Does Information Technology Flatten Interest Articulation? Evidence from Uganda

Guy Grossman* Macartan Humphreys† Gabriella Sacramone-Lutz‡

November 10, 2012

Abstract

This is a mock article: it provides the structure of analysis prior to examining real data on outcomes. All outcome data used in this mock article are simulated and do not represent real-world data in any way. We note that the mock analysis was conducted after data collection but before analysis.

We use a field experiment to study the effects of access to channels of political communication on the efforts voters take to influence their representatives. Using a custom information communication technology (ICT) platform we presented randomly sampled constituents in Uganda with an opportunity to send a text-message to their MP at one of three randomly assigned price levels. In assessing who chose to contact their representatives and what they chose to communicate, we seek to assess (a) whether ICTs can "flatten" interest articulation (or whether they amplify the claims of those already lobbying their representatives) and (b) how the volume and the content of messages sent depends on price. Of particular interest is the possibility that increasing the cost of communication results in disproportionate drops in participation by already marginalized subpopulations as well as changes in the content of what gets communicated—from more public concerns to more private concerns.

*University of Pennsylvania. We thank Heather Kashner and Simon Osborn at NDI that have supported this project. We thank the Deepening Democracy Programme, the International Growth Centre, and Innovations in Poverty Action for support for data collection and the Trudeau Foundation for support during the analysis phase. We thank especially Nicolas de Torrente, Robert Sentamu, and Doug Parkerson for their support on the survey. We thank Jeremy Weinstein, Ivan Tibemanya, Kosuke Imai, and Don Green for comments and advice on various parts of this project.

†Columbia University

‡Columbia University

1 Introduction

The quality of democratic institutions as a tool of political representation depends on interest articulation: the opportunity and willingness of voters to communicate their needs and preferences to their representatives in government. What politicians think and what they prioritize depends in part on what they are hearing from constituents. But in many low-income countries, voters typically do not communicate with their MPs with any regularity. The bulk of communication is unidirectional, clustered around election periods, and tapers off during legislative periods. Moreover, there is inequality in who can access politicians: men are more likely to have access to politicians than women, wealthier constituents are more likely to have access than poor constituents and so on.

We study how the availability and cost of communication with politicians affects who and what gets communicated. To do so we employ an experimental approach in which we provide a random set of constituents with access to an SMS based messaging system which allows them to send messages directly to their MPs. The introduction of the system to a random sample of voters allows us to assess how the representativeness of political communication depends on access. As part of the experiment we randomly vary the price for using of the system in order to assess theories of how the cost of access determines who participates and what they seek to communicate.

2 Motivation and Hypotheses

Though democratic forms of governance are now common in many low-income countries, the quality of elected governments, as measured by corruption levels and quality of public services, is low. One leading explanation— together with the dominance of clientelism (Banerjee and Pande, 2009) and the weakness of electoral institutions (Hyde, 2007) — is the effect of information deficits. Whereas studies of the determinants of good governance overwhelmingly focus on the lack of information in the hands of citizens (Pande, 2011), the lack of information in the hands of politicians may be just as important.

With limited information on the priorities and preferences of citizens, politicians have little ability to serve as representatives, and parties may be less likely to differentiate themselves based on policy-positions (Wantchekon and Fujiwara, 2011). Moreover, politicians may have little incentive to act in a representative manner when they know that their constituents know that they have such poor information (Ashworth, 2012). Instead, politicians who are unable to assess public opinion may be more likely to respond to the demands of powerful interest groups (Bartels, 2008) or serve as rubber-stamps for the executive branch (Barkan, 2009). This logic provides one explanation for why African parliaments are generally considered weak vis-à-vis the executive (Barkan et al., 2010). Critically, communication structures can affect not just how much information politicians receive but also whose voice gets heard and what issues people focus on. As in rich industrialized countries, there is inequality in who can access politicians and the political process.

In this study we focus on two levels of political empowerment: (a) *access*—the extent to which channels exist to communicate with politicians if need or want arise, and (b) *engagement*—the extent to which individuals participate in political processes.

We operationalize both *access* and *engagement* by grouping a number of related measures into a summary index (see discussion at Anderson (2008)). The summary index is a weighted mean of several standardized outcomes, where the weights—the inverse of the covariance matrix of standardized variables—are used to maximize the amount of information captured by the index. For some analyses we divide the population into groups by dichotomizing the summary indices using the median as cutoff point. Data used for constructing the access and engagement indices come from original survey data gathered by the research team in 2011.

We operationalize *access* to existing communication channels using 5 variables that capture technologies already available to individuals that could be used to contact politicians (see also top panel of Table 1): (1) an indicator of respondent's access to a phone call; (2) a continuous measure capturing the frequency of SMS usage and (3) an indicator of respondent's access to a computer. In addition we have two measures of physical barriers to connecting with politicians: (4) an indicator of whether the respondent travels ten kms or more from the place where they live now, at least a few times a month; and (5) a continuous variable measuring the geodetic distances from the respondent's home to the district capital.[1] With the exception of distance, all access variables are positively correlated.

We emphasize that our index captures existing access *technologies* available to voters and does not capture social channels such as family or ethnic ties.

In Table 1 we illustrate our operationalization of the access index and how it, and each of its constituent variables, relate to traditional indicators of marginalization in African politics: poverty, sex (female and cogender with MP), and ethnicity (being a non co-ethnic of one's MP). On most measures (except coethnicity) marginalized voters are significantly less able to access their representative through existing channels of communication.

We operationalize *engagement* using 9 indicator variables. These measures, which appear in Table 2, include: (1) active membership in any political party, (2) membership in the village governance committee, (3) attending a community meeting several times in the past year, (4) raising political issue with others at least once in the past year, (5) attending demonstrations and protest marches at least once in the past year, (6) attending elections rallies at least once in the past year, (7) writing letters to a newspaper or calling

[1]The distance measure was multiplied by -1 such that the smaller the number the larger is the distance to the district capital.

Table 1: **Access to Existing Communication Channels**

	(1) Phone Access (q44)	(2) SMS Access (q46)	(3) Computer Access (q38b)	(4) Travel outside village (q38d)	(5) Distance: district capital (GIS)	(6) Access Index (1-5)
Poor (subj)	0.84	0.76	0.05	0.66	0.46	0.43
Non-poor	0.9	0.85	0.13	0.67	0.54	0.57
Difference	0.06*	0.09*	0.09*	0.01	0.07*	0.14*
Poor (obj)	0.8	0.72	0.02	0.61	0.46	0.4
Non-poor	0.94	0.9	0.15	0.72	0.52	0.6
Difference	0.14*	0.18*	0.13*	0.11*	0.05*	0.20*
Female	0.84	0.77	0.06	0.6	0.5	0.43
Male	0.89	0.85	0.12	0.72	0.5	0.57
Difference	0.05*	0.08*	0.05*	0.12*	0	0.13*
Noncogender	0.86	0.80	0.08	0.64	0.50	0.49
Cogender	0.87	0.82	0.10	0.68	0.50	0.51
Difference	0.01	0.02*	0.02*	0.04*	0	0.02*
Noncoethnic	0.87	0.8	0.09	0.66	0.55	0.53
Coethnic	0.87	0.81	0.08	0.67	0.48	0.49
Difference	0	0.01	-0.01	0.01	-0.07*	-0.04*

Generated using survey data. $*p < 0.05$. *Poor (subj)* refers to a respondent's subjective relative measure of poverty, whereas *Poor (obj)* denotes an asset-based poverty index dichotomized using the median as cutoff.

a radio show at least once in the past year, (8) voting in recent parliamentary elections, (9) attending at least one MP organized meeting in the past year, and (10) personally talked to one's MP in the past year. We then use these variables to construct a summary index of political "engagement", which appears in the last column.

Table 2: **Politically Engaged**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Party Member (q52)	Village Committee (q53a)	Community Meetings (q55a)	Raise Issue (q55b)	Protest March (q55c)	Election Rally (q55d)	Write Letter (q55e)	MP meeting (q56)	Voted MP elections (q66)	Talked to MP (q109)	Engaged Index (1-10)
Poor (subj)	0.18	0.23	0.57	0.63	0.06	0.77	0.09	0.25	0.94	0.20	0.46
Non-poor	0.24	0.26	0.52	0.65	0.07	0.79	0.14	0.29	0.95	0.27	0.52
Difference	0.06*	0.03*	-0.05*	0.02	0.01	0.02	0.05*	0.04*	0.01	0.07*	0.07*
Poor (obj)	0.16	0.21	0.52	0.61	0.05	0.74	0.07	0.23	0.93	0.22	0.42
Non-poor	0.25	0.28	0.56	0.67	0.07	0.81	0.17	0.31	0.95	0.25	0.56
Difference	0.08*	0.07*	0.04*	0.06*	0.02*	0.06*	0.10*	0.08*	0.02*	0.03*	0.14*
Female	0.15	0.20	0.48	0.58	0.05	0.73	0.08	0.23	0.93	0.17	0.39
Male	0.26	0.29	0.60	0.71	0.07	0.82	0.15	0.32	0.95	0.29	0.58
Difference	0.11*	0.09*	0.12*	0.13*	0.02*	0.09*	0.08*	0.09*	0.02*	0.12*	0.19*
Noncogender	0.19	0.23	0.52	0.62	0.05	0.76	0.11	0.25	0.94	0.19	0.45
MP Cogender	0.22	0.26	0.57	0.66	0.07	0.79	0.13	0.29	0.95	0.27	0.52
Difference	0.03*	0.04*	0.05*	0.04*	0.02*	0.02*	0.02*	0.04*	0.01*	0.08*	0.07*
Noncoethnic	0.21	0.26	0.51	0.63	0.08	0.76	0.13	0.27	0.93	0.24	0.49
MP Coethnic	0.20	0.24	0.56	0.65	0.05	0.79	0.13	0.27	0.95	0.23	0.49
Difference	0.01	-0.01	0.05*	0.02	-0.03*	0.03	-0.02*	0	0.02*	-0.01	0
Long distance	0.21	0.25	0.58	0.65	0.05	0.79	0.11	0.28	0.95	0.23	0.50
Sh. distance	0.20	0.24	0.51	0.64	0.07	0.76	0.12	0.26	0.93	0.24	0.48
Difference	0.01	-0.01	-0.07*	-0.01	0.02*	-0.03	0.01	0.01	-0.02*	0.01	-0.02

Generated using survey data. $*p < 0.05$.

