiment—in the terms of Harrison and List (2004), a "framed field experiment" (FFE)—undertaken before the national program was rolled out. Results from the FFE suggested a relatively high underlying demand, that citizens are exceedingly to the cost of messaging, and that mobile technology could democratize political communication because marginalized constituents were willing to engage at relatively high rates. By contrast, the NFE found little citizen involvement, no price sensitivity and no change in differential access to political elites.

In the second part of the paper we take advantage of differences between the NFE and the FFE to explore the determinants of the disappointing findings in the national experiment. Since both experiments were implemented using subjects from every constituency in Uganda, they involved similar populations, eliminating a common external validity concern—that replications tend to fail because of unobserved features of the experimental subject pool (Allcott, 2015). Instead, we find relatively strong evidence that voters doubt the efficacy of contacting their MP directly, and that larger (structural) inequalities, further discussed below, prevented the ICT program from having effects at scale. We also find evidence suggesting that the tools governments have at their disposal for informing citizens matters, with different marketing strategies employed over these experiments inducing varying invitational effects. Our findings cast doubt on the utility of using short radio ads—as opposed to radio programming (Yanagizawa-Drott, 2014)—to elicit wide-scale participation. We find no evidence, however, that scale itself is driving our core results.

This paper makes several contributions to the literature on political communication, and especially to our understanding of inequalities in political participation. We do so by demonstrating that the underlying willingness to engage in politics—even when using low-cost impersonal communication channels—crucially depends on citizen beliefs about the effectiveness of political engagement, which itself depends on politicians' response to incoming messaging. Though not identified, we provide below evidence that the usage of the system was tightly connected to MP's (in)action. ICTs, we argue, in and of themselves, do not make non-responsive politicians responsive.

The paper also contributes to a growing literature on the effectiveness of using ICT innovations to improve governance outcomes. Past studies have focused on using ICTs to reduce absenteeism among frontline service providers (Duflo et al., 2012), improve election integrity (Callen and Long, 2014), and report corruption (Blair et al., 2015) and violence incidences (van der Wind and Humphreys, 2016). Ours is the first study to examine the role ICTs may play in altering citizen-MPs relationships, in the context of low-income countries.

In addition to our contribution to the study of political communication, we also contribute to ongoing methodological debates on the utility of relatively small-scale highly controlled experiments, such as the framed field experiment described here (and, a fortiori, "artefactual" field experiments or lab experiments) in shedding light on core political processes. Most field experiments—including many natural field experiments—are implemented on a small scale but seek to make claims about large-scale processes. For example, small-scale experiments may be used to test new approaches, be designed as a proof of concept, or test micro-logics that arguably underlie general features of human behavior. Indeed, the 'credibility revolution' in the study of international development is premised on the idea that small-scale field experiments can create a body of knowledge that allows promoting "what works" and eliminating programs and policies that do not (Banerjee and Duflo, 2009). Yet, it is often contestable whether the results of small-scale field experiments can accurately inform theory or form the basis for more general policy (Manski, 2013).

In the remainder of this paper we introduce the research questions that the different field experiments were designed to answer and present the design and results from the scaled-up national program. We then present analyses designed to assess mechanisms that could account for differences in outcomes. Our conclusions focus on the implications for efforts to democratize political communication, and on the implications for learning about political processes from controlled experiments.

2 Access as a Constraint on Political Communication

In many low-income countries, the aggregation of preferences is limited by the weakness of civil society organizations, labor unions and political parties. Potential preference aggregators, such as unions and non-government organizations (NGOs), tend to be located in urban centers and to have a narrow membership base. Political parties may have a wider reach, but many are weakly institutionalized and lack resources and elite cohesion (LeBas, 2011). That parties are often organized on ethnic or geographic basis—rather than by class

or religion—further contributes to the non-programmatic nature of many political parties, in Africa and beyond (Riedl, 2014).

The political implications of weak preference aggregations are manifold. Lacking information on voters' policy priorities, quintessential preference aggregators such as political parties, often focus on valence issues that offer voters little policy differentiation (Bleck and van de Walle, 2013). A focus on valence issues further entails that parties have a weak incentive to reach out to marginalized populations, such as poor people, women, and those living in remote areas. When parties are non-programmatic, the accountability relationship between office holders and voters often narrows down to local clientelistic exchange (Stokes et al., 2013).

While political parties commonly lack the capacity to elicit citizens' preferences, constituents too are often reluctant to bear the costs of political communication. Constituents likely will not invest in articulating preferences if they doubt that government officials would be responsive to citizen demands. This sort of low sense of (external) efficacy is especially prevalent where governments have low capacity and/or low levels of legitimacy (Craig et al., 1990). More so, sense of (internal) efficacy—i.e., the belief that one has the personal ability to participate effectively in politics (Niemi et al., 1991)—is especially low for marginalized populations, whether defined by gender, education, wealth or partisanship (Coleman and Davis, 1976).[1] Notably, low sense of political efficacy is compounded by the high cost of traditional forms of political communication—e.g., traveling large distances to meet public officials in person—that further reduce citizens' incentive to proactively reach out to politicians in order to articulate interests, needs and policy preferences.

2.1 Logic

A simple model clarifies the implications arising from political bias and information asymmetries and identifies the mechanisms we examine in section 5. Our goal is to explore the logic of preference articulation in weak information environments, and especially whether the incentives to bear the costs of political communication might differ across subsections of the population.

Consider a politician who has to decide what shares α_i^j of resources to allocate to group i in sector j subject to $\sum_i \sum_j \alpha_i^j = 1$. Say that each group i values only one sector but the politician is uncertain about the sectoral preference of groups and believes group i favors

[1] Although voting rates among the poor (Kasara and Suryanarayan, 2015) and less educated (Croke et al., n.d) are sometimes higher, this often reflect differences in mobilization (or repression) in different contexts and does not extend immediately to other types of political engagement.

sector j with probability q_i^j. Politicians maximize a weighted average of group welfare with bias parameter β_i, $\sum_i \beta_i = 1$:

$$\sum_i \beta_i \sum_j q_i^j u_i(\alpha_i^j)$$

This general set up can handle a large range of institutional environments. By letting one of the groups or sectors correspond to the politician themselves, or their party, we can capture variation in the degree to which politicians seek to respond to the interests of constituents. Similarly, the β_i term can be interpreted as capturing either the electoral importance of groups or non-electoral significance, such as ethnic affinity with politicians, which past work has demonstrated commonly affects distributional outcomes in the African context (Burgess et al., 2015; Kramon and Posner, 2016). The function u_i may be interpreted either as the benefits to sectors from a policy or the politician's valuation of these benefits to voters. The formulation does not however handle decision rules in which the gains from supporting constituents depend on how satisfied other constituents are, as is the case for coalitional politics (Humphreys, 2008).

The question of interest then is what preferences do politicians have over information on voters? What preferences do voters have over politicians' information and how do these affect equality of outcomes?

Suppose $u_i(\alpha_i^j) = \sqrt{\alpha_i^j}$. Then the optimal allocation is:

$$\alpha_j^i = \frac{(\beta_i q_i^j)^2}{\sum_k \sum_h (\beta_h q_h^k)^2}$$

Thus distributions to groups reflect how informed politicians are about group preferences: they allocate more where they can allocate accurately. In the case with certainty that group 1 prefers sector 1 and uncertainty over group 2's preferences, group 1 receives twice as many benefits as group 2 *even in the absence of ethnic, or other sources of, bias*:

$$\alpha_1^1 + \alpha_1^2 = \frac{\beta_1^2}{\beta_1^2 + \beta_2^2/2}$$

Say now the politician can choose whether to be informed about group 1 or group 2. In this case she compares ex-post utility:

$$\frac{\beta_1^2 + \beta_2^2/2}{\sqrt{\beta_1^2 + \beta_2^2/4 + \beta_2^2/4}} \text{ to } \frac{\beta_2^2 + \beta_1^2/2}{\sqrt{\beta_2^2 + \beta_1^2/4 + \beta_1^2/4}}$$

The former exceeds the latter if $\beta_1 > \beta_2$. Thus politicians invest more in learning about

the preferences of favored groups. Because of this *favored groups are benefited doubly*: because politicians *care* more about them and because they *know* more about them. These simple logics highlight how in the presence of bias there is inequality in information, which contributes to greater inequality in resource allocation.

Say now that voters can take actions to render politicians more informed. Will they have incentives to do so? The sensitivity of allocations to a voter's preferred sector (here across two sectors) to information is given by:

$$\frac{\partial \alpha_1^1}{\partial q_1^1} = 2\beta_1^2 \left(\frac{q_1^1}{\sum_k \sum_h (\beta_k q_k^h)^2} - (2q_1^1 - 1) \frac{(\beta_1 q_1^1)^2}{(\sum_k \sum_h (\beta_k q_k^h)^2)^2} \right) \quad (1)$$

This is positive everywhere, as long as $\beta_1 > 0$.[2] Thus there are always gains for a voter from politicians being more informed about their preferences— more information will always mean better targeting of resources. It is also immediate that the gains to more information are lower the more biased the politician is towards other sectors (the β_{-i} terms appear in the denominator only.) And similarly marginal gains are lower when politicians are more informed about other sectors.

There are however ranges in which groups for whom there is weak information have greater incentives to provide that information than groups about whom there is good information, even if biases go against them. As illustration, note that with two groups, the marginal gains are greater for group 1 if:

$$\beta_1^2 \left(q_1^1 \sum_k \sum_h (\beta_k q_k^h)^2 - (2q_1^1 - 1)(\beta_1 q_1^1)^2 \right) > \beta_2^2 \left(q_2^1 \sum_k \sum_h (\beta_k q_k^h)^2 - (2q_2^1 - 1)(\beta_2 q_2^1)^2 \right)$$

In the case with $q_1^1 = .5$ and $q_2^1 = 1$ this condition is always satisfied as long as $\beta_1 > 0$.[3]

Our model's core results are therefore the following: although marginalized voters might benefit less from informed politicians, they can have *strong incentives to make up the information gap*; given equal opportunities, marginalized citizens about whom there is weak information can have incentives to inform politicians more than less marginalized citizens, even though informed politicians prefer to allocate to less marginalized citizens. Importantly, the incentive to make up the gap comes from the greater effects of information at low levels and not as a response to the greater information in other sectors: as noted

[2] To be positive we require: $q_1^1 \sum_k \sum_h (\beta_k q_k^h)^2 - (2q_1^1 - 1)(\beta_1 q_1^1)^2 > 0$, or equivalently $q_1^1 \sum_{-i} \sum_h (\beta_k q_k^h)^2 + (1 - q_1^1)(\beta_1 q_1^1)^2 > 0$, which clearly holds.

