

Intended and unintended consequences of democracy promotion assistance to Georgia after the Rose Revolution

Research and Politics
April-June 2014: 1–13
© The Author(s) 2014
DOI: 10.1177/2053168014530279
rap.sagepub.com
 SAGE

Jesse Driscoll¹ and F Daniel Hidalgo²

Abstract

What are the political consequences of democratization assistance to regimes transitioning from authoritarian rule? By exploiting the downstream effects of a field experiment designed to encourage citizen monitoring of Georgia's 2008 parliamentary elections, we evaluate the political consequences of one type of democracy promotion aid. The intervention increased citizen activism, but it also had the unanticipated effect of suppressing overall voter turnout by approximately 5%. We hypothesize that the civic education campaign was interpreted as a sign of increased political attention to a selected voting precinct, which suppressed opposition turnout. Two additional experiments provide additional evidence for the hypothesis.

Keywords

Democracy promotion, semi-authoritarian regimes, experimental research

Introduction

The delegitimization of the public sphere implies that decent people should avoid politics, or enter them only grudgingly, on the grounds that "someone has to do the dirty work."

—Ghia Nodia, 2002

With the goal of unpacking the consequences of foreign-led empowerment of civil society, a pilot program was conducted through an international democracy promotion NGO during the 2008 Georgian parliamentary elections. The intent of the program was to increase civic engagement and deter voter fraud by providing Georgian citizens with information on (a) how to file a formal complaint and (b) how an individual could monitor and change the voter registry. Programs were implemented in randomly selected voting precincts, which allowed for the gold standard in program evaluation methodology: a randomized controlled experiment. The pilot program increased citizen activism, evidenced by a 12% rise in registered complaints and a moderate increase in voter registration.

While the intervention worked as intended, examination of the information campaign's effects on electoral outcomes revealed an unexpected result: precincts randomly selected for the "complaints" intervention had 5% lower voter turnout than those that received no treatment. For every person contacted, at least one person stayed home on

election day. In two separate replications of this result, precincts that were randomly selected for a 2008 or 2010 pre-election panel survey—even if they received no additional contact with the NGO—had a voter turnout about 3% lower than precincts where no data were gathered. In all three experiments, relatively innocuous interventions induced voter abstention. Given that these interventions are representative of standard methodologies used in democracy promotion activities and social scientific inquiry more generally, an in-depth exploration of these unexpected findings may shed light on important and under-reported political processes unfolding in the shadow of democracy promotion aid.

An individual's decision to turn out to vote on election day is typically modeled as a minimally costly action with few direct benefits.¹ Since rationalist approaches to turnout typically depend on non-material inducements (i.e. "expressive" benefits) to explain aggregate electoral participation, it

¹School of International Relations and Pacific Studies, University of California at San Diego, San Diego, CA, USA

²Department of Political Science, MIT, Cambridge, MA, USA

Corresponding author:

F Daniel Hidalgo, 77 Massachusetts Avenue, Room E53-470, Cambridge, MA 02139, USA.

Email: dhidalgo@mit.edu



is plausible that in the Georgian case a small increase in negative emotions—fear or disgust—on the part of study respondents might have tipped the scales against participation among some voters on election day. In semi-authoritarian contexts, some citizens believe that regimes can monitor individual voting decisions and use the results to tailor retaliations.² As a result, many post-Soviet elections resemble “rituals of consent” (“voting in order to be seen voting by anyone who might be interested in keeping track of who is voting”) rather than exercises in preference aggregation.³ These fears are not subject to easy falsification. They are akin to a conspiracy theory, which can vary in potency across different electoral districts within the same state or county. We hypothesize that whether a Georgian voter turned out to participate in the 2008 or 2010 election hinged on individuals’ self-assessment of two factors: (a) whether the regime was capable of monitoring his or her vote choice, and (b) potential personal costs of retaliation for voting for the opposition. Our suspicion is that our intervention triggered risk aversion behaviors: staying home on election day.

An alternative hypothesis to regime surveillance is disenchantment with politics in general, as suggested by the epigraph to this paper. Randomization of the content of a 2010 pre-election survey found that refusal to voluntarily supply information was more common the more “political” the survey instrument was. This finding was especially acute in precincts rife with mistrust of regime intentions. Perhaps citizens feared that the data being collected would be somehow used against them. Or perhaps respondents see politics as intrinsically “dirty,” and viewed the activities of the foreign-assisted youth who arrived from the capital with some mix of pity and disgust. Regardless of the precise mechanism, these findings draw attention to a possible source of bias in much of the survey data collected in semi-authoritarian regimes on politically sensitive subjects.

Local context: The Georgian case

Georgia is a post-Soviet state that enjoys an unusually strong strategic relationship with the United States. It receives more Western aid per capita than the rest of the states in the region combined.⁴ In 2003, peaceful protests swept Eduard Shevardnadze and his regime from power when exit polls did not match official vote tallies.⁵ In the wake of the “Rose Revolution,” the next president, Mikheil Saakashvili, allowed the Soros Foundation, NDI, IRI, Transparency International, and various other democracy promotion organizations freedom of action.

At the time of the study Saakashvili’s party, the *United National Movement* (UNM), dominated the political landscape by using formal institutions—notably the judiciary, the Ministry of Interior, and Parliament—to discourage electoral alternatives.⁶ As president, Saakashvili enjoyed favorable media coverage and free time on state television networks. In a pre-election survey conducted in the lead-up

to the 2008 parliamentary election, 32% of voters reported that people had to “vote a certain way to keep their jobs.” An additional 34% of voters admitted that this happens “sometimes.”⁷ For all the disappointments of the post-Rose Revolution era, however, Saakashvili’s decision to concede electoral defeat to Bidzina Ivanishvili’s *Georgian Dream* party in 2012 represents the clearest example of a party in post-Soviet Eurasia losing an election and voluntarily stepping down from power.⁸

Two facets of the 2008 Parliamentary election deserve special note. First, the election was conducted in an atmosphere of paranoia and conspiracy. Many voters believed the integrity of the secret ballot was compromised. It was taken for granted by Georgian media outlets that authorities could figure out how constituents voted, and tailor rewards or punishments accordingly.⁹ This is borne out in survey data. According to a nationally representative household survey conducted in the Spring of 2009, only 50.1% of voters believed that their vote was secret in last Georgian election—a far lower figure than was reported for the same question in either Armenia (72.8%) or Azerbaijan (62.8%).¹⁰ Second, little distinguished the various opposition parties in terms of ideology or policy proposals. Voters generally went to the polls to either affirm or reject the status quo—to either vote for or against Saakashvili’s UNM party.