Our survey data suggest that, at least in Uganda, there exist large and significant differences between the poor and non-poor, and male and female respondents with respect to both political access and political engagement. The difference between non-cogender and cogender respondents is somewhat smaller yet significant at the 95% level. However, and in contrast to classic accounts of the political economy of African development, neither access (defined in terms of technologies of access) nor political engagement is structured around ethnic lines (see Table 1 and Table 2). Note that these patterns contrast somewhat with patterns found in the US, which show little effects of gender on political participation once education is controlled for (Nam, 2010).

How do these measures relate to each other? We find a weak positive correlation between the access and engaged indices (0.09). The fact that access is only weakly correlated with engagement motivates the possibility that opening access though ICT can empower voters who do not participate in traditional forms of political engagement.

In this project we seek to assess how technologies for political communication alter these patterns of representation. To do so we draw on two literatures that provide insight into how political communication structures the behavior of constituents and politicians.

2.1 Technology, political participation and democratic consolidation

In recent years there has been a growing interest in the role technology might play in improving political communication (Balkin, 2004, Bimber, 2001). Information Communication Technologies (ICTs) can potentially contribute to democratic processes by facilitating group interaction and rapid accumulation and dissemination of information. ICTs can also allow citizens to engage in debate on political matters, and become familiar with opinions and events that affect their communities (Oates, 2003). Indeed by some accounts, access to ICTs likely has a causal effect on national levels of democracy (Shirazi, Ngwenyama and Morawczynski, 2010).

Given these potential benefits there has been a recent surge in innovations to exploit ICTs to enlarge access to politics. In Africa alone innovations include the Africa Technology and Transparency Initiative (ATTI) and the African Electronic Governance for Research Initiative. The growing role of ICTs in political communication, nevertheless, raises important questions about whether ICT initiatives can genuinely alter representative-constituent relation and whether ICTs are increasing access to marginal groups. Thus there is a concern that the groups that have the weakest access to political processes are also the least likely to access and use ICT systems. Whether the overall effect will mean greater or lower representation of the preferences of marginal groups is an important question that this study seeks to address. As noted by Thompson (2008), "while ICTs may enlarge access to political communication, they have the power to create new inequities, as well as exacerbate existing ones."

Focusing on the alleged technological gender divide, Hafkin and Huyer (2007) find that women in low-income countries are facing substantial disadvantages since they are significantly less likely than men to use ICTs. Similarly, Park (2009) finds a gender divide in developing countries, which applies not only to access but also to the frequency of usage. By contrast, Hilbert (2011), analyzing recent data from 12 Latin American and 13 African countries, finds that the reason why fewer women access and use ICT is a result of their unfavorable conditions with respect to employment, education and income. These findings suggest that there exists a pressing need to critically assess the case for "technological optimism" in the area of politics and governance. Whether ICT innovations can play a role in facilitating good governance and whether they increase or decrease political access to marginal populations are key questions our study is designed to address.

The core question we focus on from this literature is a simple one: Does the introduction of an ICT system result in representative information on constituency needs and preferences? We assess this question by focusing on two hypotheses, one regarding the representativeness of user demographics and one regarding the representativeness of user preferences.

H_1 **Flattening:** the level of communication through ICTs, relative to traditional forms of political communication, is greatest for marginalized groups (such as women and the poor) with more limited political access.

H_2 **Representation:** The priority issues for ICT users are closer to those of the general population than are those of groups exhibiting high levels of political engagement via traditional communication channels.

2.2 The cost of access and the demand for communication

How does the cost of access affect the level of political participation? We generally expect that offering a valuable service at a lower cost will increase demand. Somewhat counter-intuitively, however, offering free goods and services, can sometimes reduce demand. One reason for this is that costly options may invoke market exchange norms, whereas free products and services may invoke norms of social exchange (Heyman and Ariely, 2004). Shampanier, Mazar and Ariely (2007) show that when offered candies at a cost of 1c per piece, students take about four pieces, but when the price is set to zero, more students take candy, but almost no one takes more than one piece. Applying this logic to our context, there is a possibility that offering free messaging would result in broader usage without necessarily generating more messages (because people elect to use the system only for "good reasons").

Another possibility is that a reduction in costs can increase the number of potential participants and produce a collective action dilemma. This might arise for example if the messages for public goods act as strategic substitutes. In the example shown in Figure 1 when messages are cheap, both rich and poor would be individually willing to send a message for a public good, but each would rather the other sends. This produces a coordination problem. In the mixed strategy equilibrium each sends with a 1/3 probability and there are 2/3 messages sent in expectation. With higher prices if the perceived cost of the message is greater for poor than for rich players, then this can result in a unique equilibrium and a resolution of the coordination problem.

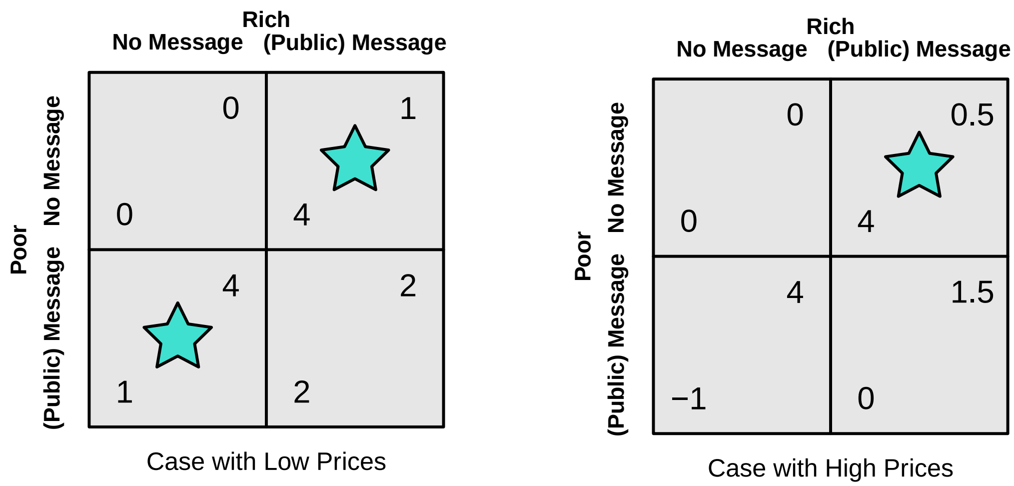


Figure 1: Higher prices remove the coordination dilemma. Pure strategy Nash Equilibria are marked with stars. The payoffs for the poor are 2 points lower on the right whenever messages are sent and 0.5 lower for the rich. The symmetric Nash equilibrium on the left has each send messages with probability 1/3 for an expected 2/3 messages sent.

This discussion highlights the fact that in strategic environments the slopes of demand curves are not obvious. Although strategic and behavioral logics suggest mechanisms through which a price increase could increase messaging, we expect that political access acts like a normal good and satisfies the law of demand. We state this as our third hypothesis.

H_3 **Demand:** Less expensive communication results in greater levels of communication.

Content: Cost may affect not only the "type" of people contacting their elected representative, but also the content of messages sent by given types. Citizens may be relatively more likely to send messages with public goods content (rather than with demands for private goods) when prices are low. The core insight is that when there is no cost, one can expect many others to contact their representative. In this case, the marginal benefits from seeking private goods, for which there is substitution, declines relative to the marginal gains from seeking public goods, from which there can be complementarities. When the cost of sending messages is high, senders may assume that competition over the resources of the politician is relatively small, and hence it is relatively more prudent to request private goods.

We illustrate the core logic using a simple full information normal form game presented in Figure 2. Consider two players deciding whether to send a message for a public good, a message for a private good, or no message at all. Say one player is rich and the other poor and payoffs have a structure like that shown below. The key features are (a) that there are strategic complementarities in public messaging but substitution in private messaging and (b) the poor player is more sensitive to the monetary cost of messaging than the rich player. In this game when costs are low there are multiple equilbria but the equilibrium involving public messaging Pareto dominates equilibria from the chicken-like game that is induced by the decision to engage in private messaging. When costs are high however there is a unique equilibrium in which the rich player sends a private message.

In this case a rise in prices produces two effects: an overall decline in messaging and a shift (for the wealthy player) from public to private messaging.

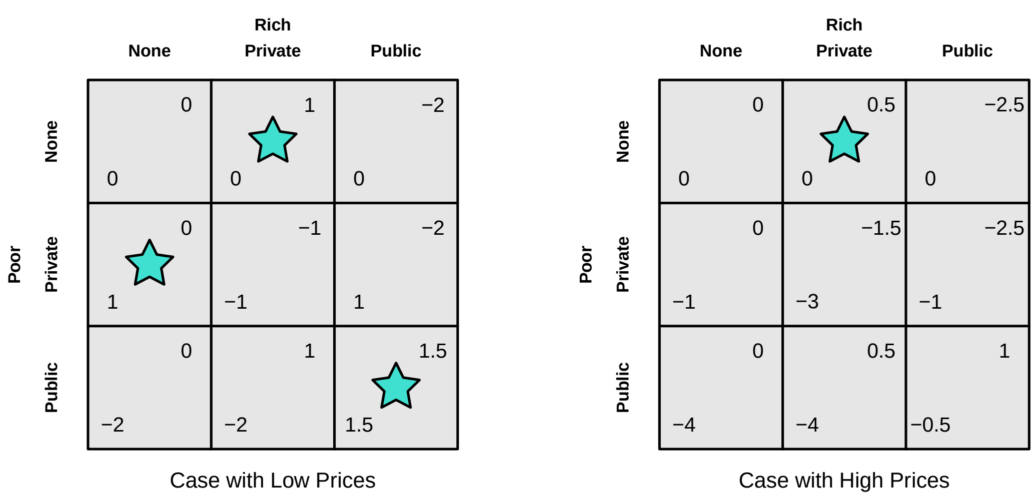


Figure 2: Higher prices remove public messaging equilibrium. Pure strategy Nash Equilibria marked with stars. The payoffs for the poor are 2 points lower on the right whenever messages are sent and 0.5 lower for the rich.

Though price variation has been examined in the context of consumer goods and health products, to our knowledge, this is the first study to experimentally vary the cost of contacting one's representative in parliament. If price variation affects both the type of people sending messages as well as the content of the messages, then different price schemes can affect the representativeness of the IT-based communication platform. This will be the case, for example, if people who cannot afford to pay a positive price have different needs and priorities to those that can afford to make use of the system. In the end, the relative benefits of various levels of subsidization of communication systems depend on (1) the elasticity of demand with respect to price, (2) the variation in preferences and priorities of public goods as a function of income, and (3) the number of messages that need to be reached in order to induce a representative to action. To the extent that the representativeness of the ICT platform may, in turn, impact MPs' attitudes and behavior–understanding the impact of the cost of messages (i.e., cost of contacting one's MP) is of theoretical and policy relevance.