[3] The condition reduces to $\beta_1^2/2 \left(\beta_1^2/2 + \beta_2^2\right) > \beta_2^2 \left(\beta_1^2/2 + \beta_2^2 - \beta_2^2\right)$ or: $\beta_1^4/4 > 0$,

above and seen from Equation 1, increases in the information available about sector a reduces the incentives of sector b to share information.

We emphasize that while these logics are consistent with optimal behavior, the results are sensitive to multiple features of this specification, such as the degree of concavity, order of play, and differential costs of access. For example, with log utility, politician allocations do not depend on information. Note also that differences in *price sensitivity* of marginalized and non-marginalized voters are ambiguous in this framework. Moreover, less marginalized voters may have stronger incentives to communicate if the differences in baseline information levels about the two groups is not great.[4] Thus although the model clarifies the types of logics in operation, it also makes clear that one cannot expect these logics to hold universally.

Several past studies confirm that politicians, at least in many parts of Africa, have limited information on the preferences and priorities of their constituents (Bleck and van de Walle, 2013). Similarly, [authors] report that Members of Parliament in Uganda who were surveyed in-person, describe themselves as insufficiently informed when they vote in plenary and in committee meetings. Weak political communication channels are one reason that, notwithstanding two decades of democratization processes, citizens across Africa generally feel disempowered and unrepresented, commonly expressing low levels of political efficacy (Lynch and Crawford, 2011).

It is interesting to note that, to date, the burgeoning literature on the relationship between information and accountability almost exclusively focuses on the (dearth of) information that citizens have about politicians (Ashworth, 2012). This research project focuses instead on the fact that representatives cannot represent if they lack reliable information on their constituents' preferences and priorities. Indeed, our research project is premised on the idea that improving the information that politicians have—and citizens' awareness that politicians possess such knowledge—may be just as important as improving the information in the hands of voters.

However, studying the logic of political communication described in the model above, and especially the conditions under which different constituents are more likely to invest in communicating their preferences to politicians, is hard. This is because there are likely unobserved characteristics at the individual and the constituency level that are correlated both with the *availability* of communication channels and *intensity* of political communication. We address this challenge by using a field experiment research design in which the

[4]For example if $\beta_1 = 0.45, \beta^2 = 0.55, q_1 = .5$ then group 1 only has a stronger incentive to communicate if $q_2 > 0.9$.

availability of a new innovative communication channel has been randomly assigned.

2.2 Is Mobile Phone Penetration a Disruptive Technology?

Our study contributes to a growing literature that explores whether the rapid penetration of mobile technology across the developing world can be harnessed to improve governance outcomes. Focusing on weak channels of political communication, we ask: can technological innovations that reduce the costs of access to politicians improve channels of political communication, thereby altering MP-constituency relations?

Specifically, we seek to test whether technological platforms that connect voters and MPs, using simple innovations such as text-messaging, can alter the nature of political participation, and ultimately strengthen political representation. Our study is therefore designed to produce evidence that helps us understand how mobile technology might affect who gets to be heard and what gets communicated to political elites. Following the logic of our model, we assess five dimensions related to this question.

First we look at the overall level of political engagement, by reporting the extent to which citizens adopt a newly introduced SMS-based communication platform to articulate their priorities and preferences (system uptake). Naturally, ICT platforms can be an engine of political change only if there exists an underlying demand to communicate preferences to their representatives via mobile-technologies.

Second, the extent to which technology can have a transformative effect on politics, in part, dependent on users' identity. Following our theoretical framework, we are especially concerned with uptake among marginalized citizens (flattening). "Flattening" occurs when the share of marginalized voters using SMS-enabled political communication is high, relative to elite users and relative to the rates at which they use traditional forms of engagement. ICT platforms, on the other hand, do not have a flattening effect when they are used primarily by citizens who are already engaged in frequent communication with political elites using traditional forms of engagement. As derived from the above model, we hypothesize that we would observe a "flattening" effect in our study.

Third, our research is designed to also explore the implications of core system design choices. Specifically, given our interest in uptake across levels of marginalization, we test whether the cost of communication (price) effects the decision to send text-messages. Price is, indeed, a key concern: the cost of sending a message can affect who communicates, and what gets communicated. If messaging is a well-behaved good, subsidization will increase the volume of communication that reaches political elites. If marginalized voters

are, on average, more sensitive to price, then ICT platforms will increase the relative share of previously excluded groups among system users, when offered at lower costs. Alternatively, we could find a larger share of marginalized voters among system users when prices are high, if marginalized voters place a higher value on new access to political elites. More advantaged individuals may be more sensitive to prices if they can switch to more traditional channels of political communication. We therefore randomize the cost of messaging and estimate price effects on the overall level of uptake, and on the type of users.

Fourth, we explore *feedback* effects–the extent to which knowledge of the use of the system by others affects uptake. If potential users view messages as complements—the more people raise an issue, the greater the likelihood that this issue would be addressed—then feedback effects should be positive. On the other hand, if potential users view messages as substitutes, then feedback would exacerbate a collective action problem, leading to overall lower uptake. In our experiment, we randomize the level of information on system usage and estimate the effect of such feedback on the level of uptake.

Finally, we explore *downstream* effects–the extent to which politicians' attitudes and behavior might be affected by exposure to citizens articulating their preferences and priorities via text-messaging. Closely related, we also investigate whether citizens' attitudes—especially their sense of efficacy and trust in government—changes when new political communication channels become readily available.

3 Research Design

The field experiment we study was part of the national strategy of Uganda's Parliament for widening citizen voice. To the best of our knowledge, it is one of the largest field experiments ever to be undertaken with consenting political elites.[5] Below we describe the political context that gave rise to this intervention—summarizing results from the framed field experiment implemented prior to the national intervention—and describe the design of the national intervention and the data used to study it.

3.1 Political Context

Uganda provides a good context for exploring changes to behavior in the wake of introducing a new political communication platform. First, Uganda shares characteristics with

[5]Our study joins a growing body of work using politicians as experimental subjects. See for example, Loewen et al. (2016) and LeVeck et al. (2014) on politicians' decision making, and Grossman and Michelitch (2016) for politicians' response to disseminating information on their performance in office.

many low-income countries on relevant dimensions. It is in the mid-range of the World Bank's low-income economies in terms of economic development (as captured by GDP per capita) and of human development (as captured by HDI ranking).[6] In addition, Uganda is in the middle range ICT ownership, use and access among African countries (*World Development Report*, 2016). These factors strengthen claims to external validity.

Second, data from Uganda supports the assumption of weak political communication channels leading to dearth of information in the hands of politicians. Consider results culled from a survey the research team has conducted with Ugandan Members of Parliament at baseline. We find that the majority of surveyed MPs describe themselves as insufficiently informed when they vote in plenary and in committee meetings. Tellingly, surveyed Ugandans report that MPs do not frequently elicit voter opinions, and generally express a low sense of political efficacy (Grossman et al., n.d.).

Third, results from the framed field experiment (FFE) conducted prior to the launch of the national field experiment, further point to Uganda as a good context for studying the questions at hand. Specifically, findings from the framed field experiment suggests that not only does there exist underlying demand in Uganda for contacting one's MP via a text-messaging platform, but also that IT communications do not necessarily widen the participation gap between more and less marginalized populations. We briefly describe the FFE below.[7]

The framed field experiment, undertaken in 2011, was delivered alongside a survey conducted in every parliamentary constituency in Uganda using a national representative sample. The FFE sought to assess whether demand existed, and to explore the validity of the concern that IT-based communication platforms exacerbate existing inequalities in political access. At the end of the survey, sampled respondents were invited to send a text message to their MP at randomly assigned prices. Discussed in more detail in [citation omitted], the uptake recorded in the FFE—about 5%—suggests that a sizable number of citizens value the opportunity to contact their MPs via SMS.

In addition, usage rates in the FFE were higher among more marginalized populations, possibly reflecting the fact that these populations have fewer opportunities to access politicians and therefore place a higher value on impersonal and inexpensive ICT channels. Experimentally manipulating the price of sending a text message to one's MP, we further found, as expected, that reducing the cost of communication encouraged usage.[8]

[6] Low-human development countries are ranked between 145 (Kenya) and 188 (Niger). Uganda is ranked 163.

[7] A more detailed description of the FFE can be found in (citation Omitted).

[8] Uptake was almost 50% higher for those randomly assigned to a free SMS treatment arm, as compared

Moreover, consistent with the idea that marginalized populations place a higher value on cheap impersonal communication, we found that marginalized populations were not more sensitive to the cost of political communication than less marginalized populations.

The FFE confirmed that Uganda offers a good context to examine the implications of harnessing technological innovations to improve political communication, and that ICT platforms have a genuine potential to alter citizen-MP relations and "flatten" political access. However, the setup of the FFE also had some clear limitations. For example, it allowed only a 'one-shot' opportunity to communicate with MPs, and thus was unable to examine usage patterns over time, in which citizens' behavior is likely (also) a function of both the usage of *other* citizens and the response of their MP to past messages. Moreover, it was implemented in the context of an in-person survey in which subjects interacted with enumerators regarding their political views. This interaction may have made politics more salient to interviewed subjects, further strengthening the invitation to use the platform. The personalized invitation to contact one's MP may have also increased both the sense of empowerment and civic obligation to raise one's voice. Subjects may have also perceived the FFE as closer to a civil society effort than an official government program.

These considerations raise the question of whether similar effects would be found when the intervention was brought to scale, and shifted from being a researcher-led framed experiment to being an institutionalized part of national politics. The natural field experiment described in the next section was designed to address these concerns.

3.2 Intervention

As part of Uganda's Parliament national strategy, a case management platform hosted in the National Parliament was developed, allowing citizens to send messages to their MP via SMS or a voice call to a call center. MPs randomly assigned to participate in the program ("uSpeak") were given access to the platform and trained in its use. The platform allowed MPs to log onto a dashboard where they could read tagged SMS messages from constituents, reply, and see simple descriptive statistics about the messages they received, such as what the priority issues in their constituency were within a selected time-frame. A screenshot of the query dashboard is presented in the Appendix (Figure 6). Only treated MPs were able to receive messages from their constituents via the case management system.