Program evaluation: Expected and unexpected findings

In the weeks leading up to the 2008 parliamentary elections, the authors worked with an international democracy promotion NGO to randomize the precincts that would receive a voter information program. The goal of this pilot program was to educate Georgian citizens about common technologies of electoral fraud and—it was hoped—to deter voter fraud by disseminating accurate information about how to file an anonymous complaint. With the aid of the implementing organization, we tailored the content of the message to address types of fraud regularly reported to be associated with Georgia’s electoral environment. The program was implemented by the Caucasus Research Resource Centers (CRRRC). The total program budget was just under \$10,000. Eligible precincts were those previously selected as primary sampling units in a national panel survey on the 2008 parliamentary elections. Canvassers delivered the message several weeks after the first wave of the survey. To increase efficiency, precincts were grouped in blocks of four, based on region, urban or rural, 2006 governing party’s vote share, and percentage of ballots cast between 5:00 p.m. and 8:00 p.m. in the previous election, which election observers use as a rough proxy of voter fraud (since ballot stuffing tended to occur at the end of the day). Within these blocks of four, two precincts were randomly selected to receive the intervention.

Table 1. The average treatment effect of the complaints treatment on whether or not any complaint was made and the number of complaints in the 2008 elections. For covariate adjusted results, the adjustment variables are the number of registered voters, 2006 turnout, and 2006 vote share of the ruling party. Standard errors are heteroskedasticity consistent and account for blocking.

	Complaints dummy		Total complaints	
	1	2	3	4
Complaints treatment	0.12 ^b (0.04)	0.12 ^b (0.04)	0.26 ^a (0.10)	0.26 ^b (0.09)
Covariates				
N	84	84	84	84

Robust standard errors in parentheses.

^aSignificant at $p < .05$.

^bSignificant at $p < .01$.

Table 2. The average treatment effect of the complaints treatment on turnout and the vote share of the ruling party. For covariate adjusted results, the adjustment variables are number of registered voters, 2006 turnout, and 2006 vote share of the ruling party. Standard errors are heteroskedasticity consistent and account for blocking.

	Turnout (%)		Pres. Vote (%)	
	1	2	3	4
Complaints treatment	-5.44 ^a (2.57)	-5.27 ^a (2.10)	0.51 (2.81)	1.42 (2.03)
Covariates				
N	84	84	84	84

Robust standard errors in parentheses.

^aSignificance at $p < 0.05$.

Description of intervention

The pilot program studied an intervention designed to encourage voters to report any observed problems on election day to either local electoral officials or to well-known NGO partners (with experience in election monitoring).¹¹ The messages were delivered through a combination of pamphletting and canvassing. In every treatment precinct, approximately 30 households were visited by canvassers and given a face-to-face oral message and a flier. Canvassers also distributed fliers to roughly 80 additional households in the precinct (2–3 neighbors per home selected). Given that average precinct size in the experimental samples is 1151 voters and assuming an average of 2 voting age adults in each household,¹² we would have directly exposed about 20% of households to the message. The canvassers were college students recruited from Tbilisi universities. Most enumerators had substantial survey administration experience and training through the CRRC. The text of the messages are in the appendix.

Expected results: Enhanced civic activism against electoral malfeasance

The intent of the “Complaints” intervention was to encourage voters to monitor local elections themselves, filing a

grievance if election regulations were violated. To check if our intervention had this effect, we compiled data on complaints of electoral malfeasance reported by prominent election monitoring organizations. We use two dependent variables: (a) a dummy variable for at least one complaint filed by a voter in that precinct and (b) the raw number of complaints filed. Results are presented in Table 1. We find strong evidence that our intervention increased the propensity of voters to challenge voting procedures on election day. The mean difference across treatment and control for the complaints dummy implies a 12% increase in the probability of a complaint being filed in treated districts. For the models using the raw number of complaints, the point estimate is 0.26—which is roughly twice the mean number of complaints filed in all precincts in the sample.

Unexpected downstream results: Turnout suppression

Subsequent investigation of the data revealed another finding: a substantial number of Georgians did not cast their votes on election day in districts selected for the “Complaints” treatment, as seen in Table 2. The average effect of the complaints treatment (unadjusted specification) on voter turnout is about -5.4%. We find a positive,

Table 3. “Survey as an Experiment.”

	Turnout %		Pres. Party vote %	
	2008	2010	2008	2010
Estimate	−3.09	−2.51	−0.03	−2.16
SE	1.47	1.38	0.02	1.84
95% CI	(−5.97, −0.21)	(−5.22, 0.2)	(−0.07, 0.01)	(−5.76, 1.45)
N	2994	3101	2994	3101

This table shows the estimated effects of a precinct being included in two different political surveys. The first survey was conducted immediately before the 2008 parliamentary election. The second survey was carried out before the 2010 local elections. Treatment effect estimates account for unequal probability of selection and control for precinct size, the variable used to sample precincts into the “treatment”.

but small, difference in the ruling party’s vote share (model 3 and 4), though the null hypothesis of no difference could not be not rejected at conventional levels.¹³ This unexpected finding prompted additional investigation.

Turnout suppression effects of two pre-election surveys

Given that the turnout dependent variable was not pre-specified prior to the experiment’s implementation, skeptics could reasonably dismiss the turnout suppression finding as a statistical fluke. This section presents two partial replications of the result, wherein two *pre-election surveys* are treated as randomized interventions. Subcontracted Georgian canvassers capital city visited randomly selected precincts, asking standard questions about political beliefs a few weeks before two separate elections. The 2008 and 2010 CRRC surveys had a multi-stage sampling structure, with precincts as the primary sampling units. In 2008, prior to the parliamentary elections, 85 precincts were selected out of a total of 3019 precincts using sampling probabilities (π_i) proportional to the number of registered voters. Using a virtually identical sampling procedure in the weeks prior to the 2010 local elections, 110 precincts were selected as primary sampling units for a similar political survey. The unit of analysis for this “experiment” is the precinct and the treatment is selection into the sample. To recover the average treatment effect of being included in the survey sample on turnout and regime vote share, we fit a weighted linear model of each dependent variable on a treatment indicator variable and the selection variable, the number of registered voters. The weights are inverse probability weights, where the weight for each treatment unit is equal to $\frac{1}{\pi_i}$ and the weight for each control unit is $\frac{1}{1-\pi_i}$. Because the selection weights are known, incorporating these weights in our estimation procedure allow us to recover unbiased treatment effect estimates.¹⁴ Balance statistics for this design are presented in the Appendix.