The third core question on price effects that we examine is then: does the cost of sending messages affect the nature of political information sent? As noted above, we expect price effects to be stronger for poor people and voters with alternative channels of communication. Moreover, when the cost of contacting one's representative decreases, we expect citizens to be more likely to send public requests rather than private ones.

H_4 **Content Filtering:** More expensive communication results in greater focus on private goods issues rather than public issues.

Our discussion suggests that the overall price effects may result in part from different effects for different subpopulations and in particular that poorer populations may be more sensitive to higher prices. Data gathered in 2010 in preparation for this study also suggests that such patterns are likely to hold in Uganda. Prior to the implementation of this study, the National Democratic Institute (NDI) conducted a small pilot in four Ugandan constituencies, using marketing teams to examine the willingness of survey respondents to send a text-message to their MP as a function of hypothetical prices. Figure 3 illustrates two key patterns. First, there is evidence of a strong sensitivity to (hypothetical) prices, suggesting that price can be a major deterrent to uptake at all income levels. Second, the reported effect of prices was different for different socio-economic groups: less educated and poorer constituents reported less willingness to pay for sending a text-message to their MP. Over 70% of respondents at all income-level groups, apart from "Much worse" express willingness to use the system at a cost of 100 USH. For the "Much worse" group the figure is 52%. Critically these patterns suggest that higher prices likely generate messaging that is more reflective of the needs and preferences of wealthier constituents.

A similar logic may hold for individuals with alternative channels of access. Individuals that are otherwise *less* marginalized may also be more sensitive to prices since they have the option to substitute to more traditional channels when prices rise.

Though striking, the pilot presented respondents with hypothetical prices, leaving open

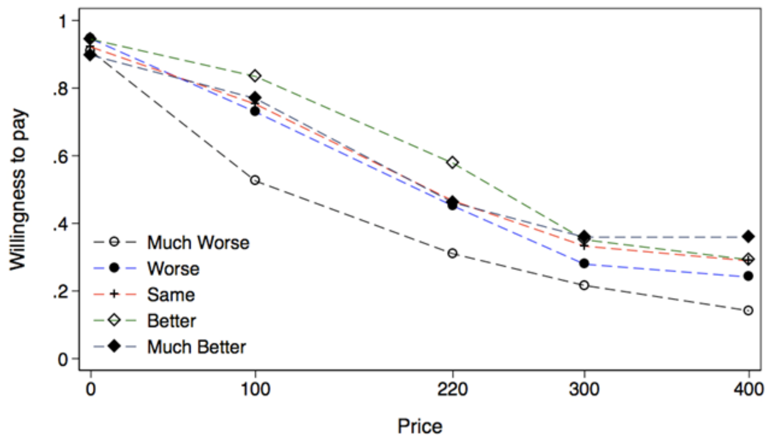


Figure 3: Price Sensitivity by Relative Economic Position. Based on pilot data. At low prices all respondents claim to be willing to contact MPs. At higher prices better-off respondents are more than twice as willing as worse-off respondents (NDI, 2010).

the question: how does the price of political communication affect the population of views communicated? We assess the question in terms of heterogeneous demand effects.

H_5 **Heterogeneous price effects:** The price effect will be stronger for (a) poorer constituents and (b) constituents with alternative channels of access to politicians

Table 3 summarizes the five hypotheses under examination; the next section describes how we seek to assess these hypotheses.

Table 3: **Hypotheses Summary**

#	**Effect type**	**Hypothesis**
H_1	Observational Hypothesis 1	Flattening: the level of communication through ICTs, relative to traditional forms of political engagement, is greatest for marginalized groups
H_2	Observational Hypothesis 2	Representation: The priority issues for ICT users are closer to those of the general population than are those raised by traditional high engagement groups.
H_3	Price effects: *Level*	Demand: Less expensive communication results in greater uptake across all groups.
H_4	Price effects: *Content*	Filtering: More expensive communication results in greater focus on private rather than public issues
H_5	Heterogeneous effects	Heterogeneous price effects: The price effect will be stronger for (a) poorer constituents and (b) constituents with alternative channels of access to politicians.

Note: Summary of hypotheses on the effects of the introduction of ICT based access to politicians.

3 Data

In this section we describe the key data to be employed. We begin by focussing first on balance of the treatment variable and second on the core outcome variables.

3.1 Research Site and Sampling Scheme

Beginning in late April 2011, the research team led a group of Ugandan researchers in conducting interviews with randomly sampled respondents in each of Uganda's 238 electoral constituencies. Cluster randomized sampling was used to select 4 villages in distinct sub-counties within each constituency. Within each village we conducted interviews with 8 respondents, 4 of whom were offered the chance to SMS their MP at a price level which was randomly pre-assigned. Assignment guaranteed balance across treatment price groups within each constituency. The service was introduced with the following script:

> We would like to offer you an opportunity to send your new MP a message using SMS. It is a chance to tell your incoming MP about issues that are important to you, or things you feel he/she should work on. This service is not associated with any political party or government agency. The service is (free/50sh/100sh).

In addition to this price variation a second variation was introduced in which a random set of respondents were read examples of public goods messages collected during the NDI pilot in order to assess whether messaging is subject to complementarities. This second treatment is not the subject of the present analysis.

3.2 Balance in Treatment Assignment

We use five variables to test pre-treatment covariate balance across the three treatment groups: subjective measure of wealth (binary)[2], objective measure of wealth (binary)[3] gender (binary), age (continuous) and education (ten category variable). In addition, we

[2] To generate a 'subjective' measure of wealth, survey respondents were asked to place themselves on a five-category wealth scale using the following question: "How would you compare your overall economic situation to those of other Ugandans?" Respondents were coded 1 if placed themselves at the lowest category ("much lower," 51% of respondents), and 0 otherwise.

[3] To generate a measure of 'objective' wealth we use a battery of questions that capture the respondents' purchasing power. These include: (a) number of radios owned by members of the household; (b) number of televisions; (c) number of bicycles; (d) number of motor vehicles; (e) number of cell-phones; (f) a computer; (g) years of education; (h) material used in the construction of the respondent's house; (i) access to clean drinking water; (j) time to nearest protected water source; (k) whether the respondent has a job that pays a cash income; and (l) monthly income estimate. All variables were grouped into a summary index as in Anderson (2008). We then dichotomized the scale using the median as cutoff point.

test the balance of the three continuous political participation indices: access, engagement and influence. In Figure 4, for each of the eight variables we provide (a) standardized mean deviations by treatment, which allows us to use a similar scale for all covariates (row 1), and (b) the full distribution of the variables in their original scale by treatment status (rows 2-4).

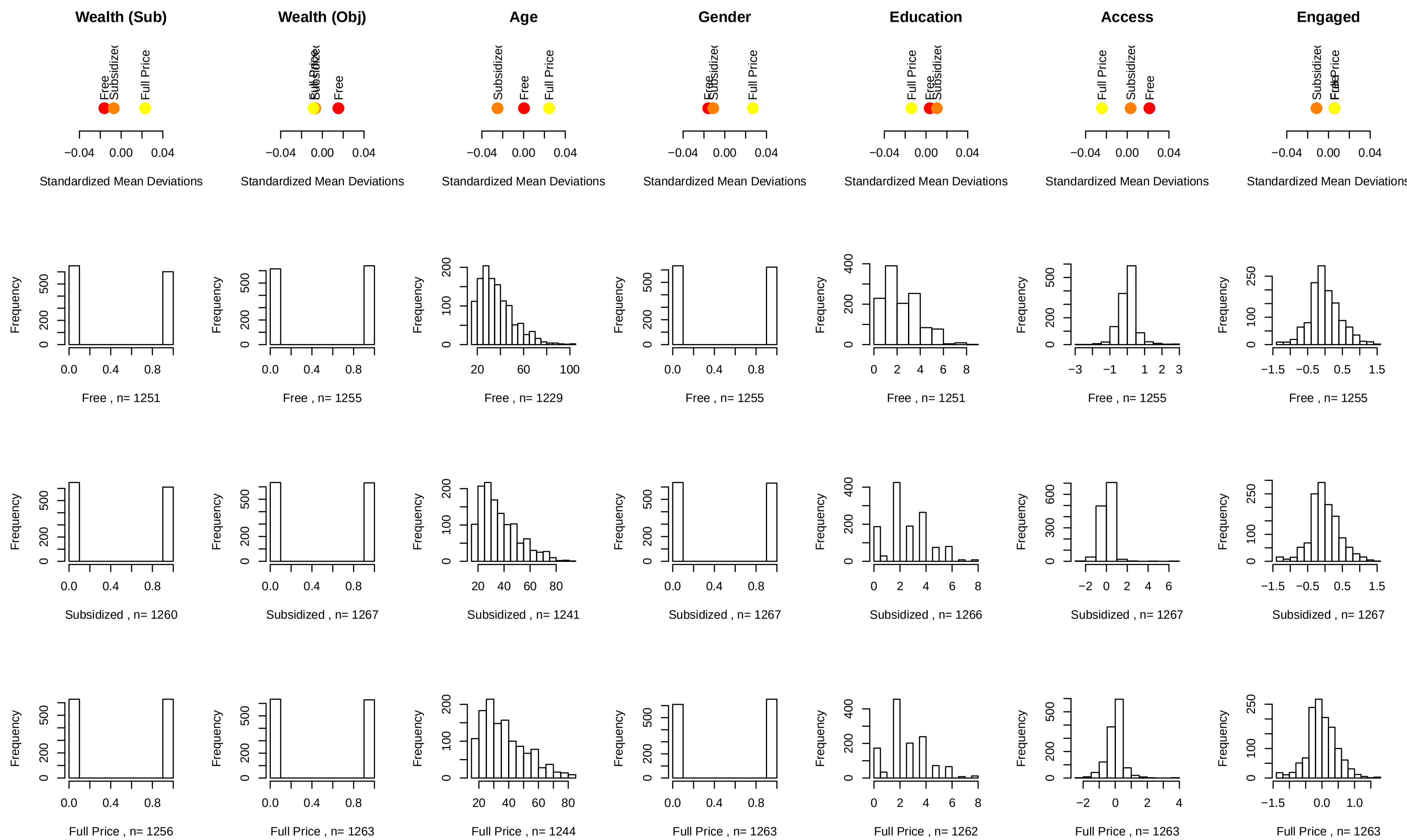


Figure 4: The distribution of key pretreatment covariates broken down by each price range (treatment groups). The top row shows means for each group in units of standard deviation of the covariate in question. See code in Appendix 7.4

3.2.1 Outcome Variables

Our two key estimands are (a) the propensity to send any message and (b) the propensity to send messages of a given type ("private" or "public"). We begin with a classification of messages into five types as follow:

1. Private/personal messages: Messages intended to benefit the individual sender or their family only. Examples include senders asking directly for money for home construction, school fees, funerals, or support finding work.