The ICT platform was promoted to citizens through 30 seconds radio advertisement spots, played twice daily on local radio stations over the study' six-months period. The

to those assigned to a treatment group that was not offered any subsidy for texting their MP.

radio ads were in local languages, and featured a skit where actors portraying constituents talked about how uSpeak could be used to draw the MP's attention to important issues, specifically service delivery deficiencies. These skits were first tested using focus groups. The second tier of randomly assigned treatments—price and feedback—was also delivered via the radio ads.

3.2.1 Treatment 1: Elite Participation

The NFE involved 186 MPs who volunteered to be part of a six-month pilot. It was expected that, if deemed successful, all MPs would be phased into the program at the end of the study. Given the sensitivities of providing a new service to only some constituencies, it was agreed that MPs would be selected into the program using a public lottery, managed by NDI. Block randomization was used to assign MPs to treatment groups; MPs were sorted into bins based on their *type* (Woman MP or Constituency MP), *party*, and *region*.[9]

3.2.2 Treatment 2: Variation in Price

To assess the effects of price on uptake, Parliament randomly varied the cost of sending a message to MPs via the uSpeak system, across and within constituencies. Each constituency was assigned 3 months in which uSpeak would be provided free of charge and 3 months without any subsidization. Being sensitive to potential sequence effects, all possible sequences of full price and free months were randomly assigned to constituencies in the treatment group using a blocked design. Note that while the variation in prices in the first period provides a clean separation into price groups, identification based on variation in subsequent months must assume no carryover effects.

3.2.3 Treatment 3: Variation in Feedback

In order to examine whether information on *others'* usage encourages greater uptake, we added a 'feedback' treatment arm delivered through modification of the base radio ads. In one version, voters heard that others had been sending messages to the system about the need to do more in the educational sector. A second variation also highlighted the educational sector but without communicating that others had been using the system to lobby in that area. To the extent that there are complementarities in public goods

[9]Each bin was used to implement a separate public lottery with a target number of MPs selected into treatment based on that MP type's prevalence in the subject pool. Block randomization was used not simply to improve balance in expectation, but also to improve ex-post equality between parties in participation.

messaging, we expect that hearing that others are sending messages about education should increase the willingness to contact one's MP. Indeed, our feedback skit was written explicitly in a way that made this sort of complementarity more apparent to radio listeners.

There were eight unique price sequences that fulfilled the balance requirement (3 of each type), and six unique combinations of the feedback treatments. Together there were 48 unique combinations of price and feedback sequences. These were assigned in a balanced way to treatment constituencies, resulting in roughly 2 constituencies of each unique treatment schedule. Figure 7 (Appendix) provides an example of treatment schemes for a subset of constituencies.

3.3 Data

Data for testing the effects of the uSpeak program come from four sources: (1) a baseline survey of Ugandan adults randomly drawn from all constituencies in Uganda, conducted immediately following the 2011 Parliamentary election, (2) the SMS messages sent by constituents to the uSpeak system, tagged with the date and time they were received, (3) a callback phone survey we conducted with uSpeak users, and (4) an endline survey of a nationally-representative sample of Ugandan adults. In addition, as described below, we conducted an additional follow up experiment with about 3,000 experimental subjects in Arua district to help adjudicate some of the conflicting findings between the natural field experiment and the framed field experiment.

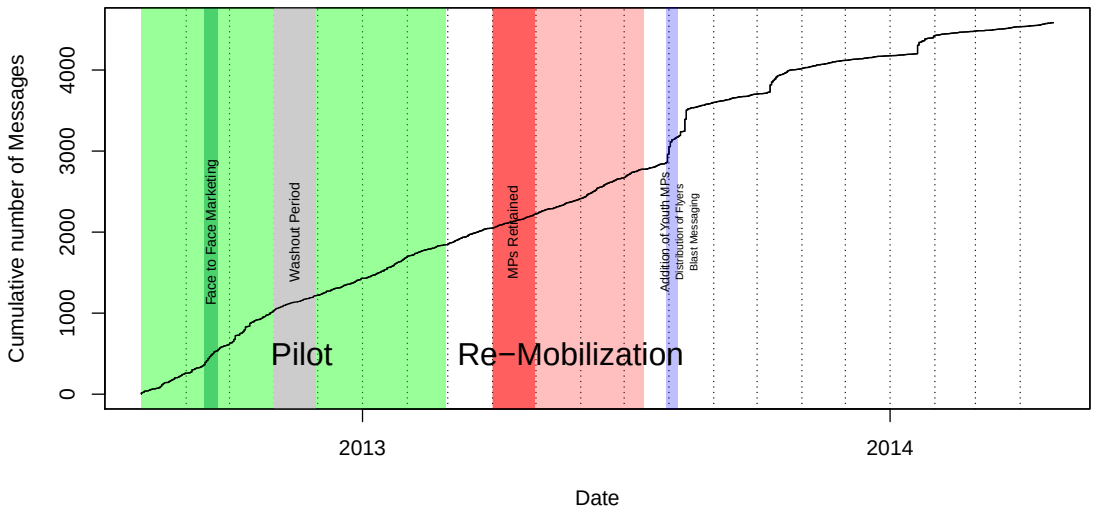
4 Main Results

We focus on core results related to overall uptake, flattening (the characteristics of participating populations), price and feedback effects, and downstream effects. We note that uptake and flattening are not experimental treatment effects in the usual sense, rather they are levels assessed under controlled conditions. Price and feedback effects draw on randomized variation within treatment and downstream effects draw on randomized MP participation in the intervention, as described above. Additional experimental results, such as price effects on the type of message sent, are available in the online appendix. Analyses implemented to explain our results on uptake and flattening are described in section 5.

4.1 Weak Uptake

Unlike the FFE described above, uptake in the NFE was very low. Despite twice daily radio ads and price subsidization throughout the country, MPs in the treatment group received a total of 1946 messages during the 6-month study period. Although census data for Uganda is outdated (2002), we estimate conservatively that the radio ads were played over an area where 10 million voters live. This uptake then corresponds to a monthly usage of about 1 in 30,000. Figure 1 (see also Figure 3) shows the cumulative messaging over time, extending beyond the study period to show uptake in the post-study period including various periods in which an assortment of mobilization efforts were used by Parliament and NDI—none of which produced sustained effects.

Figure 1: **Full Scale Intervention: Uptake**



Note: Cumulative messaging over time. Gray area represents the wash-up period in which no radio spots were played. Green areas denote the period with experimental variation. The figure also shows uptake in the post-experimental period, in which there were attempts by Parliament and NDI to further encourage usage.

4.2 No flattening effects

One of the key findings of the FFE was that the share of marginalized populations—such as women and the poor—among system users was higher than the share of marginalized con-

stituents participating in traditional forms of political engagement. That finding formed the basis of our conclusion that ICT platforms have a genuine potential of flattening political access.

To measure flattening in the national experiment, we conducted a phone survey of system users. Using a call center that the research team had set up, local enumerators contacted all uSpeak users no longer than two months after they had sent a text message to their MP. The short callback survey was designed to elicit information on users' demographics, on whether they received a response from their MP, and general satisfaction with the ICT service. More information on the logistical aspects of the callback survey is in the online appendix.

Comparing results from our callback survey to information culled from the FFE, it is clear that the scaled-up national program failed to replicate the flattening effect identified in the FFR. Specifically, the users of the uSpeak system were wealthier, more highly educated, and overwhelmingly male, compared to those sending text-messages in the FFE. Put plainly, the uSpeak program failed to elicit participation from marginalized population in the way political actors expected. See Figure 2 that provides information on the distribution of wealth, gender, and education, across the two field experiments.

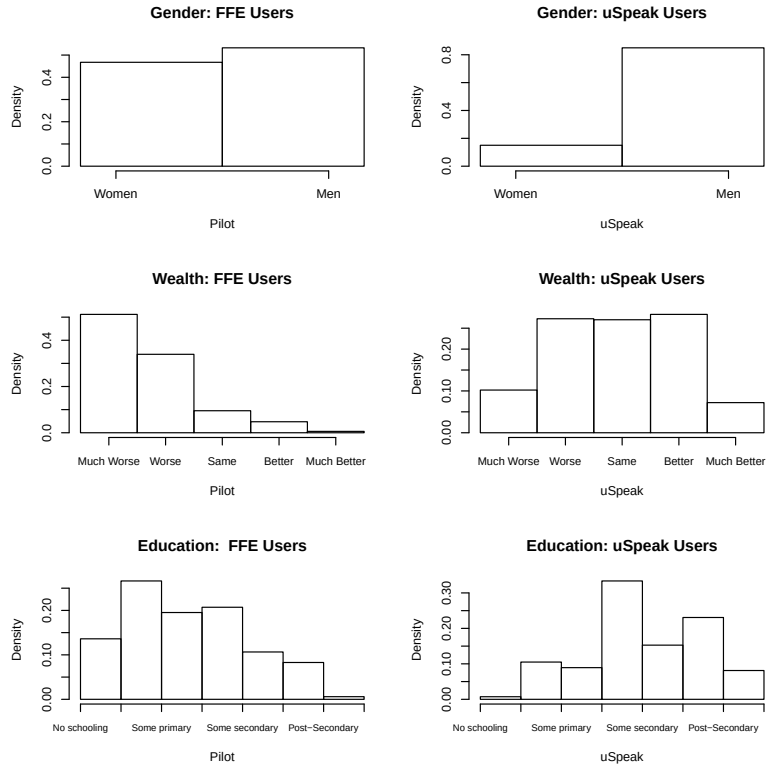
4.3 Insensitivity to Price

Unlike the FFE we find no evidence of overall sensitivity to price in the scaled-up national program. This is evidenced in Figure 3, which shows the monthly rate of messaging in the free and full-price treatment conditions. Testing for a price effect more formally, we run a linear regression of the number of messages received in a given month on price—a binary variable that takes the value of one for full price and zero for months of free messaging—controlling for the month and whether or not we also included feedback, with MPs fixed effects. Results presented in Table 1 suggests that contrary to the FFE, in the scaled national program, price did not significantly affect uptake.