The results are reported in Table 3. In both the 2008 parliamentary election and the 2010 local elections, being randomly selected for a pre-election survey had a statistically

significant voter suppression effect. The suppression effect was slightly more acute in 2008 than in 2010 (a lower turnout of approximately 3% compared to a lower turnout of approximately 2.5%). The results are consistent with the effects of the complaints intervention reported in Table 2.

Potential mechanisms

Why did the three interventions lower turnout? The interventions could have lowered turnout either by changing the behavior of Georgian voters or the functioning of the elections themselves. In the latter scenario, exposure to the information or survey intervention could have deterred would-be ballot stuffers, thus reducing the number of manufactured voters and consequently reducing measured turnout. Ballot-stuffing and tabulation manipulation are prevalent in the South Caucasus.¹⁵ In the 2008 election, however, European election monitors observed only one instance of ballot stuffing and the ballot counting procedures were generally assessed “positively” (OSCE, 2008: 7). Under this scenario, political operatives—typically allied with the ruling regime—would interpret the interventions as increasing the probability of being caught and punished for breaking the law, thus deterring electoral manipulation. According to this hypothesis, one would expect the vote share for the president’s party to decrease in treatment precincts, as ballot stuffers would no longer generate the same number of votes for the ruling party. The evidence for this implication of the fraud dampening mechanism is equivocal. The point estimates of the effect of the complaints intervention (Table 2) on the vote share of the presidential party is *positive* (albeit not statistically significant), while the effects of the survey (Table 3) are indeed negative (again not significant). Furthermore, it is not clear why a precinct being selected as a primary sampling unit for a survey would reduce ballot stuffing, since enumerators were deployed several weeks before the election and did not present themselves as election observers. Given the weak theoretical rationale for attributing reduced ballot stuffing to the “survey as an experiment” intervention and the null results on regime party vote share, the case for this mechanism is tenuous.

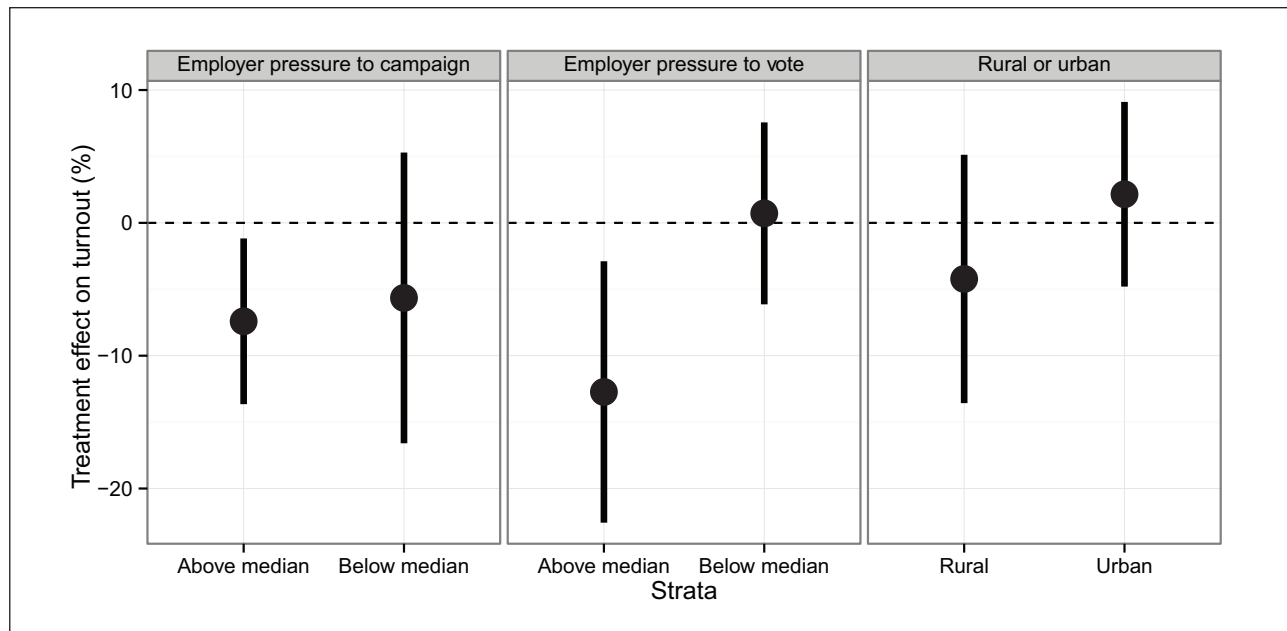


Figure 1. Heterogeneous treatment effects: Heterogeneity in the effect of the complaints treatment on voter turnout by three variables. ATE in each covariate strata is from a linear model with block fixed effects and adjusted for covariates. Standard errors are heteroskedasticity consistent.

Our favored explanation for the findings is that the interventions lowered the propensity of some Georgian citizens to vote. Being surveyed has been shown to affect consumer behavior (Zwane et al., 2011) in a variety of contexts and several cognitive mechanisms such as self-prediction and related priming effects have been proposed as explanations (Rogers et al., 2013). The particular characteristics of the Georgian political environment, however, suggest an alternative mechanism for our findings on voter behavior: increased suspicion of regime surveillance, induced by the unexpected arrival of survey enumerators, reduced political participation. This mechanism is plausible given the widespread disbelief in the secrecy of the ballot and the common view that the government punishes political opponents. Citizens in post-Soviet Georgia—many of whom came of age in a totalitarian state—may misconstrue enumerators from the capital asking about politically sensitive issues as agents of the state seeking to gather information about their political sympathies. These types of fears are frequently expressed by Georgian opposition leaders. For example, in 2009, a Georgian opposition legislator claimed that

telephone calls have been made ... to people from a number which is registered to a division of the Ministry of Internal Affairs ... An unidentified individual was conducting a survey from this number, asking whether or not they were planning to attend the 9 April rally, whether they stood in solidarity with [participants of] the rally and whether they would change their mind if the opposition used weapons [at the rally].

(BBC, n.d.)

If visits by enumerators prime these sorts of fears, it is plausible that even a small increase in the perceived risk to the voter or her family could tip the scales against participation on election day.

