Election Fairness and Government Legitimacy in Afghanistan*

Eli Berman [a] elib@ucsd.edu
Michael Callen [a**] mcallen@ucsd.edu
Clark Gibson [a] ccgibson@ucsd.edu
James D. Long [b] jdlong@uw.edu
Arman Rezaee [c] abrezaee@ucdavis.edu

Abstract:

Elections can enhance state legitimacy. One way is by improving citizens' attitudes toward government, thereby increasing their willingness to comply with rules and regulations. We investigate whether reducing fraud in elections improves attitudes toward government in a fragile state. A large, randomly assigned fraud-reducing intervention in Afghan elections leads to improvement in two indices, one measuring attitudes toward their government, and another measuring stated willingness to comply with governance. Thus, reducing electoral fraud may offer a practical, cost-effective method of enhancing governance in a fragile state.

[a] University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093 USA
[b] University of Washington, 4333 Brooklyn Ave N, Seattle, WA 98195 USA
[c] University of California, Davis, One Shields Avenue, Davis, CA 95616 USA

* We are indebted to many colleagues, and especially grateful to Glenn Cowan, Jed Ober, Eric Bjornlund, Evan Smith, and Jon Gatto at Democracy International and Nader Nadery, Jandad Spinghar, and Una Moore at the Free and Fair Elections Foundation of Afghanistan for project support and data access in Afghanistan. For comments, we thank Luke Condra, Danielle Jung, Asim Khwaja, Stefan Klonner, David Laitin, David Lake, Margaret Levi, Aila Matanock, Victor Menaldo, Jacob Shapiro, Susan Whiting, and seminar audiences at UC Berkeley, UCLA, University of Toronto, Heidelberg University, and Harvard University. This project would not have been possible without the dedicated research assistance of Randy Edwards, Mohammad Isaqzadeh, and Shahim Kabuli, or the project management of Katherine Levy of the UC Institute on Global Conflict and Cooperation. Our conclusions do not necessarily reflect the opinions of our funders.

** Corresponding author, mjcallen@ucsd.edu, 1-858-822-7455

Highlights:

- We explore if fair elections enhance government legitimacy in fragile states.
- We randomize a fraud-reducing technology in Afghanistan's 2010 election.
- We match the experimental sample with post-election household survey data.
- Improvements of elections' procedural fairness bolsters attitudes toward the state.

Keywords: election fraud, democracy, legitimacy, development, experiment, Afghanistan

JEL Classification: H41, O10, O17, O53, P16

Funding: This work was supported by USAID Development Innovation Ventures (DIV), Democracy International, and the Air Force Office of Scientific Research (grant #FA9550-09-1-0314).

1. Introduction

In this paper, we test whether improving election fairness can improve attitudes, and in particular compliant attitudes, toward government. The context is a national election in Afghanistan, a particularly interesting setting because many Afghans do not view the state as legitimate, in the sense that citizens do not feel obliged to cooperate with government and to comply with its rule.[1] Enhancing government legitimacy is a challenge of general interest to development economics: almost half of the world's poor are projected to live in fragile and conflict-affected states by 2030.[2] These states might effectively increase state capacity if citizen cooperation and compliance could be achieved at lower cost. Moreover, improved attitudes in response to an intervention would provide indirect evidence that electoral fairness and enfranchisement are directly valued by Afghans.

Our analysis builds on a nationwide fraud reduction experiment conducted during the 2010 lower house (*Wolesi Jirga*) parliamentary elections in Afghanistan (Callen and Long, 2015). We fielded a survey (both before and) following that intervention, which finds that respondents in areas that held fairer elections—due to an experimental fraud reduction treatment—reported more favorable views of their government and also more compliant attitudes. We measure attitudes using two indices, each aggregating responses to four or five survey questions. For example, regarding attitudes to government, respondents living near treated polling

---

[1] Greif and Tadelis (2010) define legitimacy of a political authority as "the extent to which people feel morally obliged to follow the authority." The obligation might flow from that authority exhibiting moral standards (Greif & Tadelis, 2010), from procedural fairness (Paternoster, Brame, Bachman, and Sherman 1997; Tyler and Huo 2002; Sunshine and Tyler 2003; Tyler 2006; (Levi, Sacks and Tyler 2009), or from policy outcomes and competent provision of public goods (Guyer 1992; Fjeldstad and Semboja 2000; O'Brien 2002; Bernstein and Lü 2003; Levi 2006; Lake 2010). Dal Bo et al (2010) demonstrate experimentally that procedural fairness increases cooperation.