The lack of evidence of a price effect could be explained by the fact that, as we mentioned above, those who chose to send an SMS to their MPs via the ICT system were a small number of politically engaged citizens from, by and large, traditionally powerful groups. Given that at the time of the study the average cost of text-messages was lower than 110 Ugandan Shillings (equivalent to about 2 US cents), it is reasonable to interpret the null effect of the price treatment as stemming from the fact that relatively well-off citizens are not price sensitive when communicating directly with MPs.[10]

[10]The divergence observed in price effect across treatments is reminiscent of the way subjects of controlled

Figure 2: **Demographic Differences: Users in the Framed Field Experiment compared to users in the Natural Field Experiment**



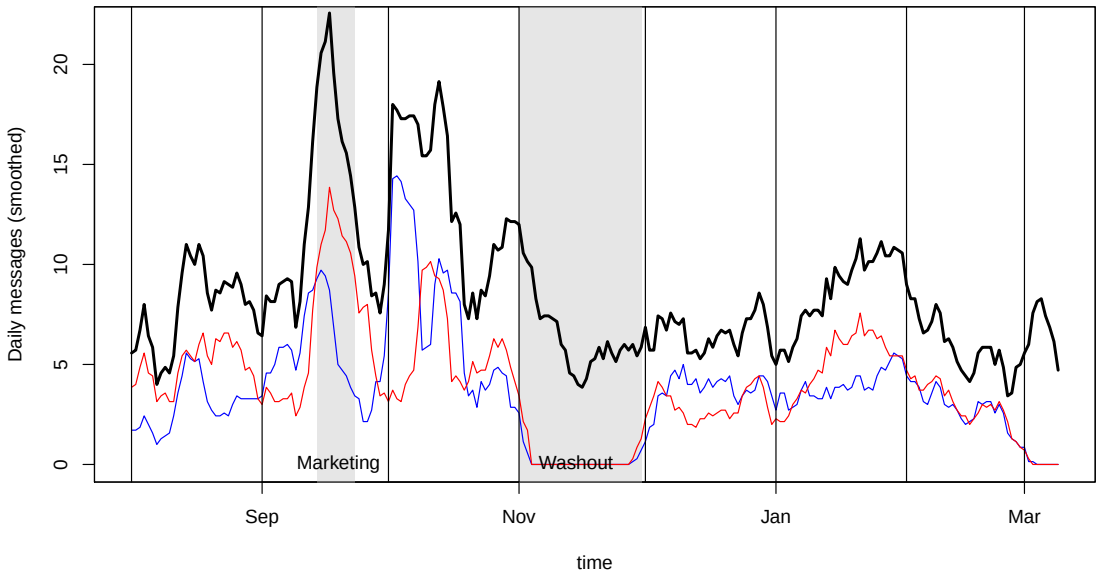
Note: Users in the scaled up national experiment were more likely to be male, better educated, and wealthier than users in the FFE. Data Sources: phone surveys of all system users.

4.4 No Evidence of Downstream Effects

Thus far we have shown that uptake in the scaled-up uSpeak program was low and that fully subsidizing the cost of messaging did not increase voters' proclivity to contact their MP via SMS. Notwithstanding the low rates of usage, it is possible that uSpeak had a positive effect on voters' sense of efficacy and their satisfaction with politics in Uganda. This would be the case if citizens view the existence of the ICT platforms, irrespective of one's own usage, as an important tool for strengthening citizen voice. This was a goal of the intervention and we report on it here briefly. Results in this section use experimental estimates of the effects of the intervention, exploiting the random assignment of the scaled-up program.

laboratory experiments react to even small monetary manipulations that are inconsequential outside laboratory settings.

Figure 3: **Price Effects**



Note: The number of messages per day sent at full price are represented by the blue line, while messages sent for free are represented by the red line. Price clearly did not affect uptake.

To test for the effect of the national program on voters' efficacy we turn to our endline survey. The survey, which took place in July-August 2014, included 2,714 adult respondents from 76 constituencies and 304 villages in 52 districts across Uganda. We provide further information, including descriptive statistics, of this survey in the online appendix.

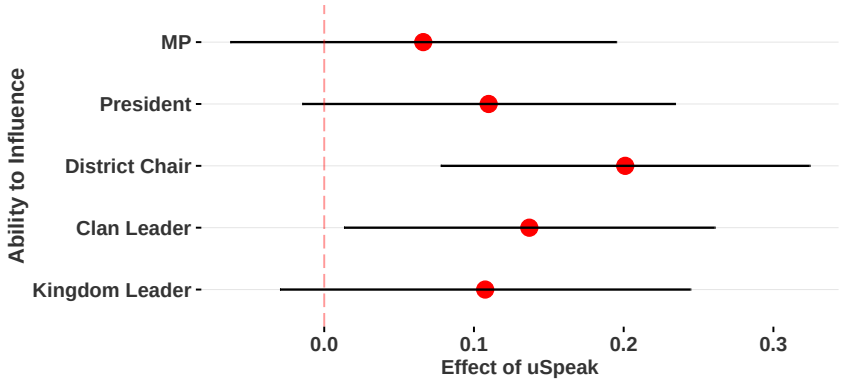
To measure efficacy, we asked survey respondents whether they agree with the following statement: *People like you can do things that can have an influence on the actions of ... [your constituency MP]*; we then repeated the question for the president, district chair, and traditional leaders, which serve as placebo tests. Our key dependent variable is a binary indicator that is equal to one for the 60% of respondents who had agreed that citizen action could influence their MP. We then run a simple OLS model regressing the efficacy outcome on a treatment indicator and district fixed effects. Results, presented in Figure 4, suggest that uSpeak had no discernible effect on voters' sense of efficacy. Note that the graph also gives results from four placebo tests, assessing increased confidence in leaders that are not related to uSpeak, and, surprisingly, passes two of these. While surprising, the pattern suggests that the intervention did not increase the efficacy of citizens with respect to MP

Table 1: Uptake as a function of price and feedback

	Dependent variable:	
	MESSAGES	
	(1)	(2)
Price	−0.081	−0.079
	(0.262)	(0.262)
Education prompt		−0.275
		(0.453)
Education plus Feedback Prompt		−0.132
		(0.452)
Observations	660	660
R^2	0.055	0.055
Adjusted R^2	0.045	0.045
F Statistic	5.251*** (df = 6; 544)	3.972*** (df = 8; 542)
Note:		*p<0.1; **p<0.05; ***p<0.01

engagement relative to the effect on engagement with other political actors.

Figure 4: **Efficacy Effects**



Note: The marginal effect of uSpeak on political efficacy measured as respondents' perception of their ability to impact their MP.

In the online appendix we further provide evidence that the scaled-up national program had no discernible effect on citizens' level of satisfaction with their MPs, or their level of political participation more broadly.

5 Discussion

5.1 Explaining Low Uptake

Experimental findings from the national program conflict with the results from the FFE. Notably, uSpeak resulted in low uptake, even when the service was offered to voters at no cost. Moreover, confirming concerns that ICTs would exacerbate existing inequalities in political access, when uSpeak was used, it was by and large used by citizens whose voice is already more likely to be heard. In other words, the groups that have the weakest access to political processes were also the least likely use the new ICT platform. In this section we explore some of the reasons that may account for uSpeak's low uptake, while in the following section we test several explanations for the fact that—contrary to the FFE—marginalized populations were significantly less likely to use the new ICT platform.[11] Our goal here is not merely to account for these diverging results, but rather to use the analysis in the following two sections to derive substantive (realistic) insights regarding the role ICTs can currently play in improving political communication in low-income countries.

Although the FFE led by the research team was meant to capture the key features of the scaled-up national ICT platform, the introduction of relatively tight experimental control introduces a number of differences.

We first explore the explanatory power of two external features of the NFE—which are common to interventions that are scaled-up from controlled pilots to larger-scale programs—that may have been consequential. We refer to these as 'scale' and 'partner' effects. In addition, we examine the implications of subtle differences in the delivery of the treatment. These 'design' effects are especially relevant for interventions that involve the dissemination of information to subjects.

Changes in *scale* are often described as a problem of general equilibrium effects (Deaton, 2010). This concern is of particular salience when treatment effects are sensitive to the share of treated in the population. Scale effects are of special concern when subjects can accurately infer the magnitude of a program from its delivery method, as is clearly the case in our study. In our setting, it is quite possible that collective action problems get altered substantially as scale increases. Insofar as political communications complement each other, or substitute for each other, increases in scale could lead to greater or lower overall levels of communication.

[11] We do not explore the reasons behind the lack of downstream effects, since the fact that the scaled-up national program generated such weak first-stage results speaks volume to why voters and politicians' attitudes and behavior were not affected by the introduction of uSpeak.

A second possible reason for the low uptake relates to *partners*. Whereas the research team implemented the FFE, the Parliament of Uganda and NDI led the scaled-up national program.[12] In our case, this change in partners might have affected citizen expectations regarding the responsiveness to their messages. In other words, the fact that the scaled-up national intervention was implemented by Parliament rather than by researchers may have reduced the incentives of the target populations to engage.

The third possibility relates to *experimental design* and specifically, to the possibility that details of the mode of treatment delivery—the nuts and bolts of executing field experiments—matter a great deal for citizens in deciding whether or not to communicate with elites. We focus here on two possibilities. The first is that the method of delivery (radio spots) introduced a *treatment compliance effect*: that Ugandans were simply unlikely to hear or internalize appeals issued through mass media, and not less likely to respond, conditional on hearing. This mode of delivery differs from the FFE where the enumerators ensured that respondents unambiguously heard and internalized the information on the SMS platform.

Closely related is the possibility that different methods of delivery have varying degree of an (implicit) *invitational effect*. It is possible that uptake was relatively high in the FFE not simply because the in-person survey context ensured that experimental subjects were aware of the new ICT platform, but also because the enumerators had personally invited respondents to contact their MP. Communicated in the context of a survey, such invitation may appear as a more personal encouragement to engage in politics. A direct personal invitation may have an empowering effect, signaling receptiveness and the possibility that political communication will make a difference. These last two 'design' mechanisms are closely related yet distinct: one is about whether an invitation was empowering and deemed personal, the other is whether an invitation gets heard at all.