Some evidence for this mechanism can be observed by examining how treatment effects in the original “complaints” experiment vary by perceptions of the prevalence of electoral coercion, as well as the geographic location of the precinct. We examined three main interactions: urban precincts versus rural precincts, prevalence of employer pressure to vote, and the degree to which employers pressured employees to campaign.¹⁶ If the intervention was indeed being interpreted as political monitoring, then we would expect that the effect would be largest in rural areas where machine politics is more entrenched. The qualitative literature suggests that employment, particularly in the public sector, is often used to mobilize supporters. To that end, we examined whether, as expected, the monitoring effect was largest in areas where survey respondents reported that employer political pressure was highest. We also check whether effect estimates are larger in areas where employers are perceived to place undue pressure on workers to participate in a campaign.¹⁷

The results can be found in Figure 1. The treatment effects all vary in the expected directions, though only one interaction—perhaps not surprising given the low sample size—is statistically significant. The intervention had a larger suppression effect in rural areas, as expected. But particularly striking is the “employer pressure to vote” interaction, which is statistically significant at the 5% level.

Table 4. The average effect of question order on non-response. Treatment respondents were asked political questions at the beginning of the survey.

	“Refuse” + “Don’t Know”	“Refuse”	“Don’t Know”
ATE	1.7	0.85	0.85
SE	1.14	0.62	0.97
95% CI	(−0.53, 3.92)	(−0.37, 2.06)	(−1.04, 2.74)
<i>n</i>	2378	2378	2378

Table 5. Heterogeneity in the effect of question order.

Covariate	“Georgia is a Democracy”		“Elections Conducted Well”		“Close to the UNM”	
Stratum	Below median	Above median	Below median	Above median	Below median	Above median
ATE	4.03	−0.62	3.51	0.05	2.97	0.79
SE	1.83	1.46	1.77	1.47	1.75	1.54
95% CI	(0.44, 7.61)	(−3.48, 2.24)	(0.04, 6.98)	(−2.84, 2.93)	(−0.46, 6.4)	(−2.22, 3.79)
<i>n</i>	1131	1151	1176	1169	1135	1155

In areas where relatively few survey respondents believed that employer pressure to vote for particular parties was widespread, the estimated effect on turnout is close to 0. In areas with a high number of people citing employer pressure as widespread, the effect estimate is −12.7%. Cumulatively, these patterns in the data provide suggestive evidence that the intervention was interpreted by opposition voters as a signal that there would be increased regime attention to their voting behavior on election day.

Politicized data collection and non-response in 2010

In order to probe the mechanism further, we implemented a survey experiment in the 2010 pre-election survey. Our new question: Is the behavioral response of Georgian citizens to a survey instrument different if the survey is clearly associated with politics? Procedurally, we front loaded half of the surveys with the “politics module” (treatment) and placed these directed political questions at the end of the survey for the other half of survey respondents (control).¹⁸ Every respondent was eventually asked political questions, so if the hypothesized mechanism is correct, the survey itself heightened concerns over regime surveillance in both the treatment and control group. The advantage of the randomization of module order is that one can observe if Georgian respondents perform differently on surveys they *know* to be collecting political behavior data compared to surveys where this was *suspected* but not known.

The dependent variable in the analysis that follows is the number of questions that the respondent answered “Don’t Know” or refused to answer.¹⁹ The average treatment effect for the full sample can be seen in Table 4. Asking political questions in the first part of the survey increased the number of “Don’t Know” or “Refuse To Answer” responses by

an average of 1.7, which does not reach standard levels of statistical significance.

Given that a variety of mechanisms could explain how priming politics could induce higher rates of non-response, we are particularly interested in heterogeneous treatment effects: is non-response attributable to treatment higher in precincts with greater distrust of Georgian political institutions and the ruling party? A complication is that it is not possible to use individuals’ responses to a question as a covariate: all answers are “post-treatment.” Since the thrust of our argument is that respondents are reluctant about providing data and answer questions strategically, examining heterogeneity by individual response risks conflating treatment with baseline characteristics. Ecological inference provides a partial solution to this problem: creating covariates, at a precinct level, based on aggregates of survey responses that could divide precincts transparently.²⁰ It becomes possible to make generalizations about which communities are essentially UNM strongholds and which communities do not believe that Georgia is a democracy.²¹ Table 5 splits the sample by the median value of the community level covariates and then displays the average treatment effect within each stratum.

Citizens were less likely to give answers to questions on the politics-loaded version of the survey (treatment) in precincts where many citizens also report they are not living in a democracy, they don’t trust the electoral process, and they are mistrustful of the UNM.²² The specific questions with the the largest difference in non-response rates across the treatment and control conditions also support the favored hypothesis.²³ The five questions with the largest gap in response rates when the politics module was put at the front of the survey are as follows, in descending order:

1. On a scale of 1 to 5 how do you feel about the following parties with “1” meaning “Very Negatively,” and “5” being “Very Positively”: [Answer: The *Alliance for Georgia Opposition Party*]
2. It has been proposed that closed circuit TV cameras be installed outside the polling stations to count the number of voters. Do you support this proposal?
3. On a scale from 1 to 5 where “1” means “Totally Falsified” and “5” means “Conducted Very Well” how do you think the local elections will be conducted?
4. On a scale of 1 to 5 how do you feel about the following parties with “1” meaning “Very Negatively” and “5” being “Very Positively”: [Answer: The *We Ourselves Opposition Party*]
5. On a scale from 1 to 5 where “1” means “Do Not Trust At All” and “5” means “Trust Completely,” how would you assess your trust towards domestic election observers?

In line with our theoretical expectations, front loading the survey with political questions most powerfully affected non-response rates on questions related to monitoring of electoral practices, the credibility of the electoral process, and feelings about opposition parties. The first and fourth largest gap in response rates were questions that asked respondents to rate particular opposition parties (*Alliance for Georgia* was a coalition of opposition parties backing the main opposition candidate in the Tbilisi 2010 mayoral race; *We Ourselves* was a single-issue opposition party focusing on the Abkhazia dispute). The three other questions with large gaps refer to the credibility of the electoral process—with particular reluctance about monitoring, either via cameras and digital recording devices or domestic election observers. We interpret this pattern as additional circumstantial evidence that the more the survey veered into questions of politics, the less willing Georgian citizens were to reveal their opinions.

All of this may, in part, explain why neither area experts nor Georgians themselves predicted that Saakashvili's UNM party would be outmaneuvered in the 2012 election by Bidzina Ivanishvili's Georgian Dream. Mullen (2012) reports that 46% of people contacted for a NDI-funded pre-election survey in the immediate lead-up to the 2012 election refused to tell CRRC survey enumerators who they were planning to vote for. He speculates: “[T]hat number was based on fear; and fear had become an important part of life under the UNM for many people. It was a fear that would lead people to avoid taking risks that could have put them on the wrong side of the authorities.” But it is important to remember that is also plausible that respondents simply did not see themselves in solidarity with foreign-funded democracy professionals doing surveys. While fear is the mechanism that we find most intuitive—and a mechanism that fits neatly into both the rationalist voting framework and anti-regime polemics—other psychological mechanisms are plausible. Simple disgust is difficult to falsify, and

broadly consistent with the heterogeneous treatment effects reported above. As Nodia's epigraph to this paper makes clear, politics in Georgia are widely viewed as dirty work.²⁴ The tendency to perceive politicians as corrupt, and respond negatively and emotionally to misused political power, is ubiquitous in democracies around the globe. The desire by Georgian respondents to avoid dirtying themselves with politics by avoiding it as much as possible could explain the behaviors reported in this study.