[2] World Bank estimate, http://pubdocs.worldbank.org/en/154641492470432833/FCV-Main-04-041717.pdf, accessed 28 March, 2019.

stations more frequently agreed that Afghanistan is a democracy, and that members of parliament provide services. Regarding compliance, for instance, respondents near treated stations were more likely to report that paying taxes is important, and that one should inform state security forces about insurgent activity. All of these measures were balanced at baseline, further supporting a causal interpretation of our results.

This study joins a group of experiments testing whether improved service delivery changes citizens' view of government in nascent democracies (Fearon, Humphreys, and Weinstein 2009, 2012; Beath et al 2012; Casey, Glennerster, and Miguel 2012; Humphreys and Weinstein 2012; Burde, Middleton, and Samii, 2016). [3] Separately, several experiments test efforts to strengthen electoral processes through direct observation (Hyde, 2007; Hyde 2009; Enikolopov et al 2013; Asunka et al. 2014; Callen and Long, 2015; Callen, Gibson, Jung, and Long, 2016), generally finding that treatment increases electoral integrity. To our knowledge, however, ours is the first study showing that experimental improvements in the *procedural fairness* of elections improves attitudes toward government.[4]

Our finding that electoral fairness improves attitudes is interesting for four reasons. First, it challenges a view that Afghan political attitudes operate solely

[3] Public attitudes and compliance may help democracies consolidate power through mechanisms familiar to economists. "Tax morale"—a social norm of voluntary compliance with taxation, reduces costs of enforcement (Luttmer and Singhal 2014). For instance, US firms owned by individuals from low tax morale countries are much less likely to pay their US taxes. Voluntary compliance with law enforcement allows improved effectiveness, especially in a community policing setting (Akerlof and Yellen 1994; Bayley, 1994; Kennedy et al 2001 (p. 10)).

[4] Grossman and Baldassarri (2012) provide evidence from a lab-in-the-field experiment showing that subjects *electing* their leaders contribute more in a public goods game, and that the same relationship between the perceived legitimacy of authority and cooperation exists non-experimentally in decisions related to the farmer cooperatives to which subjects belong.

along pre-existing ethnic, class, religious, or ideological lines. Instead, it suggests that fraud reduction can affect attitudes, even in a country with weak institutions and widespread informal governance outside of the state.

Second, in this setting compliance may include sharing information about rebel activity, which could be critical to the very survival of government.[5]

Third, the fairness-enhancing intervention, using "photo quick count" is highly cost-effective relative to traditional election monitoring, and feasible even during a violent election (Callen and Long, 2015). We successfully visited 471 polling centers, with a budget of just over US$100,000. By contrast, the largest foreign mission during this election reached about 85 polling centers, spanning much less of the country, with a budget of approximately US$10 million. Photo quick count has since been used to reduce fraud in South Africa, Kenya, Uganda, and in more recent elections in Afghanistan, broadly suggesting the value of election fraud reduction interventions.

Finally, this study provides insight into policy debates on whether and when to hold elections in post-conflict environments (Commission on State Fragility, Growth and Development, 2018). Calling an election too soon is associated with an increased likelihood of renewed fighting (Brancati and Snyder, 2011), or may result in governments that subsequently restrict further reform (Paris 2004;

[5] Berman et al (2011) summarizes this literature: "Mao Tse-Tung (1937) famously describes the people as "the sea in which rebels must swim," a perspective reinforced by a generation of twentieth-century counterinsurgency theorists (Trinquier 1961; Galula 1964; Taber 1965; Clutterbuck 1966; Thompson 1966; Kitson 1977). Twenty-first century scholarship by practitioners of counterinsurgency reinforces the enduring relevance of noncombatants (Sepp 2005; Petraeus 2006; Cassidy 2008; McMaster 2008). The most prevalent explanation for the importance of garnering popular support is that parties to insurgent conflicts use it to gain critical information and intelligence. Kalyvas (2006) demonstrates that this information increases the effectiveness of both defensive and offensive operations." (p. 771).