We use a number of strategies to adjudicate between these four explanations. First, we exploit a feature of the scaled-up field experiment in which there was variation in the *feedback* provided to voters on the behavior of others. This allows examining whether exacerbating collective action problems due to scale can, at least partially, account for the significantly lower uptake. Second, we conducted a citizen endline survey with a nationally representative sample, which allows assessing—albeit with some lag—ex-post differences in treatment compliance. Third, we hired a Ugandan private marketing firm to examine whether the radio stations NDI had contracted indeed played the ads according to the

[12]Such differences in partners across scales is common, if not typical: for example, the Millennium Villages initiative sought to assess the scope for government led development change by examining an intervention in which government was not a primary actor.

experimental design. Fourth, we implemented an additional "mechanism experiment" in one district in Uganda in which we specifically varied the invitational component. Table 2 summarizes the list of potential explanations, and the source that was used to explore their explanatory power.

Table 2: **Differences between the FFE and the NFE**

	Type	**Data Source**
1	Scale Effects	National experiment (Feedback treatment)
2	Partner Effects	Users survey data (callback)
3a	Design Effects I: Treatment Compliance	Citizens survey data (national sample), Radio monitoring data
3b	Design Effects II: Invitational Effects	Follow-up Mechanism experiment

5.1.1 Scale effects and low uptake

Political communication is subject to a collective action problem. If many others are visibly lobbying a politician, free-ridership may become more likely.

From the model in Section 2.1 above we can see two ways in which a logic of this form can play out. First, if politicians are more informed about others, and better able to target resources to them, this can reduce incentives to provide a politician with information on one's own preferences. Second, scale may also result in lower individual contributions through a simple logic of substitution for members of a given group. Equation 1 gives the marginal returns to a sector's returns from incresead information about that sector. In the exteme case in which information from constituency members were perfect substitutes and citizens faced linear costs, increases in potential information providers would not alter the amount of information provided, in equilibrium, which in turn implies a corresponding reduction in per capital information provision.[13]

Conversely, one could also construct examples in which when many others are lobbying for a common good there may be increasing returns to lobbying (or here too, there may also be increasing incentives to free ride). In short, if the incentives to engage in political communication depend on the perception of how others are engaging, then outcomes at a small scale may look very different to outcomes at a large scale. From this perspective,

[13]More formally, say, utility for an agent j in sector i were given by $u_j(\alpha_i(q_i(\sum_{k\in i} a_k))) - a_j$, keeping the politicians information about other sectors and bias fixed. Here q_j is replaced by $q(\sum_{k\in i} a_k)$ where a_k is the information-providing action of individuals in sector i, taken here to be substitutes. In this case, in any solution the individual information contributions sum to a constant, no matter how many potential information providers there are in a sector.

weaker participation from the scaled up program may reflect a simple failure of collective action.

In order to understand if changes in scale induced free-riding, we look for differences in outcomes due to our *feedback* treatment. Recall that in the scaled up national experiment, we exogenously varied the information constituents received about the level of activity by other voters in previous periods. In particular, a random subset of constituencies was informed, through the short radio ads, that other voters had been using the system to mainly raise issues around education. Under a free-riding logic, such information would depress engagement among those exposed to it.

Returning to Table 1, however, we find no evidence of sensitivity of engagement to information on uptake by others—neither the difference between feedback on education messages and standard marketing spots, nor differences between education marketing with and without feedback is significant. This is consistent with a set of analyses we conducted on the data from the FFE in which we found no evidence for strategic engagement with the system. We conclude that scale, by itself, does not seem to be a key factor driving our divergent results in terms of overall uptake.

5.1.2 Partner Effects and low uptake

Another possible reason for the low uptake witnessed in the scaled-up national program is *partner effects*. Citizens' usage of mobile messaging plausibly increases with the belief that there is a receptive representative at the other side of the interaction (Grossman et al., n.d.). In the model in Section 2.1 these effects are captured by the β terms; citizens will be less likely to communicate if they think that a political actor cares less about their interests. More subtly, they will also be less likely to participate if political actors are better informed about the interests of rival constituents.

Which system should voters expect to produce greater responsiveness by politicians? On one hand, unlike our FFE, the scaled up national program is formally owned and led by Parliament, which signals some level of commitment by politicians. In addition, the dynamic nature of the scaled-up program—i.e., the ability of MPs to interact with citizens directly via the ICT platform—further allows MPs to signal their responsiveness directly. This sort of dynamic reciprocal relationship could not have been established in the 'one-shot' controlled FFE. On the other hand, in the scaled-up program, the communication between citizens and politicians was direct, whereas in the FFE this relationship was mediated by the NDI, which was responsible for delivering the messages to survey respondents' respective MPs. Citizens may believe that their MP will take their messages

more seriously if an NGO mediates the relationship between voters and representatives; for example, if it follows up in case some messages get ignored. Thus it was hard to predict *a priori* how the change in the implementer's identity would affect citizen uptake. We explore (non-experimentally) partner effects in two ways.

First, we use the callback survey ($n = 2,517$ uSpeak users) to calculate MPs' response rate at the constituency level and then test for a correlation between MP's responsiveness and the volume of messaging at the constituency level. We find that only 9 percent of uSpeak users report ever hearing back from their MP; in fact, in almost half of the treated constituencies (44) *not a single uSpeak user had received any response from their MP*. Moreover, analyzing system login information, we find that the majority of MPs did not even read many of the messages sent to them. As expected, we find a positive correlation between messaging and responsiveness, which is consistent with citizens' low engagement being a rational response to their MP's (in)action during the scaled-up study period.

Second, though the callback survey analysis focuses on system users—a self-selected group—we, nonetheless, can also assess whether broader expectations regarding MP inaction may have contributed to the low uptake rate among the general population. Here we examine responses in the citizen endline survey, when our national representative sample was asked to indicate reasons for why people might not use SMS platforms such as uSpeak to communicate with their MPs. Figure 5 (top left) provides information on the share of respondents in treatment constituencies that indicate each possible reason. Tellingly, we find that close to 50 percent of respondents, report that they would not send a message because they do not expect their MP to be responsive, and about a quarter report a reluctance to contact their MP via text-messaging out of fear of bad repercussions.

We do not have information on the expectations of responsiveness from MPs in the FFE and so cannot compare those directly. Nevertheless, the statements by citizens and the very weak responsiveness by politicians suggests that the low engagement with the scaled up program was a rational response on the part of citizens.

5.1.3 Design Effects I: Treatment compliance and low uptake

To test for *compliance* effects we asked respondents in our citizen endline survey directly whether they have ever heard about uSpeak. The survey was implemented a year after the six-months radio campaign, though at a time when the uSpeak system was still active and promoted by Parliament. In light of the time gap, we used a deliberately strong prime, which entailed playing the original radio ad and asking respondents if they have heard of the service the ad sought to promote (uSpeak).

We find that 17% of respondents in control constituencies and 24% of respondents in treatment areas self-report that they ever heard of uSpeak. We note that this is an upper bound because of the possibility of social desirability bias. Also note that not all the control respondents are necessarily lying about their knowledge of the program; this is because radio signals normally have a range that encompass more than a single parliamentary constituency.[14] When probing deeper about respondents' knowledge of the uSpeak program we find that only 6 percent of treatment respondents were able to confidently say that their MP had participated in the program. Moreover, when asked to repeat the four-numbers short-code, less than half a percent of treated constituencies claimed to know the short-code to send a text-message to their MP and an additional 3% report they once knew the number but have since forgotten it. These findings strongly point to the limitation of radio marketing to garner sufficient awareness to the new service.

5.1.4 Design Effects II: Invitational Effects and low uptake

We turn to explore the possibility of a second 'design' effect; namely, that the marketing tools used to inform citizens about a new service or program likely have (unintentional) invitational effects. Recall that the two experiments differed in their mode of 'marketing': whereas the scaled-up national program used 30 seconds radio ads, in the FFE, respondents were invited by enumerators to contact their MP in the context of an in-person survey. As mentioned, direct personal invitation may have an empowering effect, or it may signal greater government responsiveness. Multiple logics consistent with the model in Section 2.1 could underpin these effects—an invitation could in principle change a voter's beliefs about β_i—how much the politician cares about their welfare—as well as about the collection of q_{ij} parameters—politician's knowledge of citizens' preferences. As further discussed below, if such invitational effects operate differently for marginalized and non-marginalized populations, this could account for the differences in observed flattening effects.

To assess the role personal invitations plays in the decision to politically engage using an ICT platform, we implemented a third (mechanism) experiment. In so doing, we take advantage of an existing SMS platform, UBridge, which has been operating in Arua district since late 2014. UBridge was developed in partnership between UNICEF's Ureport platform and Uganda's Governance, Accountability, Participation and Performance [GAPP] project.[15] Unlike uSpeak that connects citizens with national MPs, UBridge was

[14] We minimized spillover by using the name of treated MPs in the 30 second adverts.

[15] With some loss in external validity, our design aims to keep the *treatment compliance* effect constant by focusing on respondents in the UBridge system. We hope that parsing the outcome compliance effect will be the focus of future studies.

designed to open a new channel of communication from citizens to local government officials to specifically report problems of public service delivery. UBridge was launched as a pilot study in over 100 villages across Arua. A study evaluating the effect of getting access to the UBridge system is underway and is not the subject of this paper. At the time of our 'mechanism' experiment, UBridge had 4,568 registered users, out of which 2,720 were explicitly verified by the research team.[16]

On June 13, 2015, UBridge conducted a baseline poll using a robocall system asking users about their attitudes toward budgetary processes. The key outcome of interest is a binary variable that receives the value of 1 if the UBridge user responded to the poll, and 0 otherwise. Of the 2,720 verified users, 12% responded to the opinion poll and shared their views with UBridge. To explore the role of direct invitations on levels of ICT-based political engagement, we asked UBridge to run a modified version of their baseline poll but now experimentally introducing a modest variation in their outreach activities. All users would be invited to participate in an opinion poll regarding taxation, similar to the previous UBridge poll. In a randomly selected treatment group, however, UBridge preceded the call with a set of (blast) text-messages that explicitly invited participants to take part in the weekend poll and that highlighted the importance of individual responses. Further details on the block randomization used in this experiment, as well as the full text of the treatment text-messages, are provided in the online appendix.