Conclusion

In recent years a number of studies have used natural and field experiments to demonstrate that the presence of international actors—as election monitors, scientific observers, or activists—can have a causal influence on local electoral outcomes. Unfortunately, as this study emphasizes, it is likely that well-meaning interventions are occasionally misinterpreted. And while we are certain that our intervention interacted in an unexpected way with prevalent beliefs in Georgian society, we emphasize that our study does not allow us to distinguish precisely *which* inherited beliefs interacted with the interventions to reduce turnout. Did the intervention accidentally prime fear of regime monitoring? Or did it accidentally activate ambivalence, reminding respondents of some other distasteful aspect of Georgian political life? Future research may be able to discover the mechanism at work.

The Rose Revolution is an exemplary case of civil society actors nonviolently seizing control of the state apparatus. One of the unforeseen consequences was a blurring of the boundary of where the regime ends and civil society begins. Well-meaning efforts by foreigners to strengthen civil society, in this context, were likely interpreted by many Georgians as strengthening the state. While this study provides limited grounds for generalization beyond the Georgian case, or even beyond the Saakashvili era, inherited beliefs from the Soviet legacy—a culture of bureaucratic compliance that is exploited by incumbent parties, expectations that private behaviors can and will be monitored and reported to authorities, and a sense that the political sphere is “dirty”—are ubiquitous in the region.²⁵ Residual fear of state surveillance is a common phenomenon in post-authoritarian societies. This study should be seen as an invitation for further research on how authoritarian cultural legacies interact with the standard social science toolkit.

Unexpected results are not unusual in the nascent field experimental literature on democracy promotion and election observation. Hyde (2010), for example, reports the results of a randomized evaluation of election observers in Indonesia, where she found that precinct visits by international observers increased the vote share of the incumbent presidential candidate by 32%. Hyde (2010: 521) proposes several possible explanations, but underscores that the specific result is “somewhat idiosyncratic” and unanticipated. In a voter mobilization experiment in Uganda, Ferree et al. (2011) unexpectedly found that increasing the salience of

the visibility of voting suppressed turnout, particularly among women. These results highlight one of the strengths of field experiments in that they enable researchers to uncover unanticipated effects of interventions. But these results also point to our incomplete knowledge about how interventions common in democracy promotion activities interact with local political practices and cultural beliefs.

This study has policy implications. Investments in improving the secrecy of the ballot—and then funding a public information campaign to publicize this fact—would probably indirectly empower opposition parties by reducing the perceived risks from regime surveillance. But primarily these findings represent a sobering reminder that U.S. policy ambitions outstrip the current state of knowledge on how foreign assistance can alter the relationship between the state and civil society. A healthy dose of humility is justified. While aggregate analysis of democratization aid have found positive effects on democracy indices (Finkel et al., 2008), our results suggest that the consequences of any particular program can be highly heterogeneous and dependent on hard-to-measure local factors, particularly in countries just emerging from long bouts of authoritarian rule. There is much that we do not know about the technologies used by electoral authoritarian regimes to conduct elections in an environment that keeps citizens insecure, atomized, and dependent on the party in power.

Supplementary Material

The replication file is available at: <http://thedata.harvard.edu/dvn/dv/researchandpolitics>

Acknowledgements

We would like to thank John Bullock, Donald Green, Hans Gutbrod, Nahomi Ichino, Charles King, Jennifer London, Jas Sekhon, and the participants of the HALBI working group at the University of California at San Diego and the CRITICS working group at Georgetown for helpful comments that contributed to this paper in its current form. Most of all, we would like to acknowledge the assistance of the Caucasus Research Resource Centers, particularly Aaron Erlich, Nana Papiashvili, Arpine Porsughyan, and Koba Turmanidze, who generously provided logistical support for the implementation of the experiment itself.

Declaration of conflicting interest

The author declares that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. Arguments from psychology and sociology have been formalized, tested, and refined to enrich this rationalist account. For a synthesis and formalization of the literature on voter behavior in fully consolidated democracies see Gerber et al.

(2008). Brody and Page (1973) and Anderson and Glomm (1992) discuss the decision to abstain strategically rather than vote for an “alienating” party in the American context.

2. The journalist Thomas Friedman often told the following joke: “A minister goes to the Syrian dictator Hafez Assad after a national election. Minister: I have excellent news, Mr. President! You won 98.6% of the vote in the election! Less than 2% of the people dissented! What more could you possibly want? Assad: Their names.” Like much black humor, it draws attention to a set of rational fears that are not remotely funny for citizens born into authoritarian rule. This intuition is formalized in Stokes (2005), who provides a seminal model that parameterizes machine ability to monitor individual vote choice. For the related model of “turnout buying,” see Nichter (2008). See also Wedeen (1999) Chapters 1 and 6.
3. The evocative phrase “ritual of consent” is borrowed from Heathershaw (2009: 95).
4. Jones (2013: 136–139) reports that between 1992 and 2010 Georgia received \$3.37 billion in U.S. aid. Jones (2013: 244) notes that after its war with Russia over disputed secessionist territories in the summer of 2008, the United States was quick to pledge Georgia \$1 billion in reconstruction and development aid, and the European Commission promised \$500 million more. At the time that this study was conducted, it was common for political scientists to describe Georgia as an electoral authoritarian (Linz, 2000) or competitive authoritarian (Levitsky and Way, 2002) political system. Still, the year before the experiment was implemented, the 2007 “Nations in Transition” data from Freedom House—which ranks every post-Socialist nation on a 1–8 scale from most to least democratic on 8 different dimensions—rated Georgia at 4.68, the most democratic of the eight post-Soviet states of Central Asia and the Caucasus (the average score of all eight states in this region was 5.96).
5. The late Shevardnadze era was eloquently described as “Potemkin Democracy” by King (2001). The events of the Rose Revolution are described in Wheatley (2005: 181–209). See also De Wall (2005), Nodia and Scholtbach (2006), Areshidze (2007), and Wheatley and Zürcher (2008).
6. When votes were counted after the 2008 Parliamentary elections, the UNM, controlled 119 seats in the 150 member parliament—a very comfortable cushion of safety above the 100 votes that would be necessary to rewrite the constitution. The second largest faction—the United Opposition, consisting of eight small parties and a few individual politicians—had only 17 seats.
7. As shall be discussed below, the percentage of respondents in a particular voting precinct that believed it was widespread varied substantially. 23% of survey respondents in the average precinct in the study reported that these beliefs were widespread, with a standard deviation of 22—which is to say, in some precincts no one reported this fear and in others it was the majority belief.
8. For a lurid English-language account of how UNM tactics were portrayed at the time, see Fairbanks (2012).
9. Certificates were distributed prior to elections for electricity, fuel, or food, with the implied caveat that they could be invalidated by opposition victory at a district level.
10. The survey was conducted by the Caucasus Research Resource Centers (CRRC) as part of their Data Initiative, which seeks to gather comparable, nationally representative