Mansfield and Snyder 2007; de Zeeuw 2008;). This may be because elections immediately following conflict are often affected by fraud, for a number of reasons, including the interests of those staging the elections, a lack of trustworthy electoral institutions, and the disorganization of the opposition (Bjornlund 2004; Hyde 2011; Kelley 2011). We find that enhancing electoral fairness, during active conflict, positively affects attitudes; which, in turn, might assist the consolidation of a responsive political authority, rather than its disintegration.

It is important to acknowledge that our attitude measures come from survey questions, so they may not reflect respondents' true views. However, a broad literature correlates survey responses on cultural norms (such as the World Values Survey) to real-world outcomes such as conflict, public good provision, work, and fertility decisions (Fortin 2005, Alesina, Giuliano, and Nunn 2013; Desmet, Ortuno-Ortin and Wacziarg 2017). Additionally, a recent study finds that stated views of Pakistani men about the United States predicts their revealed anti-Americanism in a lab setting (Bursztyn, Callen, Ferman, Gulzar, Hasanain, and Yuchtman 2016). [6]

The paper proceeds as follows. Section 2 describes context, an election in a fragile state. Section 3 describes the intervention, our data, and our research strategy. Section 4 provides results and discusses mechanism. Section 5 concludes.

[6] In a similarly fragile environment, and drawing from multiple sources, Berman, Felter, and Shapiro (2018) document that survey-based measures of civilian attitudes toward government (including willingness to share tips with authorities) respond to violence suffered by civilians the same way that subsequent attacks on government forces do.

2. Background to Afghanistan's 2010 *Wolesi Jirga* election

Afghanistan provides a compelling case which resonates with the challenges of all fragile states attempting to enhance their legitimacy by building effective governance. To this end, promoting elections has been a core component of the United States' policy in Afghanistan. Following the US invasion and the fall of the Taliban in 2001, Coalition forces immediately began developing democratic institutions, hoping to promote stability by establishing a functioning central government which had been undermined by two previous decades of internecine conflict, civil war, and Taliban rule. Soon after the invasion, Coalition forces empaneled a *Loya Jirga* to create a new constitution. In 2005, Afghans voted in the first elections for the lower house of parliament (*Wolesi Jirga*). In 2009, Hamid Karzai won re-election as president amid claims of rampant election fraud (Callen and Weidmann 2013). General Stanley McChrystal, NATO commander in Afghanistan at the time, argued that fraud in that election created a "crisis of confidence" in the government, which would ultimately undermine the war effort (McChrystal 2009).

Afghans had good reason to believe that the 2010 parliamentary elections would not be fair. The international community nearly unanimously blamed the IEC for failing to prevent widespread vote manipulation during the 2009 presidential race: Politicians and their agents intervened at all levels, from stuffing ballot boxes and inflating counts at polling centers to manipulating counting processes at the provincial and national levels. So flawed was the 2009 election that while Hamid Karzai claimed victory initially, the IEC would not certify the results, leading to a diplomatic crisis and a second round run-off that the opposition boycotted. The government failed to implement reforms before 2010 elections, so

that it, the IEC, and international donors expected these problems to recur (Democracy International, 2010); in section 3 below, we provide evidence of fraud in numerous parliamentary contests (Callen and Weidmann, 2013).

We study the effects of a fraud-reducing intervention implemented during the 2010 *Wolesi Jirga* elections, which occurred amid a growing insurgency and a U.S. commitment to begin withdrawing troops the following year. The international community viewed these elections as a critical benchmark in the consolidation of democratic institutions given doubts about the Karzai government's ability to exercise control in much of the country and the growing influence of the Taliban. The Taliban significantly increased their attacks on security forces and election officials during this period (Condra et al., 2019). Despite that direct threat of violence, roughly five million voters cast ballots on election day.