Our primary measure is the response (or non-response) by UBridge users to the weekend opinion poll. The encouragement text-messages were delivered on 24, 25, and 26 June 2015, and the poll took place on 26 June. We estimate average treatment effects using a regression that accounts for block fixed effects. Our analysis takes account of the variables used for blocking but introduces no further controls. Our primary regression uses only the verified subset of UBridge users, whereas our secondary analysis includes all registered users whether or not they have been positively verified.

Results, reported in Table 3 (column 1), suggest that invitation had a large positive effect on response rate: 2 percentage points from a base rate of 9.4% for the control group. These results are consisted with findings reported by Grossman et al. (n.d.) in a similar context, and by Dale and Strauss (2009) and Malhotra et al. (2011) in the USA. We note that even though the invitation tested in the mechanism experiment was relatively weak (three text-messages) compared to the in-person invitation used in the FFE, it was able to increase participation rates by over 20%. The evidence at hand allows at to conclude that consistent with insights from the voter mobilization literature (e.g., (Green and Ger-

[16]We verified the identity of registered users through a call center that we set up with the help of Innovation for Poverty Action, Uganda.

ber, 2008)), personal invitations can have a powerful effect on rates of participation. We can further conclude that, at least in low-income countries with characteristics similar to Uganda, short radio ads likely represent a marketing strategy that is too impersonal to mobilize large-scale participation.

Table 3: Mechanism Experiment

	Dependent variable:		
	Provided policy input		
	Base	Primary	Secondary
	(1)	(2)	(3)
Invitation	0.021*	0.021*	0.019**
	(0.011)	(0.011)	(0.008)
Flattening (Male*Invitation)		0.0004	−0.003
		(0.024)	(0.017)
Observations	2,717	2,717	3,957
R^2	0.153	0.153	0.165
Adjusted R^2	0.108	0.107	0.115
Note:	*Male* normalized to have 0 mean. *$p < 0.1$		

5.2 "Flattening": Differential uptake by marginalized groups

The above section has focused on explaining the overall low uptake witnessed in the scaled-up national experiment, especially when compared to the (relatively) high demand observed in the FFE. In this section we build on the theoretical framework developed above (and summarized in Table 2) to explore some of the reasons that may account for the lack of "flattening" effect; namely, that when brought to scale, marginalized populations used the ICT platform to articulate interests at significantly lower rates as compared to more powerful populations.

5.2.1 Scale effects and differential uptake

Returning again to the simple logic described in Section 2.1, scale can affect not simply per capita information provision but also the profiles of providers. Most simply, reduced flattening could result from increases in scale if substitution effects were stronger within more marginalized groups. Unfortunately, while a theoretical possibility, our data do not allow us to exploit the variation in feedback to assess flattening effects due to scale. Nev-

ertheless given the generally weak evidence for any scale effects we think it unlikely that scale could also produce flattening.

5.2.2 Partner effects and differential uptake

It is also theoretically possible that partner effects are responsible for differences in flattening effects. This could arise if marginalized voters are more doubtful of government willingness to engage and are encouraged more by the presence of NGOs. As NGOs often provide services to marginalized communities where governments fail to prove, this explanation has face validity. The data reported in figure 5 do not bear this out however. Although less educated voters were marginally more likely to doubt MP responsiveness, the difference is not large and trends in the opposite direction for poorer citizens.

5.2.3 Design Effects I: Treatment compliance and differential uptake

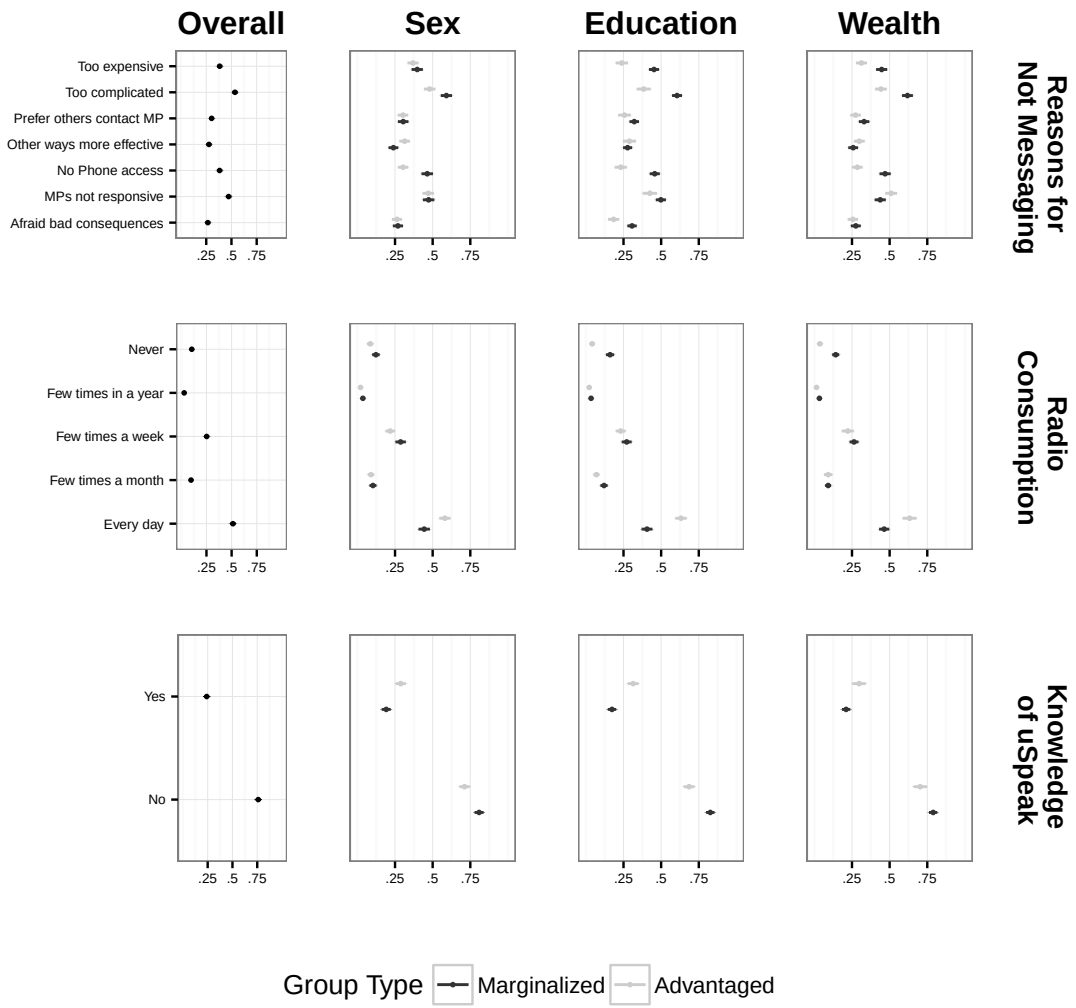
Recall that to test for compliance effects we asked respondents in our endline survey whether they have ever heard about uSpeak. Treatment compliance can help account for differential uptake if more marginalized populations have lower awareness of the new political communication service. Figure 5 thus breaks down the overall awareness of uSpeak by sex, wealth, and education. The figure demonstrates clearly that, marginalized groups, such as women and less-educated (denoted in black), are significantly less likely to be aware of the uSpeak program, compared to non-marginalized groups (denoted in gray).

To further test whether a treatment compliance effect is responsible for the large gaps in participation, we asked survey respondents about radio ownership. Starting with gender gaps, we find that 0.71 of female respondents reside in a household that owns a radio in good working condition, compared to 0.76 of men. Since owning and listening to radio are distinct, we further asked our representative sample about their frequency of radio consumption on a five-point scale (from "never" to "daily").[17] Figure 5 that provides information on the distribution of responses broken down by gender, indicates clearly that women listen to radio much less frequently than men. Moving to another form of marginalization, poverty, we find that poorer respondents were only marginally less likely to have heard of uSpeak though they were significantly less likely to listen to the radio.

Together, the above findings provide evidence of treatment compliance effect as one explanation for the absolute and relative low participation rate of marginalized groups in the uSpeak program.

[17] The question verbatim was: *How often do you get news from the following sources?*.

Figure 5: **Treatment Compliance Differentials**



Note: Differences in the proportion of each group in each response category across three measures of treatment compliance. Means and 95% confidence intervals, $N = 1,324$)

5.2.4 Design Effects II: Invitational effects and differential uptake

In our final analysis we test whether 'invitational' effects can help account for the differential uptake across groups defined by their level of marginalization. The intuition is that personal invitation to participate in politics may have a large marginal effect for marginal-

ized populations who have, on average, lower political (internal) efficacy to begin with.

Recall that to assess the role of invitations in political communication we implemented a third 'mechanism' experiment, using an existing political IT communication platform, UBridge. Recall also that in this experiment our key outcome of interest is whether UBridge users responded to an opinion poll. It is important to note that in implementing the 'mechanism' experiment, we were limited by the amount of information we had on registered UBridge users. Specifically, the only marginalization information the research team had on UBridge users is their gender.

Consistent with past findings demonstrating gender gaps in participation in Africa (Isaksson et al., 2014), we find that, at baseline, 6.6% of women responded to the poll compared to 12.3% of men. Thus in the baseline setting in which ICT is used to communicate with elites one's preferences, but invitations are weak and impersonal, we find evidence of a large and significant gender gap in participation, similar to the one observed in the uSpeak program. Recognizing that gender is only one form of marginalization among others (such as low education attainment, membership in a minority group or poverty), we designed the mechanism experiment to explicitly assess whether the gender gap observed in the baseline would be reduced if messages are accompanied by a stronger invitational component.

Moving to the experimental setting, we begin with a simple cross-tabulation of the raw data by treatment and gender for our verified users subsample. We find that in the control condition, the response rate of registered male users was almost double that of female users (11.2% as against 5.7%). Moving to the personalized invitation treatment, for both male and female UBridge registered users, overall response rate has increased by about 2 percentage points (13.4% against 7.5%). In order to more formally test whether there is a differential effect of the invitational treatment by gender we implement analyses pre-registered at EGAP's registration web page.[18] Results, reported above in Table 3 (columns 2 and 3), confirms that the invitation effect was almost identical across men and women.

We conclude that though personal invitations can have a powerful effect on rates of participation, they do not necessarily have a differential effect by gender. It remains to be explored in future work whether invitational effects have a differential effect for other forms of marginalization.