2. Local geographic group benefit messages. Message intended to benefit a geographic group below the constituency level. The group can be defined by gender, location ("village residents") and profession ("farmers"). Note that messages that do not indicate a particular subgroup within the constituency, but that say "we need" are treated as local geographic group benefit for the purposes of coding. Examples include messages on the need for electricity in a village, or the need for infrastructure or equipment in subcounties.

3. Large geographic group messages. Messages that make requests or provide information on behalf of geographic groups corresponding to the constituency or district.

4. Large non-geographic group messages. Message intended to benefit a group such as women, veterans, and farmers, beyond the constituency level.

5. Public messages. Messages with an unrestricted beneficiary group. This category includes items such as corruption, inflation, the institution of third terms, cost of living and other features related to national laws and policies.

This more detailed measure is then collapsed into a simple binary measure of public messaging (such that 0 refers to a private message and 1 refers to a public message) by selecting a cutoff point that maximizes the variance of the measure.

4 Mock Tests

4.1 Test of Hypothesis 1: Flattening

Table 1 and Table 2 demonstrate clearly that marginalized groups, especially the poor and women, are significantly less engaged in political life across a broad range of political participation indicators, compared to the non-poor and to male. For example, compared to their male counterparts, women are about 11% less likely to be members of political parties, 12% less likely to attend community meetings and 10% less likely to report writing a letter to a newspaper. As mentioned above, the findings, however, are more ambiguous with respect to respondents who are non-coethnics of their respective MPs. Does the opening of a new IT-based communication channel flatten access to representatives in the national assembly?

The first hypothesis we examine is that the composition of SMS message senders is more heavily weighted towards politically marginalized groups than is typical for other types of access to politics. Note that this hypothesis is tested using observational rather than experimental data. To assess this hypothesis, we compare the share of marginalized populations among SMS message senders to their share in those who take part in other political activities, as captured in the political engagement summary index. Results are presented in Table 4. The significance of the difference in the share of vulnerable groups using SMS versus the share using existing forms of political participation is the probability that the difference in the coefficients exceeds the critical chi-square value.

Table 4: **Flattening Participation: Test of H_1**

1	Share of poor respondents (subj) among engaged type		0.48
	Share of poor respondents (subj) among SMS-based access population		0.52
	Difference	H_1	0.04
			(0.013)**
2	Share of poor respondents (obj) among engaged type		0.43
	Share of poor respondents (obj) among SMS-based access population		0.46
	Difference	H_1	0.03
			(0.065)*
3	Share of women respondents among engaged type		0.4
	Share of women respondents among SMS-based access population		0.26
	Difference	H_1	-0.14
			(0.000)***
4	Share of non-coethnic respondents among engaged type		0.31
	Share of non-coethnic respondents among SMS-based access population		0.31
	Difference	H_1	0.00
			(0.958)
5	Share of non-cogender respondents among engaged type		0.46
	Share of non-cogender respondents among SMS-based access population		0.43
	Difference	H_1	-0.03
			(0.212)
6	Share of distant respondents among engaged type		0.51
	Share of distant respondents among SMS-based access population		0.52
	Difference	H_1	0.01
			(0.546)

Note: p value from χ^2 test in parenthesis. *Data on SMS usage is simulated.*

4.2 Test of Hypothesis 2: Representation

That the share of politically marginalized respondents among SMS-users is larger than their share among engaged types, by itself does not tell us whether the priority issues for ICT users are closer to those of the general population than are those raised by traditional high engagement groups. To test whether ICTs have a potential of being more representative than existing traditional political communication channels we first elicited the political priorities of our sample of constituents. Table 5 provides information on the top priorities of all respondents against the priorities of two groups of interest: groups that are traditionally highly engaged and groups that employ the SMS system. We also report the distribution of priority categories as captured by incoming SMS messages although we note that we do not have a measure of the typical subject of political communication with which to compare this quantity.

Table 5: **Patterns of Representativeness of Messaging**

	All subjects	Engaged subjects	SMS senders
Issue A	0.24	0.17	0.25
Issue B	0.28	0.25	0.25
Issue C	0.23	0.27	0.26
Issue D	0.25	0.31	0.24

Generated using mock data representing 4 potential issues.

To test H_2 we construct a measure of the non-representativeness of preferences of engaged constituents relative to preferences of the population and a measure of the non-representativeness of SMS senders and compare these two quantities. Our "non-representativeness statistic" (NR) measures the distance between the distribution of responses from subpopulation A and subpopulation B (not necessarily distinct) as:

$$NR(A,B) = \frac{1}{2}\sum_{k=1}^{m}(\alpha_k^A - \alpha_k^B)^2$$

where α^j denotes the vector of share of members of group j selecting different options. The NR statistic is calculated then as half the sum of squared deviations of shares in each of m categories. The maximum deviation is 1, which would arise if the message sending group all valued one area but others valued another.

To estimate the significance of differences in non-representativeness across groups we set the values of the non-SMS sending non-engaged group as the references distribution, we then estimate a multinomial logit model of sector choice as a function of group membership and using the estimated distribution of parameters, simulate a distribution of NR statistics relative to the (fixed) reference distribution as well as a distribution of differences in NRS's relative to the reference distribution both for non-SMS (and non engaged) populations and non-engaged (but SMS sending) populations. The significance test for the NRS of each group is calculated relative to the estimated distribution of NR statistics we would expect in the reference group (relative to the reference distribution), due to sampling error alone.[4]

The results of this analysis are illustrated in Figure 5.

4.3 Main Effects (H_3 and H_4)

We now turn to test H_3, namely that less expensive communication results in greater levels of communication. We begin with a summary presented in Table 6, of uptake, our key outcome variable, by price category.

Table 6: **Message Send Rate as a Function of Price**

	Public SMS	Private SMS	Any SMS
Free	0.23	0.27	0.50
Subsidized	0.20	0.19	0.39
Full	0.23	0.09	0.32

Note: Uptake is based on mock data simulated for the purpose of this mock article. Number of observations in parenthesis.

Table 6 shows the incidence of messaging by treatment condition (i.e., price) but it does not provide formal tests of H_3 and H_4. To test these hypotheses and estimate treatment effects we first create three treatment dummies:

T_{0S} equals zero for those assigned to free price and one for those with subsided price

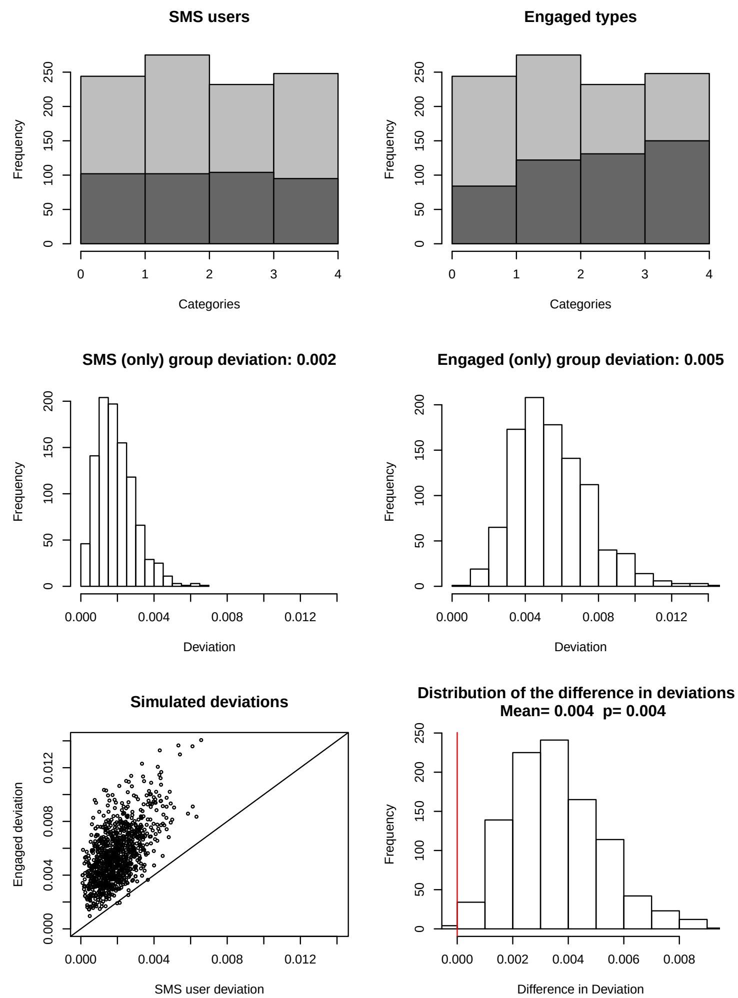
T_{SF} equals zero for those assigned to subsided price and one for those with full price

T_{0F} equals zero for those assigned to free price and one for those with full price

We then estimate sample average treatment effects associated with each of these binary treatments. These estimates take account of the blocked randomization by using gender and constituency as strata. ATEs are calculated as mean difference, whereby coefficients and p-values are derived using randomization inference, while taking account of the structure of blocking in the randomization scheme. To generate a summary measure of effects we

[4]Letting x_1 indicate membership of the SMS group and x_2 membership of the engaged group, we estimate a multinomial logit model under the assumption that $Pr(y=k|\beta) = \frac{e^{\beta_{k0}+\beta_{k1}x_1+\beta_{k2}x_2}}{\sum_{k=1}^{m} e^{\beta_{k0}+\beta_{k1}x_1+\beta_{k2}x_2}}$ (where we set $\beta_{1j}=0$. We estimate $\hat{\beta}$ using maximum likelihood and use draws from the resulting distribution of $\hat{\beta}$ to estimate the distributions of $(\alpha|x_1=1, x_2=0)$ and $(\alpha|x_1=0, x_2=1)$ from which we estimate a distribution of NR statistics.

Figure 5: Flattening Representation: Test of Hypothesis 2



Note: The upper row shows the distribution of preferences over categories for SMS users (left) and Engaged users (right) relative to the population as a whole. The second row shows estimates of the NR index for each group. The last row compares estimates of NR indices for each group. See code in Appendix 7.7

report, in addition, the linear trend in which the dependent variables are regressed on a three-category treatment variable. Linear trend (coefficients and p-values) are similarly derived using randomization inference linear model in which blocks are introduced as fixed effects.

Table 7 illustrates findings on H_3 and H_4.