household survey data from Armenia, Azerbaijan, and Georgia. 1837 individuals participated in the survey, with 921 (50.14%) reporting that they believed their vote was secret, 345 (18.78%) reporting that they believed it was not secret, and 501 (27.27%) reporting that they didn't know. The remaining 70 individuals either refused to answer or did not complete the survey.

11. A second intervention encouraged voters to check the public voter rolls in the lead-up to the election, to ensure that their and their families' registration status was accurate. This second intervention was only implemented in urban areas and in fewer precincts. Results from this intervention are presented in the Appendix.
12. According to the 2008, Georgian households have an average 3.5 members, but this figure includes children.
13. Electoral statistics are from the Georgian Central Electoral Commission. The President's party vote share is for candidates of the UNM running for the single member district tier of the national legislature.
14. This estimation procedure is equivalent to inverse propensity score weighting with a known propensity score. Unlike most applications using inverse propensity score weighting, unbiasedness of the estimator is guaranteed in this case because the true propensity score—the sampling weights—are known ex ante (Tsiatis, 2006).
15. These activities are almost always assumed to benefit the regime (Tucker, 2007; Sjoberg, 2012).
16. The variables "political pressure by employer," and "percent reporting pressure to campaign," are aggregated from pre-treatment survey data. The variables are the percentage of survey respondents that said the following phenomena are widespread: "Voters feeling they have to vote a specific way to keep their jobs," and "Employers forcing employees to take part in a campaign."
17. For these variables, with the exception of urban–rural, we split the sample into "high" and "low" samples by the median. For subsample treatment effect estimates, we used point estimates and their associated standard errors from a regression model that allows the treatment effect to vary with a binary interaction variable:

$$y_i = \alpha_0 + \beta_1 T_i + \beta_2 Z_i + \beta_3 (Z_i \times T_i) + \mathbf{X}_i' \gamma + \epsilon_i$$

In this equation, Z_i is the dummy indicator variable for the subgroup, T_i is an indicator for selection into treatment, and \mathbf{X}_i' is a vector of covariates.

18. The politics module for the 2010 local election contained a standard battery of questions that contemporary political scientists working in comparative political behavior use to make inferences about likely voter behavior: questions about past voting history, intended vote choice in the upcoming elections, and the like.
19. This analysis excludes "cascade" questions that were only asked of respondents who answered in a particular way to a previous question. Summary statistics for the dependent variable can be found in the Appendix.
20. When using individual level data to estimate heterogeneous treatment effects, the same general patterns are found.
21. The questions used to construct the community level variables are: "In your opinion, is Georgia a democracy now?";

"On a scale of 1 to 5, where '1' means 'Totally Falsified' and '5' means 'Conducted Very Well' how do you think the local elections will be conducted?"; "Which party is closest to you?" Community level variables are the fraction of respondents who answered that Georgia was a democracy, the fraction that rated the elections as a "3" or above, and the fraction that stated that the UNM (Saakashvili's party) was closest to them. Descriptive statistics for the community-level covariates can be found in Table 10 in the Appendix.

22. The treatment effect is statistically significant at the 5% level in two of the three substrata and statistically significant at the 10% level for the third. The interaction (i.e. the difference between the above and below median strata) is statistically significant at the 5% level in the case of the "Georgia is a Democracy" comparison and significant at the 10% level in the "Elections Conducted Well" comparison.
23. Treatment effects on individual questions are generally small, even for the five questions with the largest difference in response rates. The treatment effect on nonresponse for these five questions was about 4%. While small in absolute terms, the average non-response rate among control respondents was about 8%, making the treatment effect an increase of about 50% over baseline.
24. For a description of politics as "dirty," see Wheatley (2005: 180).
25. For a historically grounded explanation of why a wide stratum of Georgian voters tend to see politics as dirty, see Wheatley (2005: 5). For ethnographic observations on post-totalitarian behaviors, see Paxson (2005) and Harris (2004).
26. The restriction to urban precincts for the voters list intervention was determined before randomization.

References

- Anderson SP and Glomm G (1992) Alienation, indifference, and the choice of ideological position. *Social Choice and Welfare* 9(1).
- Areshidze I (2007) *Democracy and Autocracy in Eurasia: Georgia in Transition*. East Lansing, MI: Michigan State University Press.
- BBC (n.d.) Georgian opposition: Government sowing fear amongst public.
- Brody R and Page B (1973) Indifference, alienation, and rational decisions. *Public Choice* 15(1).
- De Wall T (2005) Georgia and its distant neighbors. In: Coppieters B and Legvold R (eds) *Statehood and Security: Georgia After the Rose Revolution*. Cambridge, MA: MIT Press, pp.307–338.
- Fairbanks C (2012) Georgia's prison rape scandal—And what it says about the Rose Revolution. *The Atlantic*. Available at: <http://www.theatlantic.com/international/archive/2012/09/georgias-prison-rape-scandal-and-what-it-says-about-the-rose-revolution/262720/>.
- Ferree K, Dowd R, Jung D, et al. (2011) Getting out the vote, uganda style: Social and political context and turnout in an African election. Working Paper.
- Finkel S, Perez-Linan A, Seligson M, et al. (2008) Deepening our understanding of the effects of US foreign assistance on democracy building final report. Technical report, USAID.