[18]Note the primary and secondary analyses are as registered. The base column (1) is added to provide information on the unconditional invitation effect; i.e., without including heterogeneous gender effects.

6 Conclusion

This study integrates three related field experiments designed to assess whether innovations in information communication technologies (ICTs) can be harnessed to improve weak political communication, prevalent in many low-income countries. Evidence from a framed field experiment (FFE) conducted before rolling up a national program suggested not only that there is underlying demand to contact representatives using mobile technology, but also that ICTs have a genuine potential to increase levels of political engagement in a way that flattens access for marginalized populations. By contrast, when brought to scale using a natural field experiment (NFE) imlemented nationwide, we find significantly lower levels of citizen engagement, with marginalized populations especially refraining from using the ICT platform to raise voice. These results have implications for theory, policy, and research methodology.

Our study contributes primarily to our understanding of the promises and pitfalls of ICT-based political communications, at least in the context of low-income countries. Consider three findings that help account for the weak uptake of Uganda's national parliamentary communication system.

First, we learned from the FFE that a nontrivial share of citizens, including especially marginalized citizens, want to communicate with their representatives in government using new technological innovations, and are willing to pay to do so. This stands in contrast to accounts of disengagement as reflecting alienation or apathy. We also know that many—though clearly not all (see Figure 5)—have the capacity and means to do so. The results from the FFE support the idea that mobile technology could, under the right conditions, change the relationships between voters and representatives in the developing world. An examination of the scaled up system alone would have masked this core insight.

Second, from an experimental manipulation in the NFE we found little evidence that the differences across field experiments is due to pure scale effects. Specifically, we do not find evidence that system usage is a startetigic response to how many others are contacting their MP via the ICT platform. We believe that improving our understanding of the conditions under which constituents might view IT-based communication with public officials as complements or substitutes, is an important research avenue for future work to explore.

Third, from the 'mechanism' experiment we learned that there is a strong responsiveness to personal invitations to engage politically when interest articulation is at stake, but we do not find evidence of the kind of differential responsiveness that would be needed to

account for differences in flattening effects across experimental settings.

These three findings suggest that the disappointing results of the uSpeak program are *not driven by weak demand*. In contrast, survey evidence suggests weaknesses in the marketing of the system itself. Moreover, our analysis suggests that *partner effects*—i.e., that the change in the identity of the implementer, which was easily observed by experimental subjects—likely have been very consequential. Specifically, we find strong evidence that general trust in the responsiveness of politicians is preventing engagement but is also rational. Interestingly in our case, partner effects do not stem from motivation differences between implementers (as for example identified by Berge et al. (2012)), but rather from the way partner identities interact with citizen expectations. In Ugnada, as in many electoral authroitarian regimes—the most common regime type in Africa—low levels of political efficacy are discouraging political action; ICT innovations, by themseleves, cannot force non-responsive politicians to become responsive.

With the multiple pieces of evidence available to us we infer that the failure of the nationwide program is not simply a function of weak demand on the part of citizens but is a function of larger inequalities. Some of these, such as unevenness in receipt of invitations from Parliament, might be addressable through improved interventions. However, some reflect more fundamental weaknesses in the broader political system, most notably cynicism regarding the competence and motivations of politicians, which Parliament likely cannot address easily through technological innovation.

Our study also has broader implications for research methodology, and especially for the extent to which outcomes of scaled-up programs can be gleaned from results of controlled small-scale interventions. The literature on scaling up has largely focused on assessing the extent to which experimental estimates in one context apply in another. Some of this literature highlights the problems in using a small handful of studies as the basis for inferences to different contexts (Collaboration, 2015). Other work highlights the costs of extrapolation. Comparing non-experimental and experimental estimates that rely on the same data, Pritchett and Sandefur (2013) conclude that non-experimental estimates with the same subject population can better predict treatment effects as compared to experimental results from other contexts, becuase contextual variation can drive bigger differences in the estimated effectiveness of a program than selection bias.[19]

[19]There is a growing literature debating the tradeoffs associated with different approaches to generating out-of-sample predictions based on experimental data. Hotz et al. (2005) suggest using subject's observed characteristics as predictive of treatment effects independent of context. Gechter (2015) proposes a method that uses differences in outcome distributions for individuals with the same characteristics and treatment status in the original study and the context of interest to learn about unobserved differences across context.

Importantly, there should not have been significant differences in the subject population between the FFE and the NFE. The FFE was offered to a random sample of subjects from every constituency in Uganda, while the scaled-up program was offered to a random subsample of 186 constituencies, out of a total of 238.

Thus the difference we observe draw attention to a distinct problem, largely overlooked by the extant external validity literature: the external validity across the nuts and bolts of *interventions* and not necessarily across *populations*.[20] This kind of validity problem is especially critical when lessons from carefully controlled small-scale studies are intended to inform policies to be implemented at a larger scale. Our results provide a cautionary tale for researchers and policy makers seeking to make such claims.

In our analysis we identified several distinct reasons why outcomes of experiments may fail to replicate when brought to scale. These include already well-appreciated effects that relate directly to scale (see also Deaton (2010) on general equilibrium effects). In addition we highlight possible effects related to the changing partners involved when interventions are implemented at scale (see also Bold et al. (2013) on capacity and motivation of implementing organizations), and we identify differences related to details in the design between controlled interventions and interventions implemented in the political wild, of the form that may be relevant for other studies.

Ironically when design details matter, a first response is to resort to controlled conditions "to get those details right." This might be an appropriate approach when seeking to control for all factors but a manipulated parameter of interest, but one core lesson from our study is that the importance of those details may only became apparent once researchers' control is removed.

References

Allcott, Hunt, "Site Selection Bias in Program Evaluation," *Quarterly Journal of Economics*, 2015, *130* (3), 1117–1165.

Ashworth, Scott, "Electoral accountability: recent theoretical and empirical work," *Annual Review of Political Science*, 2012, *15*, 183–201.

[20]Some of our divergent findings do relate to endogenous changes in populations as a consequence of the factors outside the control of the research team. For example, one reason we do not find price effects in the scaled-up program similar to those found in the controlled experiment, can be attributed to the fact that the national intervention was taken up by relatively well off and engaged citizens who are unlikely to be sensitive to a small price subsidy.

Banerjee, Abhijit V and Esther Duflo, "The Experimental Approach to Development Economics," *Annual Review of Economics*, 2009, *1*, 151–78.

Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, Kartika Sari Juniwaty, and Bertil Tungodden, "Business training in Tanzania: From research-driven experiment to local implementation," *Journal of African economies*, 2012, *21* (5), 808–827.

Blair, Graeme, Rebecca Littman, and Elizabeth Levy Paluck, "Inciting Action Against Corruption in Nigeria," February 2015. Unpublished manuscript.

Bleck, Jaimie and Nicolas van de Walle, "Valence Issues in African Elections: Navigating Uncertainty and the Weight of the Past," *Comparative Political Studies*, November 2013, *46* (11), 1394–1421.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur, "Scaling up What Works: Experimental Evidence on External Validity in Kenyan Education," *Center for Global Development Working Paper No. 21*, 2013, (321).

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria et al., "The value of democracy: evidence from road building in Kenya," *The American Economic Review*, 2015, *105* (6), 1817–1851.

Callen, Michael and James D Long, "Institutional corruption and election fraud: Evidence from a field experiment in Afghanistan," *The American Economic Review*, 2014, *105* (1), 354–381.

Coleman, Kenneth M and Charles L Davis, "The Structural Context of Politics and Dimensions of Regime Performance: Their Importance for the Comparative Study of Political Efficacy," *Comparative Political Studies*, 1976, *9* (2), 189.

Collaboration, Open Science, "Estimating the reproducibility of psychological science," *Science*, 2015, *349* (6251).

Craig, Stephen C, Richard G Niemi, and Glenn E Silver, "Political efficacy and trust: A report on the NES pilot study items," *Political Behavior*, 1990, *12* (3), 289–314.

Croke, Kevin, Guy Grossman, Horacio A Larreguy, and John Marshall, "Deliberate disengagement: How education decreases political participation in electoral authoritarian regimes," *American Political Science Review*, n.d.

Dale, Allison and Aaron Strauss, "Don't Forget to Vote: Text Message Reminders as a MobilizationTool," *American Journal of Political Science*, 2009, *53* (4), 787–804.

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

Duflo, Esther, Rema Hanna, and Stephen P Rya, "Incentives work: Getting teachers to come to school," *The American Economic Review*, 2012, *102* (4), 1241–1278.

Gechter, Michael, "Generalizing the Results from Social Experiments: Theory and Evidence from Mexico and India," *Unpublished Manuscript*, 2015.

Green, Donald P and Alan S Gerber, *Get Out the Vote: How to Increase Voter Turnout*, 2nd edition ed., Brookings Institution Press, 2008.

Grossman, Guy and Kristin Michelitch, "Information Dissemination, Competitive Pressure, and Politician Performance between Elections: A Field Experiment in Uganda," *Unpublished manuscript*, 2016.

_ , _ , and Marta Santamaria-Monturiol, "Texting Complaints to Politicians: Name Personalization and Politicians' Encouragement in Citizen Mobilization," *Comparative Political Studies*, n.d.

Harrison, Glenn W. and John A. List, "Field Experiments," *Journal of Economic Literature*, Dec., 2004, *42* (4), 1009–1055.

Hotz, V Joseph, Guido W Imbens, and Julie H Mortimer, "Predicting the efficacy of future training programs using past experiences at other locations," *Journal of Econometrics*, 2005, *125* (1), 241–270.

Humphreys, Macartan, "Coalitions," *Annu. Rev. Polit. Sci.*, 2008, *11*, 351–386.

Isaksson, Ann-Sofie, Andreas Kotsadam, and Måns Nerman, "The gender gap in African political participation: Testing theories of individual and contextual determinants," *Journal of Development Studies*, 2014, *50* (2), 302–318.

Kasara, Kimuli and Pavithra Suryanarayan, "When Do the Rich Vote Less Than the Poor and Why? Explaining Turnout Inequality across the World," *American Journal of Political Science*, July 2015, *59* (3), 613—627.

Kramon, Eric and Daniel N. Posner, "Ethnic Favoritism in Education in Kenya," *Quarterly Journal of Political Science*, 2016, *11* (1), 1–58.