Table 7: **Main Results (H_3, H_4)**

Treatment	Effect	(1) Any	(2) Public	(3) Private	(4) H_4 test (Col 4 - Col 3)
T_{OS}	ATE	-0.113	-0.086	-0.026	
	(p)	0.007	0.013	0.454	
	(N)	666	666	666	
T_{SF}	ATE	-0.073	-0.095	0.023	
	(p)	0.071	0.001	0.51	
	(N)	666	666	666	
T_{OF}	ATE	-0.186	-0.182	-0.004	
	(p)	0	0	0.903	
	(N)	666	666	666	
Trend	Trend	-0.093	-0.091	-0.002	-0.089
	(p)	0	0	0.93	0
	(N)	999	999	999	(999, 999)
		H_3 test			

Note: ATEs (estimates and p-values) are derived from blocked randomization inference. Linear trends (coefficients and p-values) are derived from randomization inference linear model in which blocks are introduced as fixed effects. Numbers are based on mock data simulated for the purpose of the mock article. See code in Appendix 7.8

4.4 Heterogeneous Effects (H_5)

To assess heterogeneous effects we estimate the price effects and the *differences in price effects* for wealthier and poorer constituents and for those with greater access (as defined above). We expect that higher prices will reduce uptake by poorer constituents more than richer constituents—resulting in more disproportionate messaging by the wealthy; and that a price increase will result in reduced use of the system by individuals with greater alternative channels of access. Since wealthier constituents also have more access we seek to examine the effects of wealth on the effect of price *conditional* on access and vice versa. The key results are shown in Table 8.

5 Price induced changes in message type choice

While our primary interest with respect to heterogeneous effects is whether poorer individuals and individuals with better alternative means of communication are more sensitive to price increases, our discussion of strategic interactions also suggested that wealthier indi-

Table 8: Heterogeneous Effects (H_5)

		Poor	Rich		All (Rich v Poor)		Difference
Low Access	Private:	-0.035	-0.079		-0.04		-0.004
		(0.1)	(0.1)		(0)		(0.3)
	Public	-0.107	-0.082		-0.099		0.03
		(0)	(0.1)		(0)		(0.7)
	Any	-0.142	-0.161		-0.139		0.026
		(0)	(0)		(0)		(0.6)
High Access	Private	-0.035	0.052		0.026		0.11
		(0.4)	(1)		(0.9)		(1)
	Public	-0.005	-0.1		-0.093		-0.065
		(0.4)	(0)		(0)		(0.3)
	Any	-0.04	-0.049		-0.068		0.045
		(0.2)	(0.2)		(0)		(0.6)
All Access	Private	-0.037	0.031		-0.002		0.048
		(0)	(1)		(0.2)		(0.9)
	Public	-0.089	-0.086		-0.091		0.001
		(0)	(0)		(0)		(0.5)
	Any	-0.126	-0.054		-0.093	H_{5a}	0.048
		(0)	(0)		(0)	(+)	(0.9)
Difference	Private	-0.03	0.092		0.037		
		(0.4)	(0.9)		(0.9)		
	Public	0.052	-0.039		0.004		
		(0.9)	(0.2)		(0.6)		
	Any	0.022	0.053	H_{5b}	0.041		
		(0.7)	(0.8)	(-)	(0.9)		

Note: Each cell shows the estimated marginal effect of an increase in price from a linear OLS model (or differences in marginal effect). p values estiated using randomization inference and talking account of the blocking strategy are provided in parentheses. Number of simulations: 5,000. See code in Appendix 7.9

viduals that participate when costs are higher shift the content of their messaging towards private goods.

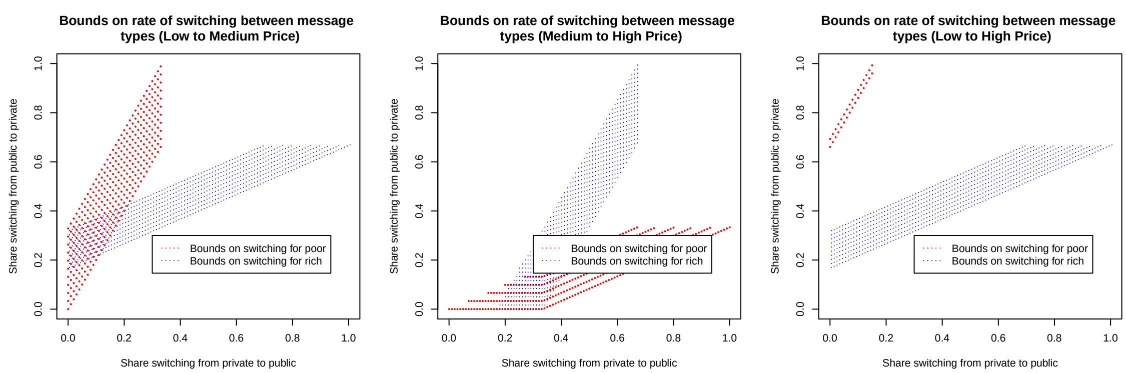
We can look for evidence of these patterns in Table 8 by assessing whether the marginal effect of price for wealthier constituents is more negative for public than for private messaging. However, even if such patterns hold, they can arise from two distinct processes. It may be that individual wealthier voters are more likely to send private messages when costs are higher, or it may be that those constituents that otherwise would have sent public messages have stopped sending messages due to the price increase.

More generally we seek to understand whether changes in the distribution of message types in different price environments results from changes in the population sending messages or changes in the kinds of messages senders sent (or both).

Distinguishing these processes is difficult because we do not know what given individuals would do under different price conditions. Treated subjects in our experiment were only offered the service at a single price level. We therefore use two approaches to assess the extent to which whether price increases result a change in the *population* of senders or changes in the *content* of messaging from given senders, or both.

First based on the distribution of outcomes at each price range we generate bounds on the shares switching categories *under the assumption that constituents that send at a high price would also send at a low price.* For example if 40% of subjects send public messages and 40% sent private messages when prices are low and 20% sent public messages and 30% sent private messages when prices are high we know that at least 50% of public message senders stopped sending public messages but no more than 75% of them could have switched to sending private messages (and only then if all the private message senders stopped sending public messages.) Similarly at least 25% of private message holders stopped sending private messages and no more than 50% of them switched to public messages. For any distribution of outcomes we can calculate a full set of bounds in this manner.

Figure 6: Bounds on estimated shares switching between public and private messaging



Note: Bounds on the shares switching from public to private and private to public broken down by wealth category. See code in Appendix 7.10

These bounds are shown in Figure 6 for both wealthy and less wealthy senders. Our discussion suggested that we expect that shifts from private to public are more likely than shifts from public to private; moreover we expect these shifts to be larger for the wealthy group (for the less wealthy group we expect a larger drop in the share of messages sent which will result in wider bounds in this analysis). [Note that this analysis on bounds for each group will be undertaken only of the quantity of messages sent is lower when prices are higher]

In a second approach we employ a model to assess more precisely the share of messages sent that address private issues for a pool of senders in the low cost condition that are matched to senders in the high cost condition. Specifically, for each high cost sender we seek a match from among the senders in the same constituency[5] that is as close as possible with respect to gender, wealth, and access. We then compare the types of messaging in the high cost pool and the matched pool. The results of this analysis are provided in figure 7. [Note again that this analysis on bounds assumes that high price senders can be matched with some senders in the low price group; this analysis is not supported if sending rates increase with the cost of sending]

6 Conclusion

Conclusions depend on findings.

7 Appendix: Code for Tables and Figures

7.1 Code for Tables 1 and 2 (Stata)

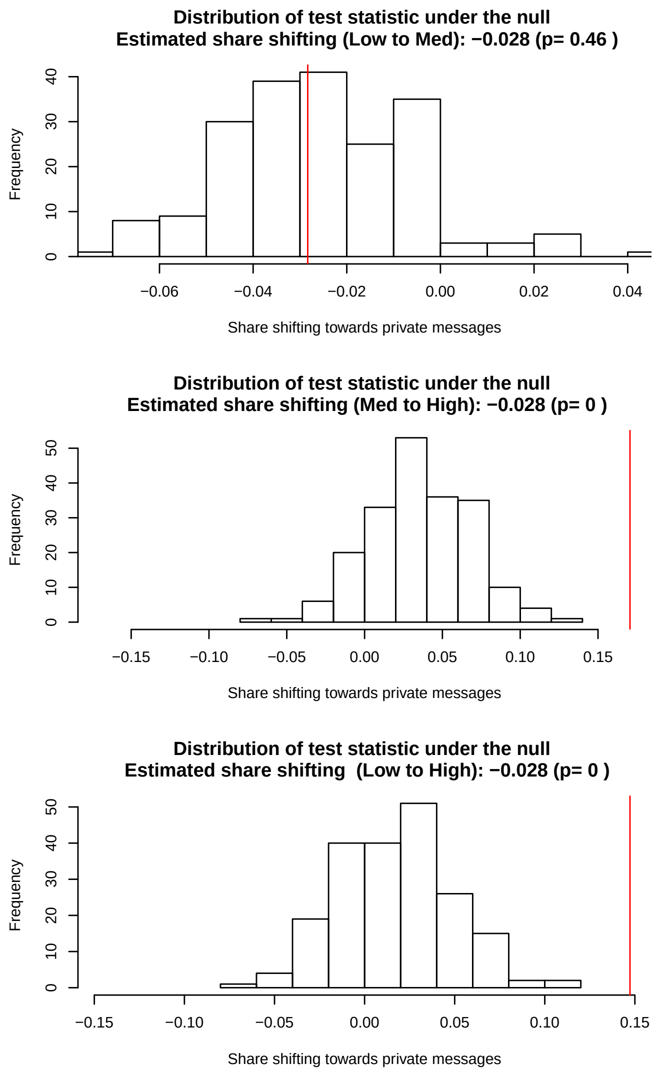
```
use data, clear

*** Table 1: ACCESS *****
eststo clear
foreach var in q44 q46 q47 computer travel geo_access accessA {
display "****************************************************
foreach xvar in  wealth_sub wealth_objA gender COG_SMP COE_SMP{
display "****************************************************
eststo a‘var’: reg ‘var’ i.‘xvar’, cl(V1_NAME)
margins ‘xvar’, post
parmest, label format(estimate \%8.2f) list(parm estimate min95 max95) level(95)
}
}

*** Table 2: ENGAGED *****
foreach var in engaged_party engaged_lc1 community_meeting raise_issue \\
demonstration rally letter q56_new q66bb q119_SMP engagedA{
display "****************************************************
foreach xvar in  wealth_sub wealth_objA gender COG_SMP COE_SMP geo_access{
display "****************************************************
eststo a‘var’: reg ‘var’ i.‘xvar’, cl(V1_NAME)
margins ‘xvar’, post
parmest, label format(estimate \%8.2f) list(parm estimate min95 max95) level(95)
```

[5] If data constraints prevent us from finding matches for more than 10% of our pool we will seek matches at the district level rather than constituency level