Afghanistan's 34 provinces serve as multi-member districts that elect members of the *Wolesi Jirga*. Each province is a single electoral district. The number of seats allocated to a province is proportional to its estimated population. Candidates run "at large" within the province, without respect to any smaller constituency boundaries. Voters cast a Single Non-Transferable Vote (SNTV) for individual candidates, nearly all of whom run as independents.[7] Winning candidates are those who receive the most votes relative to each province's seat share. For example, Kabul province elects the most members to Parliament (33) and Panjsher province the fewest (2). The candidates who rank one through 33 in Kabul and one through two in Panjsher win seats to the *Wolesi Jirga*.

[7] SNTV systems provide voters with one ballot that they cast for one candidate or party when multiple candidates run for multiple seats. If a voter's ballot goes to a losing candidate, the vote is not re-apportioned. During this election, parties played only a very minor role in Afghan politics. The SNTV system was adopted partly to dissuade their creation.

SNTV rules create strong incentives for fraud. SNTV in large districts, without political parties, generates dispersion of votes across candidates: vote margins separating the lowest winning candidate from the highest losing candidate are often small. This creates a high expected return for even small manipulation for many candidates. (In contrast, electoral systems with dominant parties guarantee victory with large vote margins, and so the many non-viable candidates are less likely to rig results.) These strong incentives to manipulate voting were compounded by a weak election commission, which had failed to prevent widespread fraud during the 2009 presidential election. We document clear evidence of election fraud in the experimental sample studied here during the 2010 parliamentary contest.

3. Research design and data

Our results use data from a randomized evaluation of an original anti-fraud monitoring package that some of us conducted during Afghanistan's 2010 *Wolesi Jirga* election (Callen and Long, 2015), and which we recount here. In this section, we first revisit that anti-fraud monitoring experiment as a prelude to investigating the effect of that fraud reduction on attitudes toward the Afghan government.

On election day, and again on the day after, a team of Afghan researchers traveled to an experimental sample of 471 polling centers (7.8 percent of polling centers operating on election day). Because Afghanistan was an active war zone during this period, we selected polling centers that met three criteria to ensure the safety of our staff: (i) achieving the highest security rating given by the International Security Assistance Force (ISAF) and the Afghan National Police (ANP); (ii) being located in provincial centers, which are much safer than rural

areas;[8] and (iii) being scheduled to operate on election day by the Independent Electoral Commission (IEC). Figure 1 maps our experimental sample.

[Figure 1 about here]

In a randomly chosen 238 of those polling centers, [9] researchers delivered a notification letter to Polling Center Managers (PCMs) between 10AM and 4PM, during voting. Researchers then visited all 471 polling centers the following day to photograph the publicly posted election returns forms (which we term "photo quick count"). [10] Letter delivery constituted the experimental treatment. The letter announced to PCMs that researchers would photograph election returns forms the following day (September 19) and that these photographs would be compared to results certified by the IEC. (Neither treatment nor control sites would be affected by measurement the day after the election, as polling staff were absent.) Figure 2 provides a copy of the notification letter in English (an original in Dari is attached as Figure 3). PCMs were asked to acknowledge receipt by signing the letter. PCMs at seventeen polling centers (seven percent of those receiving letters) refused to sign. A polling center was designated treated if the PCM received a letter (*Letter Delivered = 1*, Table 1).[11]

To measure the fairness of the election, our field staff recorded whether election materials were stolen or damaged during polling. We also examined the

[8] Given budget and security issues, we only deployed researchers in 19 of 34 provincial centers. Thus the sample is not nationally representative but biased toward safer areas. It does however cover each of Afghanistan's regions, including those with a heavy Taliban presence. See Figure 1.
[9] We stratified treatment on province and, in the 450 polling centers for which we had baseline data (we added an additional 21 to the experimental sample after baseline on obtaining additional funding), we also stratified treatment on the share of respondents from the baseline survey reporting at least occasional access to electricity and on respondents reporting that the district governor carries the most responsibility for keeping elections fair.
[10] Of 471 polling centers, six did not open on election day. We drop these from our analysis.
[11] Results below are robust to redefining treatment as receiving *and signing* the notification letter.

reason that materials went missing. Staff were careful to investigate irregularities by interviewing local community members (while not engaging IEC staff, so as not to create an additional treatment in the original fraud experiment). We received reports of candidate agents stealing or damaging materials at 62 (13 percent) of the 465 operating polling centers, a clear violation of the law. We define *Election Tally Removed* as an indicator equal to one if materials were reported stolen or damaged by a candidate agent at a given polling center.