- Gerber AS, Green DP and Larimer CW (2008) Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review* 102(1).
- Harris C (2004) *Control and Subversion*. London; Sterling, VA: Pluto Press.
- Heathershaw J (2009) *Post-Conflict Tajikistan: The Politics of Peacebuilding and the Emergence of Legitimate Order*. London; New York: Routledge.
- Hyde S (2010) Experimenting in democracy promotion: International observers and the 2004 presidential elections in Indonesia. *Perspective on Politics* 8: 511–527.
- King C (2001) Potemkin democracy. *The National Interest*, Summer.
- Levitsky S and Way LA (2002) The rise of competitive authoritarianism. *Journal of Democracy* 13(2): 51–65.
- Linz JJ (2000) *Totalitarian and Authoritarian Regimes*. Boulder, CO: Lynne Rienner.
- Mullen M (2012) Both sides are wrong. *Georgian Journal*. Available at: <http://www.georgianews.ge/index.php/mullen/11257-both-sides-are-wrong>.
- Nichter S (2008) Vote buying or turnout buying? machine politics and the secret ballot. *American Political Science Review* 102(1).
- Nodia G and Scholtbach AP (2006) *The Political Landscape of Georgia: Political Parties: Achievements, Challenges, and Prospects*. Delft, The Netherlands: Eburon Academic Publishers.
- OSCE (2008) *Georgia parliamentary elections (21 May 2008)—OSCE/ODIHR election observation mission final report*. OSCE Report.
- Paxson M (2005) *Solovoyovo: The Story of Memory in a Russian Village*. Washington DC: Woodrow Wilson Center Press.
- Rogers T, Fox C and Gerber A (2013) Rethinking why people vote. In: Shafir E (ed) *The Behavioral Foundations of Public Policy*. Princeton, NJ: Princeton University Press.
- Sjoberg F (2012) Making voters count: Evidence from field experiments about the efficacy of domestic election observation. Working Paper.
- Stokes S (2005) Perverse accountability: A formal model of machine politics with evidence from Argentina. *American Political Science Review* 99(3).
- Tsiatis A (2006) *Semiparametric Theory and Missing Data*. New York: Springer Verlag.
- Tucker JA (2007) Enough! Electoral fraud, collective action problems, and post-communist colored revolutions. *Perspectives on Politics* 5.
- Wedeen L (1999) *Ambiguities of Domination: Politics, Rhetoric, and Symbols in Contemporary Syria*. Chicago, IL: University of Chicago Press.
- Wheatley J (2005) *Georgia: From National Awakening to the Rose Revolution—Delayed Transition in the Former Soviet Union*. Aldershot, UK: Ashgate.
- Wheatley J and Zürcher C (2008) On the origin and consolidation of hybrid regimes: The state of democracy in the Caucasus. *Taiwan Journal of Democracy* 4(1): 1–31.
- Zwane AP, Zinman J, Van Dusen E, et al. (2011) Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences* 108(5): 1821–1826.

Appendix

Data and methods

To examine the effects of our intervention we employ multiple sets of data: administrative data (collected at the precinct level) and individual panel survey data (aggregated to the precinct level). Administrative data includes voting behaviors, changes to the voter lists, and registered official complaints of election day malfeasance. Panel data provides a rich set of variables that we can employ to probe hypothesized causal mechanisms. We aggregate the survey data to the precinct level, which allows us to examine treatment effect interactions with theoretically interesting pre-treatment covariates.

Electoral statistics are from the Georgian Central Electoral Commission and complaints data was compiled by from the electoral authorities and various election monitoring NGOs. The President's party vote share is for candidates of the UNM running for the single member district tier of the national legislature.

We present treatment effect results for a number of distinct dependent variables: a dummy variable for whether a complaint was filed with accusations of electoral malfeasance in a precinct, the total number of complaints, three different measures of whether a district had changes to the voter list, voter turnout by district, and percentage of votes received by the president's party by district. To estimate the average effect of each intervention on these variables, we employ a simple linear regression of the outcome variable on a treatment indicator, a vector of covariates (in some specifications), and block dummy variables:

$$y_i = \alpha_0 + \beta_1 T_i + \mathbf{X}_i' \gamma + \epsilon_i$$

where y_i is a variable measuring electoral outcomes (complaints, voter list changes, turnout, or vote share) in precinct i . T_i is an indicator variable for either the complaints or voters list treatment, the β_1 coefficient captures the average impact of each intervention on the experimental population. \mathbf{X}_i is a vector of covariates (including block indicators) and ϵ_i is an unobserved error term. For the specifications that include covariates, the adjustment variables (by precinct) are the number of registered voters, voter turnout in 2006, and vote share of the president's party in 2006. In each case we present unadjusted and adjusted results. The standard errors for the regression estimates are calculated using White's heteroskedasticity-consistent estimator.

Map of treatment and control precincts

A visual depiction of precincts assigned to treatment and control groups in the "Complaints" field experiment can be found in Figure 2.

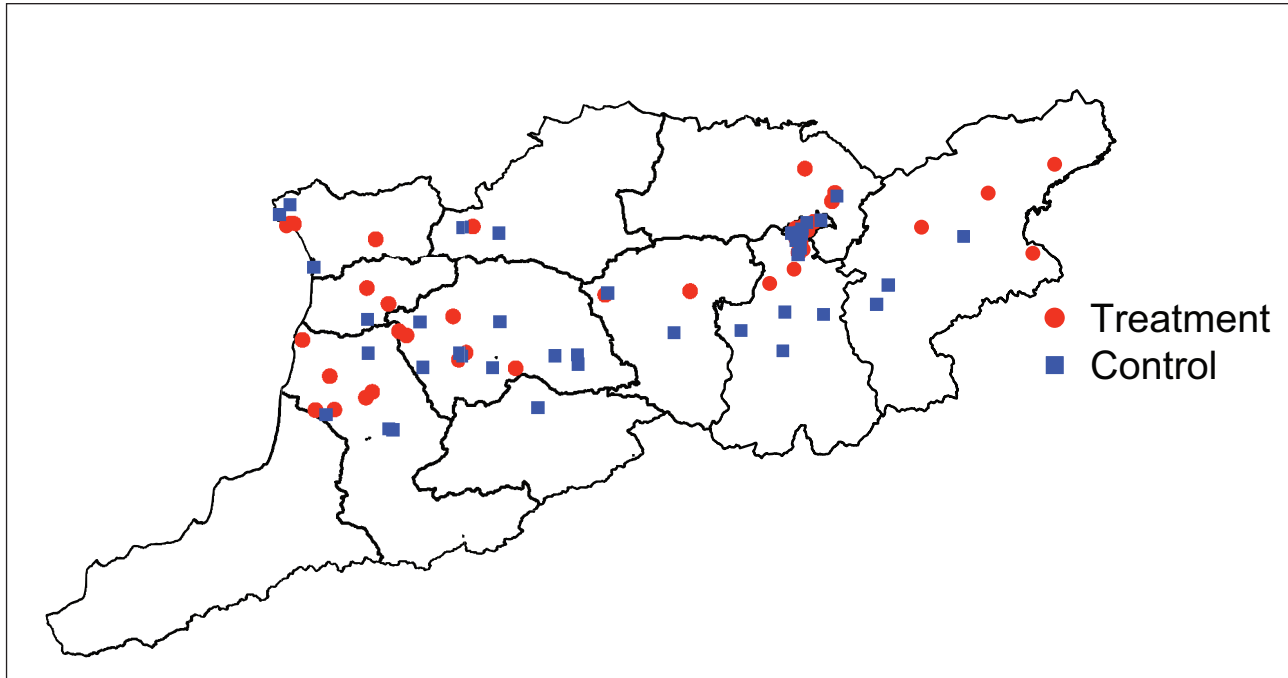


Figure 2. Map of treatment and control precincts for the “Complaints” field experiment.