Figure 7: Estimates of shares switching from public to private messaging



Note: Estimates of the differences in types of messages sent between high cost senders and matched low cost constituents. Matching is based on mahalanobos distance within constituencies and over gender, wealth, and access. p values constructed using randomization inference. See code in appendix 7.11

```
}
}

cor ACCESSA ENGAGEDA
```

7.2 Code for Table 4 (Hypotheses 1) (Stata)

```
use data
foreach Xvar of varlist _poorS _poorO _women _nonCOE _nonCOG _distant {
display "**************************************************"
display "**************************************************"
quietly reg `Xvar' engagedA  if V6_PRICE!=. [weight = DUMMYWEIGHT]
quietly est store model1
quietly reg `Xvar' DUMMY_SMS  if V6_PRICE!=. [weight = DUMMYWEIGHT]
quietly est store model2
quietly suest model1 model2, cl(V1_NAME)
display "Share of engaged that are vulnerable:" ([model1_mean]_cons + [model1_mean]engagedA)
display "Share of senders that are vulnerable:" ([model2_mean]_cons + [model2_mean]DUMMY_SMS)
display "The difference in share of senders that are vulnerable:" ([model2_mean]_cons + \\\
 [model2_mean]DUMMY_SMS - [model1_mean]_cons - [model1_mean]engagedA)
test ([model1_mean]_cons + [model1_mean]engagedA =[model2_mean]_cons + [model2_mean]DUMMY_SMS)
}

foreach Xvar of varlist _poorS _poorO _women _nonCOE _nonCOG {
sureg (eq1: engagedA `Xvar' if V6_PRICE!=.) (eq2: DUMMY_SMS `Xvar' if V6_PRICE!=.) [weight = DUMMYWEIGHT]
display "**************************************************"
display "**************************************************"
display "The difference in the share of vulnerable groups using the method" ([eq2]`Xvar' - [eq1]`Xvar')
test [eq1=eq2]
}
```

7.3 Housekeeping Code for Analysis in (R)

```
rm(list=ls())
set.seed(1)
library(foreign)
library(optmatch)
library(xtable)
require(nnet)
library(MASS)


n  = 999
ID      = 1:n # ID
y  = floor(4*(runif(n))) # Some outcome, Y
BLOCK    = floor((1:n)/20) # Block
PRICE    = rep(c(0,1,2), n/3) # Price (Treatment variable)
ENG  = (y + 9*runif(n))>6 # Covariate: Engaged, correlated with Y
ACCESS   = 1*(runif(n)<.5) # Covariate: Access
ENGAGED  = 1*(runif(n)<.5) # Covariate: Engaged
WEALTHY  = 1*((runif(n)+.5*ACCESS)>.75) # Covariate: Wealth
GENDER   = runif(n)>.5   # Covariate: Gender
Z        = runif(n) # Latent variable for messaging outcome
```

```
SMS_V    = Z>(.8)    # Private message
SMS_B  = Z<(.3-.1*PRICE)   # Public messages (corr neg with price)
SMS   = SMS_V |SMS_B # Any message
DATA2  = data.frame(SMS=SMS,ENG=ENG,y=y)# Data frame for H2
```

7.4 Balance Table (Figure 4)

```
DB = read.dta("DataBalance.dta", convert.factors=FALSE)
DB = DB[!is.na(DB$V6_PRICE),]
T1 = DB$V6_PRICE
T1.names = c("Free", "Subsidized", "Full Price")

pdf(file="../L/figures/Balance.pdf", width=13, height=8)
vars = c("wealth_sub", "wealth_objA","q10b", "gender", "q17", "ACCESSA", "ENGAGEDA")
varnames = c("Wealth (Sub)", "Wealth (Obj)", "Age", "Gender", "Education",  "Access", "Engaged")
par(mfrow=c(4,length(vars)))
for(j in 0:3) for(v in 1:length(vars)) {
if(j>0) {
V= DB[,vars[v]]
nobs = sum(!is.na(V[T1==j]))
hist(V[T1==j], xlab=paste(T1.names[j], ", n=", nobs), main = "")}
if(j==0) {
V= DB[,vars[v]]
V = (V-mean(V, na.rm=TRUE))/sd(V, na.rm=TRUE)
plot(c(mean(V[T1==1], na.rm=TRUE), mean(V[T1==2], na.rm=TRUE), mean(V[T1==3], na.rm=TRUE)),
  rep(1,3), xlim = c(-.05,.05), ylim =c(.5, 2.5), axes=F, pch=19, cex=2, ylab="",
col=heat.colors(3), xlab="Standardized Mean Deviations", main = varnames[v])
axis(1)
text(c(mean(V[T1==1], na.rm=TRUE), mean(V[T1==2], na.rm=TRUE), mean(V[T1==3], na.rm=TRUE)),
  rep(1.2,3), T1.names, srt=90, adj=0)
}
}
dev.off()
```

7.5 Table 5 Patterns of Representativeness in Messaging

```
A = table(DATA2$y)/sum(table(DATA2$y))
E = table(DATA2$y[DATA2$ENG==1])/sum(table(DATA2$y[DATA2$ENG==1]))
S = table(DATA2$y[DATA2$SMS==1])/sum(table(DATA2$y[DATA2$SMS==1]))
ISSUET = round(cbind(A, E, S),3)
rownames(ISSUET) <- paste("Issue", LETTERS[1:length(A)])
colnames(ISSUET) <- c("All subjects", "Engaged subjects", "SMS senders")
xtable(ISSUET)
```

7.6 Main Outcomes Table (Table 6)

```
MainT=cbind(sapply(0:2, function(i) mean(SMS_V[PRICE==i])),
sapply(0:2, function(i) mean(SMS_B[PRICE==i])),
sapply(0:2, function(i) mean(SMS[PRICE==i])))
rownames(MainT) <- c("Free", "Subsidized", "Full")
colnames(MainT) <- c("Public SMS", "Private SMS", "Any SMS")
xtable(MainT)
```

7.7 H2: Flattening Representation (Figure 5)

```
# Function to construct test
Dev.Dist.Out  = function(D, sims, rond=3,
xyxlab="SMS user deviation", xyylab="Engaged deviation",  breaks=0:4){
  s = as.vector(table(D$SMS,D$ENG)) # (00), (SMS=1, ENG=0), (SMS=0, ENG=1), (11)
  s = s/(sum(s))    # Shares in each group

  # Function for prediction from Multinomial logit
  mlpred=function(SMS, ENG, B){(exp(B)%*%c(1,SMS,ENG))/sum(exp(B)%*%c(1,SMS,ENG))}
  mlpred_b=function(s,B){s[1]*mlpred(0,0,B)+s[2]*mlpred(1,0, B)+s[3]*mlpred(0,1,B)+s[4]*mlpred(1,1,B)}
  # Function to calculate deviation score between  subgroup distribution and pop distribution
  DEVIATION = function(SMS,ENG, B){ sum((mlpred(SMS,ENG, B) - mlpred_b(s,B))^2)/2}
  # Run DEVIATION for matrix of simulated betas
  Dev.Dist = function(SMS, ENG, BS,  sims){ sapply(1:sims, function(i) {
mu = t(matrix(BS[i,],3)); DEVIATION(SMS,ENG, mu)})}

  # Multinomial Model
  M <- multinom(y ~ SMS + ENG, data = D)
  # Simulated betas
  BS = mvrnorm(sims, mu = as.vector(t(coef(M))), Sigma = vcov(M))
  BS = cbind(matrix(0, sims, 3), BS)
  # Quantity of interest
  SMSDist = Dev.Dist(1,0, BS, sims=sims)
  ENGDist = Dev.Dist(0,1, BS, sims=sims)
  DiffDist = ENGDist - SMSDist  # Extent to which ENG (engaged) dist is large
  par(mfrow=c(3,2))
  hist(D$y, right=FALSE, breaks=breaks, col="grey", xlab="Categories", main="SMS users");
  hist(D$y[D$SMS==1],  right=FALSE,  breaks=breaks, add=TRUE, col=grey(.4))
  hist(D$y, right=FALSE, breaks=breaks, col="grey", xlab="Categories", main="Engaged types");
  hist(D$y[D$ENG==1], right=FALSE, breaks=breaks, add=TRUE, col=grey(.4))
  hist(SMSDist,   main = paste("SMS (only) group deviation:",
round(mean(SMSDist),rond)), xlab="Deviation", xlim=c(0, max(c(SMSDist, ENGDist))))
  hist(ENGDist,   main = paste("Engaged (only) group deviation:",
round(mean(ENGDist),rond)), xlab="Deviation", xlim=c(0, max(c(SMSDist, ENGDist))))
  plot(SMSDist, ENGDist, ylab = xyylab, xlab = xyxlab, main="Simulated deviations", cex=.5,
       xlim=c(0, max(c(SMSDist, ENGDist))), ylim=c(0, max(c(SMSDist, ENGDist)))); abline(a=0, b=1)
```

```
  hist(DiffDist, main = paste("Distribution of the difference in deviations  \n Mean=",
round(mean(DiffDist),rond),   " p=",  round(mean(DiffDist<0),rond)),
xlab="Difference in Deviation", xlim=c(min(DiffDist), max(DiffDist)))
  abline(v=0, col="red")
}

# note: for simulated data SMS group is close to pop; ENG (Engaged) group is not
with(DATA2, table(y,SMS))
with(DATA2, table(y,ENG))
# Graph
Dev.Dist.Out(DATA2, 5000)
pdf(file="../L/figures/NonRep.pdf", width=6.5, height=9)
Dev.Dist.Out(DATA2, 1000)
dev.off()
```