We have several reasons to think that stealing or damaging tallies reflects an intention to manipulate the ballot aggregation process. Many of the Electoral Complaints Commission (ECC) complaints reported in (Callen and Long, 2015) speculated that the purpose of stealing materials was to take them to a separate location, alter them, and then reinsert them into the counting process. Alternatively, candidates might seek to destroy all evidence of the polling center count, and then manufacture an entirely new returns form at the Provincial Aggregation Center.

These activities could plausibly send a signal to communities in the vicinity of the polling center regarding the fairness of the election. Appendix Figure 1 provides a picture of citizens looking at a tally sheet depicting the polling outcomes.

The treatment (i.e., delivery of a notification letter) induced dramatic reductions in three separate measures of fraud: the removal or defacement of a required provisional vote tally return form (*Election Tally Removed)*; votes for candidates likely to be engaged in fraud based on their political connections[12] (*Votes*); and that same candidate gaining enough votes to rank among the winning candidates in that polling station (*Enough Votes to Win Station*). Table 1 reports

[12] The political connections of candidates were coded in advance. We surmised that a connection to a provincial polling aggregator was a predictor of engagement in fraud. See (Callen and Long, 2015) for details.

estimates of the effect of treatment on these three measures, reproducing results reported in (Callen and Long, 2015), adjusted to include only the sample of polling centers where we conducted our post-election survey. Treatment reduced the damaging and theft of forms by about 11 percentage points (columns 1 - 3), votes for candidates likely to be engaged in fraud (*Treatment x Provincial Aggregator Connection* = 1) by about seven (columns 4 – 6) and the likelihood that those candidates would rank among winning candidates by about 11 percentage points (columns 7 – 9). These results represent large treatment effects of the intervention on measures of fraud. Tally sheets are highly visible, as, by law, they need to be posted on the outside of the polling center. Because they are the only means immediately visible to communities regarding how they voted, many citizens check them (see Appendix Figure 1 for an example).

[Table 1 about here]

3.1 The Post-Election Survey

To measure the effect of increased election fairness on attitudes toward government, the focus of this paper, we combine the results of the letter intervention with data from a post-election survey. We conducted a baseline in August 2010, the month before the election, followed by a post-election survey in December 2010, roughly three months after the election, deliberately timing it to be immediately after the Independent Election Commission certified final results. This timing ensured that election outcomes would be both finalized and still potentially salient in the minds of voters. Respondents came from households living in the immediate vicinity of 450 of the 471 polling centers in our experimental sample, for a total of 2,904 respondents. To obtain a representative sample of respondents living near polling centers---generally neighborhood landmarks such as mosques, schools or markets---enumerators employed a random walk pattern starting at the

polling center, with random selection of every fourth house or structure until either six or eight subjects had been surveyed. In keeping with Afghan custom, men and women were interviewed by field staff of their own gender. Respondents within households were randomly selected using Kish grid. The survey had 50 percent female respondents. Enumerators conducted the survey in either Dari or Pashto.

We measure attitudes toward government using individuals' responses to nine questions. The first four questions (1 through 4 below) probe attitudes toward government; the remaining five questions (5 through 9 below) measure compliance with governance. We use these four and five responses respectively to address our primary two research questions, since any single question is unlikely to fully capture citizen's views.[13] In all three cases, we design indices [following Kling, Liebman, and Katz (2007) and Casey, Glennerster, and Miguel (2012)], standardizing outcomes by subtracting means and dividing by standard deviations so that each is measured in standard deviation units. Indices are then simply the arithmetic average of the standardized outcomes.[14]

1. *Who is mainly responsible for delivering services in your neighborhood (RANDOMIZE ORDERING): the central government, your Member of Parliament, religious or ethnic leaders, the provincial government, or the community development council?*

The variable *MP Provides Services* is equal to one if individuals respond "Member of Parliament" to this question. This question is intended to capture whether

[13] We did not specify these two sets of outcomes in a registered pre-analysis plan, although we designed these survey questions to measure the effect of election fraud on attitudes related to legitimacy. The timing of the survey (immediately after election outcomes were certified) and its' content (principally questions on attitudes toward government) should also indicate that our intent was to measure attitudes related to legitimacy of government.