Table 6. Balance Statistics for the 2008 field experiment.

Covariate	Complaints treatment	Voter list treatment
Registered voters (2007)	33.14 (52.3)	50.33 (65.88)
% turnout (2006)	0.02 (2.01)	−0.92 (2.34)
% vote for presidential party (2006)	−0.55 (1.81)	−3.42 (2.08)
Pre-treatment non-response %	0 (0.03)	0.05 (0.04)
% saying employer pressure is widespread	0.06 (0.04)	0.03 (0.06)
% believing Georgia not a democracy	0 (0.03)	−0.01 (0.05)
% believing ballot stuffing is widespread	0.01 (0.03)	0.04 (0.05)
% believing officials biased to one party	0.02 (0.03)	0.05 (0.05)
% believing voter harassment is widespread	0.01 (0.03)	0.05 (0.05)
% believing activist harassment is widespread	0.01 (0.03)	0.05 (0.04)

Table shows covariate balance on selected pre-treatment covariates for both the complaints intervention ($N = 84$) and the voters list intervention ($n = 48$). Point estimates are from a regression with a treatment dummy and block fixed effects. Standard errors (in parentheses) are heteroskedasticity consistent.

Balance statistics

While randomization ensures unbiased treatment effects in expectation, any given realization of the treatment assignment may result in chance differences across treatment and

control groups. To check that randomization adequately balanced background characteristics in both treatment conditions, we present the treatment effect of both the voters list intervention and the complaints intervention on

Table 7. Balance statistics for “Survey as an Experiment” results.

	Previous turnout %		Previous pres. Party vote %	
	2008	2010	2008	2010
Estimate	−2.26	−1.14	−0.02	−2.47
SE	2.27	3.21	0.02	3.04
95% CI	(−6.71, 2.19)	(−7.42, 5.14)	(−0.06, 0.01)	(−8.42, 3.48)
N	2994	3101	2994	3101

This table shows the balance statistics for the “Survey as an Experiment” results. “Previous turnout” is turnout in the election prior to the survey. “Previous pres. Party vote” is vote share of president’s party in election prior to the survey. Estimates account for unequal probability of selection and control for precinct size, the variable used to sample precincts into the “treatment.”

Table 8. The average treatment effect of the complaints treatment on changes to the voters list. For covariate adjusted results, the adjustment variables are number of registered voters, 2006 turnout, and 2006 vote share of the ruling party. Standard errors are heteroskedasticity consistent and account for blocking.

	Any changes		Early changes		Late changes	
	1	2	3	4	5	6
Voters list treatment	0.29 (0.41)	0.44 (0.35)	0.26 (0.21)	0.36 ^a (0.20)	0.04 (0.26)	0.07 (0.23)
Covariates						
N	48	48	48	48	48	48

Robust standard errors in parentheses.

^aSignificant at $p < .10$.

10 pre-treatment covariates in Table 6. Standard errors are presented in parentheses. There were no statistically significant differences between the treatment and control groups on any of the covariates.

For the “Survey as an Experiment” design, we have only a limited set of variables to check balance. Fortunately, these variables are pre-treatment values of the outcome variables. These are presented in Table 7. None of the estimated differences across the treatment and control conditions are statistically significant at conventional levels.

Voters list intervention results

The intent of the “Voters List” treatment was to encourage voters to check their district voters list and make modifications if errors were found. To check if our intervention had this effect, we used the complete national voter registration database to track any alterations to existing entries in the voter registry by precinct (such changes in address, updating erroneous information, removing names, or the like). We examined three dependent variables: (a) the percent of

voters list entries that were altered at any point in the post-treatment period, but before the election; (b) the percent of voters list entries altered in the week after the intervention (as recorded on 4 and 5 May); and (c) the percent of voters list entries altered immediately before the election, as recorded on 11 and 12 May. Results are presented in Table 8. Due to logistical constraints, the voters list intervention could only be administered in urban precincts, so consequently these estimates are substantially less precise than those from the complaints experiment and are not generalizable to non-urban areas.²⁶ On average, the voters list intervention, when administered without the complaints message, induced an increase of about 0.3–0.4% to the voter registry over the entire post-treatment period. These estimates, however, are generally not statistically significant, with the exception of the covariate adjusted estimated effect on “Early Changes” (significant at the 10% level). The effect appears to be concentrated in the days immediately after the intervention as the “Early Changes” coefficients are substantially larger than the “Late Changes” coefficients.

Additional tables

Table 9. Descriptive statistics for 2008 experiment.

Statistic	N	Mean	St. dev.	Min	Max
Any Change to the Voters List	84	2.318	1.562	0.000	7.840
Registered Voters (April 17, 2008)	84	1141.048	333.761	266	1564
% turnout (2006)	84	47.155	16.314	20	90
% vote for presidential party (2006)	84	73.679	12.938	35	100
% registration changes 1 week after intervention	84	1.478	1.012	0.000	5.312
% registration changes 2 weeks after intervention	84	0.839	0.888	0.000	4.858
% turnout (2008)	84	50.805	15.274	7.712	95.342
% vote for presidential party (2008)	84	47.292	17.839	13.929	87.228
% reporting political pressure by Employers widespread	83	0.235	0.222	0.000	1.000
% reporting employers force employees to campaign	83	0.186	0.197	0.000	0.688
Number of registered voters (election day)	84	1254.881	361.711	269	1922

Table 10. Descriptive statistics for the 2010 survey.

Statistic	N	Mean	St. dev.	Min	Max
Treatment	2,378	0.488	0.500	0	1
# of "Refuse to Answer"	2378	10.579	23.800	0	144
# of "Don't Know"	2378	32.605	30.027	0	195
Closest to the UMN	2378	0.393	0.488	0	1
Believes Georgia is a democracy	2378	0.347	0.476	0	1
Voted for Saakashvili	2378	0.493	0.500	0	1