7.8 Table 7 H3 + H4 Main results: Effects of price on use and on relative propensity to send public

```
# FUNCTION FOR PERMUTATIONS
P = function(k, X, BLOCK=rep(1, length(X)), seed=1){
  set.seed(seed)
  p = function(z) {unlist(sapply(unique(BLOCK), function(i) sample(X[BLOCK==i], length(X[BLOCK==i]))))}
  sapply(1:k, p)}

# FUNCTION FOR RI inference taking account of block structure
# (code uses fragments developed together with J Fearon)
ri = function(Y, X, k=1000, BA=rep(1, length(Y)), BP=BA, C=TRUE, seed = 1,
id=TRUE, PERM = P(k, X, BP, seed)) {
set.seed(seed)
PERM = PERM[C,];  Y = Y[C]; X = X[C]; BA = BA[C]
BANAMES = sort(unique(BA))
n = length(Y) # Number of observations
f = function(x) {t = tapply(Y,x,mean,na.rm=TRUE); t[2]-t[1]}  # Function estimates ATE | No strata
t = tapply(Y,X,mean,na.rm=T); t0 = t[1]; t1 = t[2]   # Y(D=1), Y(D=0)
w = rep(1/n, n)
if(length(unique(BA))>1){ # Replace function if multiple strata
w = (sapply(BANAMES, function(i) sum(BA==i)))/n # Stratum weights
f <- function(x){t = tapply(Y,list(x, BA), mean,na.rm=TRUE); (w%*%(t[2,]-t[1,]))[1,1]}
t = tapply(Y,list(X, BA), mean,na.rm=TRUE); t1=(w%*%t[2,])[1,1]; t0 =(w%*%t[1,])[1,1]}
    est = f(X)                # ATE
    null = apply(PERM,2,f)     # ATEs for permutated data
    pL = mean(null<=est, na.rm=TRUE)               # p: Ha diff<0
    p  = mean(null>=abs(est), na.rm=TRUE) + mean(null<=-abs(est), na.rm=TRUE)        # p: Ha diff!=0
    pR = mean(null>=est, na.rm=TRUE)     # p: Ha diff>0
neyman.se = sum(sapply(BANAMES, function(i)  {
(sum(BA==i, na.rm=TRUE)/n)^2*(  (var(Y[BA==i & X==1], na.rm=TRUE )/length(Y[BA==i & X==1]))
```

```
+(var(Y[BA==i & X==0], na.rm=TRUE )/length(Y[BA==i & X==0] ))  ) } ))^.5
out = c(t0, t1,est, round(n,0), pL, p, pR, neyman.se);
names(out) = c('T = 0', 'T = 1', 'ATE', 'n', 'p (neg, ri)', 'p (two sided, ri)', 'p (pos, ri)', 'Neyman se')
cbind(out)}

# Linear function with ri for p values (two sided); Blocks introduced as fixed effects
ri_lm = function(y, x,  BA=rep(1, length(y)), BP=BA,  k=1000, C=TRUE, control=rep(0, length(y)),
PM =  P(k, x, BP), round=3){
  f = function(x){ coef(lm(y[C]~x[C]+as.factor(BLOCK[C])))[2]}
  null=apply(PM,2,f)
  p  = mean(null>=abs(f(x)), na.rm=TRUE) + mean(null<=-abs(f(x)), na.rm=TRUE)
  c(f(x),p,length(y[C]))}

k = 5000
# Two sided tests
Y=cbind(SMS, SMS_B, SMS_V)
OUT = matrix(NA, 12,3)
for(y in 1:3){
OUT[1:3,y]     <-ri(Y[,y], (PRICE==1), k=k, BA=BLOCK, C = (PRICE!=2))[c(3,6,4)]  # T_OS
OUT[4:6,y]     <-ri(Y[,y], (PRICE==2), k=k, BA=BLOCK, C = (PRICE!=0))[c(3,6,4)]  # T_SF
OUT[7:9,y]     <-ri(Y[,y], (PRICE==2), k=k, BA=BLOCK, C = (PRICE!=1))[c(3,6,4)]  # T_OF
OUT[10:12,y]   <-ri_lm(Y[,y],  PRICE,     k=k, BA=BLOCK)   # Linear
}
# rownames(OUT)<-c("T_OS", "p","n", "T_SF", "p","n", "T_OF", "p","n", "Linear", "p","n")

### Test that marginal effect on share public (y1) is smaller than marginal effect on share private (y2)
ri_lm2 = function(y1, y2, x,  BA=rep(1, length(y)), BP=BA,  k=1000,
C=TRUE, control=rep(0, length(y)),PM =  P(k, x, BP), round=3){
  f = function(x){ coef(lm(y1[C]~x[C]+as.factor(BLOCK[C])))[2] - coef(lm(y2[C]~x[C]+as.factor(BLOCK[C])))[2]}
  null=apply(PM,2,f)
  p  = mean(null>=abs(f(x)), na.rm=TRUE) + mean(null<=-abs(f(x)), na.rm=TRUE)
  c(round(f(x),digits=3),round(p, digits=3),paste("(",length(y1[C]),", ", length(y2[C]),")",sep=""))}
TEST = ri_lm2(SMS_B, SMS_V,  PRICE,   k=k, BA=BLOCK)

OUT= rbind(cbind(round(OUT, digits=3), c(rep("",9),TEST)), c("$H_3$","","",""))
C1 = c("$T_{OS}$", "", "","$T_{SF}$","", "","$T_{OF}$","", "", "Trend", "","", "")
C2 = c(rep(c("ATE", "($p$)", "($N$)"),3), c("Trend", "($p$)", "($N$)"), "")
OUT = cbind(C1,C2, OUT)
colnames(OUT)<-c("Treatment","Effect", "Any", "Public","Private", "H4 test")
print(xtable(OUT), include.rownames=FALSE)
```

7.9 H5: Table 8 of Heterogeneous Effects

```
# PERMUTATIONS
P = function(k, X, BLOCK=rep(1, length(X)), seed=1){
set.seed(seed)
p = function(z) {unlist(sapply(unique(BLOCK), function(i) sample(X[BLOCK==i], length(X[BLOCK==i]))))}
```

```
sapply(1:k, p)}

# Calculate results by cell; function estimates b or interaction term on Z for data
# conditional on C and accounting for blocks using fixed effects
results2 = function(y, x,  BLOCK,  sims=1000, C=TRUE, ANBLOCK=TRUE, Z=NA,
control1=rep(0, length(y)), control2=control1, interaction=FALSE,
PM =  P(sims, x, BLOCK), round=3){
if(!interaction) f = function(x){
coef( lm(y[C]~x[C]+control1[C]+control2[C]+as.factor(BLOCK[C])))[2]}
if(interaction)  f = function(x){
INT= x*Z; coef(lm(y[C]~x[C]+Z[C]+INT[C]+control1[C]+control2[C]+as.factor(BLOCK[C])))[4]}
bs= apply(PM,2,f)
p = mean(bs<f(x))
c(round(f(x),round),paste("(",p,")", sep=""))}

# Row and Column Conditions for Table
rC = cbind((ACCESS==0), (ACCESS==1), TRUE)
cC = cbind((WEALTHY==0), (WEALTHY==1), TRUE)
DV = cbind(SMS_V, SMS_B, SMS)

# Output
PM =  P(10, PRICE, BLOCK)
zeros=rep(0, length(SMS))
X = matrix(NA, 24,6)
for(d in 1:3)for(r in 1:3)for(c in 1:3){X[((r-1)*6+(d-1)*2+1):((r-1)*6+(d-1)*2+2), c]<- {
results2(DV[,d], PRICE, BLOCK,  C=(rC[,r]&cC[,c]), PM=PM)}}
# INTERACTIONS ON ACCESS (FILL LOWER ROWS)
for(d in 1:3)for(c in 1:2){X[(19+(d-1)*2):(20+(d-1)*2), c]<- {
results2(DV[,d], PRICE, BLOCK,Z=ACCESS, C=(cC[,c]), interact=TRUE, PM=PM)
}}
for(d in 1:3)for(c in 3){X[(19+(d-1)*2):(20+(d-1)*2), c]<- {
results2(DV[,d], PRICE, BLOCK,Z=ACCESS, C=(cC[,c]), interact=TRUE, PM=PM,
control1=WEALTHY, control2= (WEALTHY*PRICE))
}}
# INTERACTIONS ON WEALTH (FILL RIGHT COLUMNS ROWS)
for(d in 1:3)for(r in 1:2){X[((r-1)*6+(d-1)*2+1):((r-1)*6+(d-1)*2+2), 4]<- {
results2(DV[,d], PRICE, BLOCK,Z=WEALTHY,  C=(rC[,r]), interact=TRUE, PM=PM)
}}
for(d in 1:3)for(r in 3){X[((r-1)*6+(d-1)*2+1):((r-1)*6+(d-1)*2+2), 4]<- {
results2(DV[,d], PRICE, BLOCK,Z=WEALTHY,  C=(rC[,r]), interact=TRUE, PM=PM,
control1=ACCESS, control2=(ACCESS*PRICE))
}}
# Export Table
colnames(X)<-c("Poor", "Rich", "All", "Difference", "", "")
X=cbind(c("Low Access", rep("", 5), "High Access", rep("", 5),
"Any Access", rep("", 5),"Difference", rep("", 5)),
  c(rep(c("Private","","Public","", "Any", ""),4)),
  X[,1:2],
```

```
  c(rep("",22), "H5b", "(-)"),
  X[,3],
  c(rep("",16), "H5a", "(+)", rep("", 6)),
X[,4])
colnames(X)<-c("", "", "Poor", "Rich", "", "All", "", "Difference")
xtable(X)
```

7.10 H5: Graph Bounds (Figure 6)

```
# say probability that a private type shifts to public is pVB when
# price goes up; then we seek to estimate a range for pBV, the
# probability that a public type shifts to private
# Data Structure: Summary data in matrix M; Row 1 is no message,
# row 2 is private, row 3 is public

# minimum share given moving from row A to row B assuming no one who sent
# no message under low price sends under high price
min_pAB = function(M, a, b) max(c(0,M[b,2]-M[b,1]))/M[a,1]

# maximum share given moving from row A to row B assuming no one who sent no
# message under low price sends under high price
max_pAB = function(M, a, b) min(M[a,1], M[b,2])/M[a,1]

# Range of shifts in one direction given particular shifts in other
r_pBV_given_pVB = function(pVB, M) c(max(0, min(M[2,2]-(1-pVB)*M[2,1],M[3,1])),
  min(M[3,1]-(M[3,2]-pVB*M[2,1]), M[3,1], M[2,2]))/M[3,1]

# Different matrices possible for different groups (eg wealthy, poor etc)
M0 = matrix(c(1/4,1/2,1/4, .26,.665,.075),3)  # Example with tight bounds
M1 = matrix(c(1/4,1/2,1/4, 1/3,1/2,1/6),3)    # Example with drop in pubic and no change in private
M2 = matrix(c(1/4,1/4,1/2, 1/3,1/3,1/3),3)    # Drop in public, smaller rise in private
M3= M1[c(1,3,2),]; M4= M2[c(1,3,2),]
rownames(M2) <- rownames(M1)<- rownames(M0) <- c("None", "Private", "Public")
colnames(M2) <- colnames(M1) <- colnames(M0)<- c("Low", "High")

# Function to graph bounds using bands
# Thickness of bands show uncertainty resulting from group that stops sneding messages
bounds.graph = function(M1, M2, lty=3, gap=.01,
main="Bounds on rate of switching between message \n types (Low to Medium Price)"){
plot(c(0,1), c(0,1), type="n",
xlab = "Share switching from private to public",
ylab = "Share switching from public to private",
main = main)
s1 = seq(min_pAB(M1, 2, 3) , max_pAB(M1, 2, 3),gap)
s2 = seq(min_pAB(M2, 2, 3) , max_pAB(M2, 2, 3),gap) + .005
segments(s1, sapply(s1, function(i) r_pBV_given_pVB(i, M1)[1]),
 s1, sapply(s1, function(i) r_pBV_given_pVB(i, M1)[2]), col="red", lty=lty, lwd=2)
```

```
segments(s2, sapply(s2, function(i) r_pBV_given_pVB(i, M2)[1]),
 s2, sapply(s2, function(i) r_pBV_given_pVB(i, M2)[2]), col="blue", lty=lty )
legend(.3, y=.3, c("Bounds on switching for poor", "Bounds on switching for rich"),
 col = c("red", "blue"), lty=3, bg="white")
 }

# Graph Output
pdf(file="../L/figures/Bounds.pdf", width=12, height=4)
par(mfrow=c(1,3))
bounds.graph(M1,M2)
bounds.graph(M3,M4, main = "Bounds on rate of switching between message \n types (Medium to High Price)")
bounds.graph(M0,M2, main = "Bounds on rate of switching between message \n types (Low to High Price)")
dev.off()
```