LeBas, Adrienne, *From Protest to Parties: Party-building and Democratization in Africa*, Oxford University Press, 2011.

LeVeck, Brad L, D Alex Hughes, James H Fowler, Emilie Hafner-Burton, and David G Victor, "The role of self-interest in elite bargaining," *Proceedings of the National Academy of Sciences*, 2014, *111* (52), 18536–18541.

Loewen, Peter John, Lior Sheffer, Stuart Soroka, Tamir Sheafer, and Stefaan Walgrave, "Are Politicians Better Decision Makers?," *Unpublished manuscript*, 2016.

Lynch, Gabrielle and Gordon Crawford, "Democratization in Africa 1990–2010: an assessment," *Democratization*, 2011, *18* (2), 275–310.

Malhotra, Neil, Melissa R Michelson, Todd Rogers, and Ali Adam Valenzuela, "Text messages as mobilization tools: The conditional effect of habitual voting and election salience," *American Politics Research*, 2011, *39* (4), 664–681.

Manski, Charles F, *Public Policy in an Uncertain World: Analysis and Decisions*, Cambridge, Mass.: Harvard University Press, 2013.

Niemi, Richard G, Stephen C Craig, and Franco Mattei, "Measuring internal political efficacy in the 1988 National Election Study.," *American Political Science Review*, 1991, *85* (4), 1407–1413.

Pritchett, Lant and Justin Sandefur, "Context matters for size: why external validity claims and development practice do not mix," *Journal of Globalization and Development*, 2013, *4* (2), 161–197.

Riedl, Rachel Beatty, *Authoritarian Oigins of Democratic Party Systems in Africa*, Cambridge University Press, 2014.

Stokes, Susan C, Thad Dunning, Marcelo Nazareno, and Valeria Brusco, *Brokers, voters, and clientelism: The puzzle of distributive politics*, Cambridge University Press, 2013.

van der Wind, Peter and Macartan Humphreys, "Crowdseeding in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time," *Journal of Conflict Resolution*, 2016, *60* (4), 748–781.

World Development Report

***World Development Report**, Washington DC: The World Bank, 2016.*

Yanagizawa-Drott, David, *"Propaganda and conflict: Evidence from the Rwandan genocide," The Quarterly Journal of Economics, 2014, 129 (4), 1947–1994.*

7 Appendix: Extra Figures and Tables

Figure 6: Dashboard screenshot

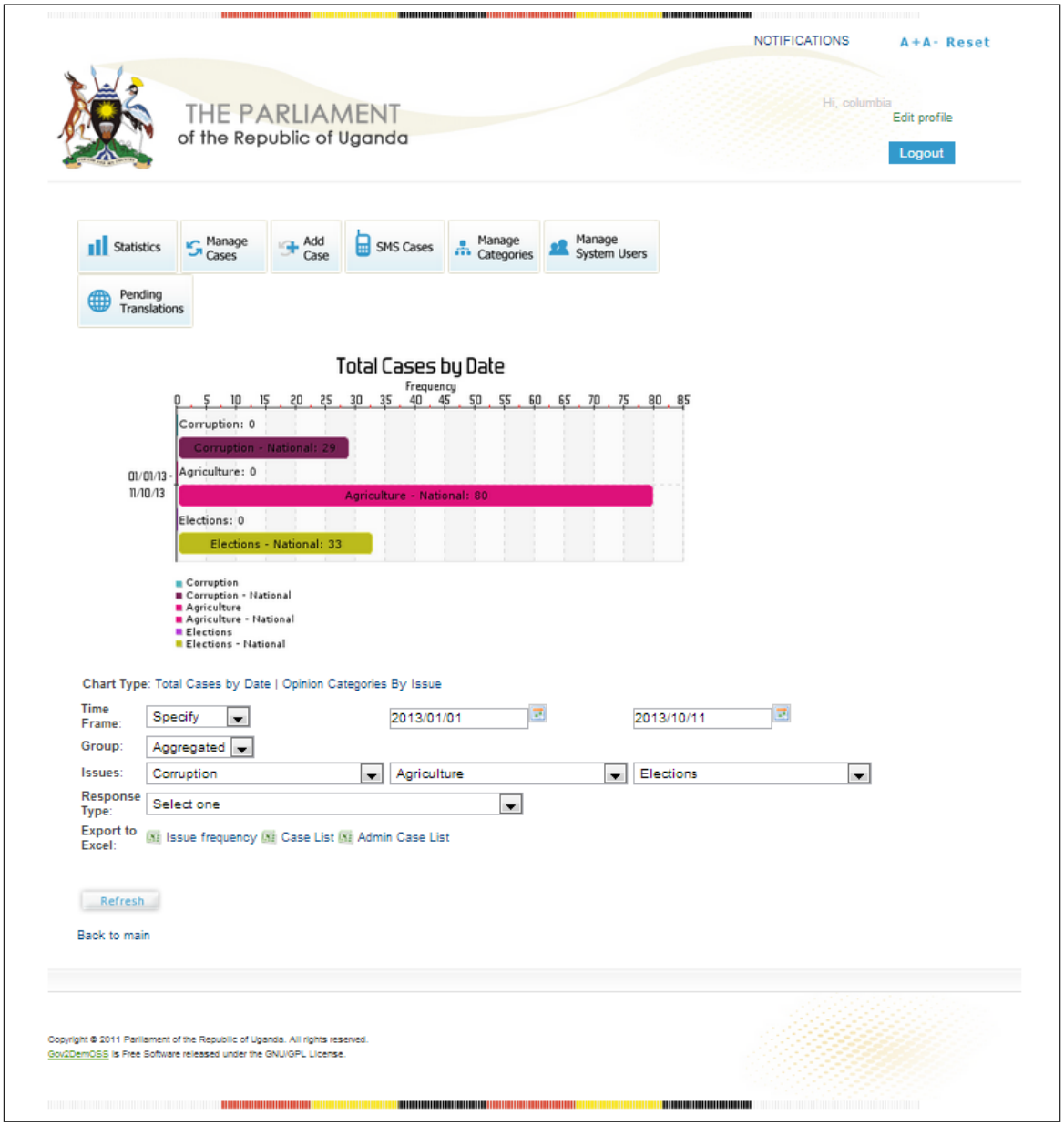


Figure 7: Treatment Scheme Example

District Code	District Name	District Type	MPs	Number of MPs	Month	Price	Feedback
1	APAC	MIXED	AJOK LUCY & AKORA MAXWELL EBONG PATRICK & AYOO TONNY	3	1	FREE	NONE
1	APAC	MIXED	AJOK LUCY & AKORA MAXWELL EBONG PATRICK & AYOO TONNY	3	2	FULL	NONE
1	APAC	MIXED	AJOK LUCY & AKORA MAXWELL EBONG PATRICK & AYOO TONNY	3	3	FULL	NONE
1	APAC	MIXED	AJOK LUCY & AKORA MAXWELL EBONG PATRICK & AYOO TONNY	3	4	FREE	SCHOOL + USERS
1	APAC	MIXED	AJOK LUCY & AKORA MAXWELL EBONG PATRICK & AYOO TONNY	3	5	FULL	NONE
1	APAC	MIXED	AJOK LUCY & AKORA MAXWELL EBONG PATRICK & AYOO TONNY	3	6	FREE	SCHOOL
2	ARUA	OPP	OKUONZI SAM AGATRE & WADRI KASSIANO EZATI	2	1	FULL	NONE
2	ARUA	OPP	OKUONZI SAM AGATRE & WADRI KASSIANO EZATI	2	2	FREE	NONE
2	ARUA	OPP	OKUONZI SAM AGATRE & WADRI KASSIANO EZATI	2	3	FULL	NONE
2	ARUA	OPP	OKUONZI SAM AGATRE & WADRI KASSIANO EZATI	2	4	FREE	SCHOOL + USERS
2	ARUA	OPP	OKUONZI SAM AGATRE & WADRI KASSIANO EZATI	2	5	FREE	NONE
2	ARUA	OPP	OKUONZI SAM AGATRE & WADRI KASSIANO EZATI	2	6	FULL	SCHOOL
3	BUNDIBUGYO	NRM	NTABAZI HARRIET	1	1	FULL	NONE
3	BUNDIBUGYO	NRM	NTABAZI HARRIET	1	2	FREE	NONE
3	BUNDIBUGYO	NRM	NTABAZI HARRIET	1	3	FREE	NONE
3	BUNDIBUGYO	NRM	NTABAZI HARRIET	1	4	FULL	SCHOOL
3	BUNDIBUGYO	NRM	NTABAZI HARRIET	1	5	FULL	NONE
3	BUNDIBUGYO	NRM	NTABAZI HARRIET	1	6	FREE	SCHOOL + USERS
5	GULU	MIXED	OULANYAH JACOB L'OKORI & ACIRE CHRISTOPHER	2	1	FREE	NONE
5	GULU	MIXED	OULANYAH JACOB L'OKORI & ACIRE CHRISTOPHER	2	2	FULL	NONE
5	GULU	MIXED	OULANYAH JACOB L'OKORI & ACIRE CHRISTOPHER	2	3	FREE	NONE
5	GULU	MIXED	OULANYAH JACOB L'OKORI & ACIRE CHRISTOPHER	2	4	FULL	NONE
5	GULU	MIXED	OULANYAH JACOB L'OKORI & ACIRE CHRISTOPHER	2	5	FREE	SCHOOL
5	GULU	MIXED	OULANYAH JACOB L'OKORI & ACIRE CHRISTOPHER	2	6	FULL	SCHOOL + USERS
6	HOIMA	NRM	KAAHWA TOPHACE BYAGIRA	1	1	FULL	NONE
6	HOIMA	NRM	KAAHWA TOPHACE BYAGIRA	1	2	FREE	NONE
6	HOIMA	NRM	KAAHWA TOPHACE BYAGIRA	1	3	FULL	NONE
6	HOIMA	NRM	KAAHWA TOPHACE BYAGIRA	1	4	FREE	SCHOOL + USERS
6	HOIMA	NRM	KAAHWA TOPHACE BYAGIRA	1	5	FULL	SCHOOL
6	HOIMA	NRM	KAAHWA TOPHACE BYAGIRA	1	6	FREE	NONE