7.11 Matching using Mahalanobis distance (Figure 7)

```
# Function for estimating change in message type as a function of price by matching on
# high price senders
# DATA contains, in order:  ID DEPVAR XVAR PREDICTORS
# NOTE WEALTHY IMPOSE ASSUMPTION THAT MATCHES FOR HIGH PRICE SENDERS ARE IN THE LOW PROCE SENDER GROUP
# IT IS POSSIBLE THOUGH THAT SOME OF THESE WOULD NOT HAVE SENT UNDER LOW CONDITION
# Want to be sure there is at least one treated unit in each section (see if command on dim() in block)

distdiff.block = function(DATA, sims=10, seed=1, Perm = TRUE, SUB=rep(TRUE, nrow(DATA))){
set.seed(seed)
PREDICTORS=DATA[,4:ncol(DATA)]
DATA$ID = 1:nrow(DATA)
DATA=DATA[SUB,]
DATA$PERM_X = DATA$XVAR
OUT = rep(NA, sims)
for(k in 1:sims){
if(Perm) DATA$PERM_X =sample(DATA$PERM_X)  #SCRAMBLE FOR RI
M = mahal.dist( as.formula(paste("PERM_X~", paste(colnames(PREDICTORS), collapse="+"))),
DATA,  structure.fmla=XVAR~BLOCK)
tIDs <- cIDs<-  MATCHOUTCOME<-MIN<- list()
for(j in 1:length(M)){
if(dim(M[[j]])[1]!=0){
tIDs[[j]] <- as.numeric(rownames(M[[j]]))
cIDs[[j]] = as.numeric(colnames(M[[j]]))
MIN[[j]] = apply(M[[j]],1,min) # Minimum distance for each row
 MATCHOUTCOME[[j]] = rep(NA, length(tIDs[[j]]))
# Generate average outcome for MIN matches; allowing possibly multiple per match
for(i in 1:length(tIDs[[j]])){
MATCHOUTCOME[[j]][i]<-mean(DATA$DEPVAR[(DATA$ID)%in%(cIDs[[j]][M[[j]][i,]==MIN[[j]][i]])])}
}}
MATCHOUTCOMES=unlist(MATCHOUTCOME)
OUT[k]=mean(DATA$DEPVAR[DATA$PERM_X==1]) - mean(MATCHOUTCOMES)
```

```
}
OUT
}

# POTENTIAL OUTCOMES FOR H5: Generate a data frame with more or
# less switching from public to private messaging
gen.DATA = function(switch=0){
  NEVER = (ACCESS+runif(n))<.5 # People who would never send a message
  PRIV_L=(ACCESS+.2*GENDER)>.5 # People who would do private at low price (Men more likely)
  DROPPER = .8*GENDER+runif(n)>.9 # People who would drop out if prices high (Men more likely)
  SWITCHER = GENDER+WEALTHY < switch # People who would switch to private messaging if prices high
  SWITCHER[PRIV_L]<-0 # Only public messagers switch
  PRIV_L[NEVER]<-FALSE
  PUB_L = (!NEVER) & (!PRIV_L)
  PUB_H=PUB_L
  PUB_H[SWITCHER | DROPPER]<-FALSE
  PRIV_H= PRIV_L
  PRIV_H[SWITCHER&PUB_L] = TRUE
  PRIV_H[DROPPER] = FALSE
  data.frame(PRIV_L, PUB_L, PRIV_H, PUB_H, SWITCHER, DROPPER, NEVER)
}

# EXAMPLE OF SIMULATED *REALISED* DATA (with switching to private)
D3     = gen.DATA(.7)
HIGH   = PRICE==2
SMS_V2  = D3$PRIV_H*HIGH + D3$PRIV_L*(1-HIGH)   #Realised private messages  (function of price)
SMS_B2  = D3$PUB_H*HIGH + D3$PUB_L*(1-HIGH)  #Realised public messages (function of price)
SUB = (SMS_V2 | SMS_B2)

# ESTIMATES AND RI Distribution
sim=200
R1=distdiff.block(data.frame(DEPVAR=SMS_V2,  XVAR=(PRICE==1), BLOCK=BLOCK,
GENDER, WEALTHY, ACCESS), sims=1, SUB=(SUB&(PRICE!=2)), Perm=FALSE)
Z1=distdiff.block(data.frame(DEPVAR=SMS_V2,  XVAR=(PRICE==1), BLOCK=BLOCK,
GENDER, WEALTHY, ACCESS), sims=sim, SUB=(SUB&(PRICE!=2)))
R2=distdiff.block(data.frame(DEPVAR=SMS_V2,  XVAR=(PRICE==2), BLOCK=BLOCK,
GENDER, WEALTHY, ACCESS), sims=1, SUB=(SUB&(PRICE!=0)),Perm=FALSE)
Z2=distdiff.block(data.frame(DEPVAR=SMS_V2,  XVAR=(PRICE==2), BLOCK=BLOCK,
GENDER, WEALTHY, ACCESS), sims=sim, SUB=(SUB&(PRICE!=0)))
R3=distdiff.block(data.frame(DEPVAR=SMS_V2,  XVAR=(PRICE==2), BLOCK=BLOCK,
GENDER, WEALTHY, ACCESS), sims=1, SUB=(SUB&(PRICE!=1)),Perm=FALSE)
Z3=distdiff.block(data.frame(DEPVAR=SMS_V2,  XVAR=(PRICE==2), BLOCK=BLOCK,
GENDER, WEALTHY, ACCESS), sims=sim, SUB=(SUB&(PRICE!=1)))

pdf(file="../L/figures/MatchEstBlocks3.pdf", width=5, height=8)
par(mfrow=c(3,1))
hist(Z1,  xlab = "Share shifting towards private messages", xlim=c(min(c(Z1, -R1)),max(c(Z1, R1))),
main=paste("Distribution of test statistic under the null \n Estimated share shifting (Low to Med):",
```

```
round(R1,3), "(p=", mean(Z1>R1),")"))
abline(v=R1, col="red")

hist(Z2,  xlab = "Share shifting towards private messages", xlim=c(min(c(Z2, -R2)),max(c(Z2, R2))),
main=paste("Distribution of test statistic under the null \n Estimated share shifting (Med to High):",
round(R1,3), "(p=", mean(Z2>R2),")"))
abline(v=R2, col="red")

hist(Z3,  xlab = "Share shifting towards private messages", xlim=c(min(c(Z3, -R3)),max(c(Z3, R3))),
main=paste("Distribution of test statistic under the null \n Estimated share shifting  (Low to High):",
round(R1,3), "(p=", mean(Z3>R3),")"))
abline(v=R3, col="red")
dev.off()
```

References

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103(484):1481–1495.

Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15(1):183–201.

Balkin, Jack M. 2004. "Digital speech and democratic culture: a theory of freedom of expression for the information society." *New York University Law Review* 79.

Banerjee, Abhijit and Rohini Pande. 2009. "Parochial Politics: Ethnic Preferences and Politician Corruption." *Typescript. Department of Economics, Massachusetts Institute of Technology* .

Barkan, Joel D. 2009. *Legislative power in emerging African democracies*. Boulder: Lynne Rienner Publishers.

Barkan, Joel D., Robert Mattes, Shaheen Mozaffar and Kimberly Smiddy. 2010. "The African Legislatures Project: First Findings." *CSSR Working Paper No. 277* .

Bartels, Larry M. 2008. *Unequal democracy: the political economy of the new gilded age*. New York: Russell Sage Foundation.

Bimber, Bruce. 2001. "Information and Political Engagement in America: The Search for Effects of Information Technology at the Individual Level." *Political Research Quarterly* 54(1):53–67.
URL: *http://prq.sagepub.com/content/54/1/53.abstract*

Hafkin, Nancy J. and Sophia Huyer. 2007. "Women and Gender in ICT Statistics and Indicators for Development." *Information Technologies and International Development* 4(2):25–41.

Heyman, James and Dan Ariely. 2004. "Effort for Payment: A Tale of Two markets." *Psychological Science* 15(11):787–793.

Hilbert, Martin. 2011. "Digital gender divide or technologically empowered women in developing countries? A typical case of lies, damned lies, and statistics." *Women's Studies International Forum* 34(6):479–489.

Hyde, Susan D. 2007. "The Observer Effect in International Politics: Evidence from a Natural Experiment." *World Politics* 60(1):37–63.

Nam, Taewoo. 2010. "Whither Digital Equality? An Empirical Study of the Democratic Divide." *System Sciences (HICSS), 2010 43rd Hawaii International Conference* pp. 1–10.

Oates, Briony J. 2003. "The potential contribution of ICTs to the political process." *Electronic Journal of e-Government* 1(1):31–39.

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

Park, Sora. 2009. "Concentration of internet usage and its relation to exposure to negative content: Does the gender gap differ among adults and adolescents?" *Women's Studies International Forum* 32(2):98–107.

Shampanier, Kristina, Nina Mazar and Dan Ariely. 2007. "Zero as a Special Price: The True Value of Free Products." *Marketing Science* 26(6):742–757.

Shirazi, Farid, Ojelanki Ngwenyama and Olga Morawczynski. 2010. "ICT expansion and the digital divide in democratic freedoms: An analysis of the impact of ICT expansion, education and ICT filtering on democracy." *Telematics and Informatics* 27(1):21–31.

Thompson, Mark. 2008. "ICT and development studies: Towards development 2.0." *Journal of International Development* 20(6):821–835.

Wantchekon, Leonard and Thomas Fujiwara. 2011. "Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin." *Working Paper* .