[14] We have also weighted these indices by the covariance of the standardized outcomes within each index. No results in the paper are changed meaningfully in magnitude or significance by weighting.

respondents link service provision to the elected government official voted on in this particular election, rather than to more traditional local religious or ethnic leaders or to other bodies (largely unelected) whose standing should not be as directly affected by the 2010 elections—the central government, provincial government, and community driven councils.[15]

2. *In your opinion, is Afghanistan a democracy or not a democracy?*

Afghanistan is a Democracy is an indicator equal to one for the response "is a democracy." This question could be interpreted by respondents narrowly, in the technical sense of democratic procedures being followed, or broadly as a positive endorsement of government. We cautiously choose the latter interpretation below.

3. *Do you think that voting leads to improvements in the future or do you believe that no matter how one votes, things never change?*

Voting Improves Future is an indicator set equal to one for the response "improvements." This measure aims to capture whether citizens believe that voting materially affects their future. If the government is viewed as incompetent, or elections are viewed as hopelessly marred by fraud and mismanagement, then citizens should not hold this attitude.

4. *Does the central government do an excellent, good, just fair or poor job with the money it has to spend on services?*

Gov. Ext. or Good Job of Prov. Serv. is an indicator set equal to one responses "excellent" or "good" to this question. This question directly assesses whether citizens believe that government is effectively providing services.

5. *In your opinion, how important is it for you to share information about insurgents to the Afghan National Security Forces (ANSF) (for example, pending IED attacks*

---

[15] Note that "central government" is generally understood to be the unelected central bureaucracy, not the national parliament, or the two combined. The same is true for the provincial government.

or the location of weapons caches): is it very important, somewhat important, or not at all important?

Important to Report IED to ANSF is an indicator set equal to one for responding "very Important" or "somewhat Important." The question is intended to measure whether or not citizens comply with ANSF requests for information, a critical component of the ANSF's ability to provide security. A substantial policy and research literature related to counterinsurgency argues that citizens' support for the government, and, consequently, their willingness to undertake the costly action of providing information to government forces, determines who wins intrastate conflicts (Berman, Felter, and Shapiro, 2018).

6. *If you had a dispute with a neighbor, who would you trust to settle it (RANDOMIZE ORDERING): head of family, police, courts, religious leaders, shura, elders, ISAF, or other?*

Police Should Resolve Disputes is an indicator set equal to one for the response "police." This question reflects compliance with police adjudication of disputes, as opposed to informal dispute adjudication mechanisms (which might include the Taliban).

7, Courts are in principle another relevant institution, but much less so in Afghanistan, because they are essentially absent in much of the country. Nonetheless, we consider the potential relevance of courts, defining *Courts Should Resolve Disputes* as an indicator set equal to one for the response "courts."

8. *In your opinion, how important is it for you to pay taxes to the government: is it very important, somewhat important, or not at all important?*

Paying Taxes is Some. or Very Imp't is an indicator set equal to one for the responses "very important" or "somewhat important." This directly measures whether citizens voluntarily comply with a government rule that otherwise would be extremely costly for government to enforce.

9. *Let us suppose that your friend has been accused of a crime. Who do you trust to determine whether your friend is guilty: head of your qawm or the Afghan government?*

Trust Afg. Gov. to Determine Guilt is an indicator set equal to one for the response "Afghan government." This measures whether citizens trust the government to make costly determinations regarding a persons innocence. Though this is literally a question about attitudes, we interpret it as an indicator of willingness to bring criminal cases to government.

[Table 2 about here]

Table 2 reports summary statistics for these variables from the post-election survey. The data depict a country with uneven support for government. About 67 percent of respondents view Afghanistan as a democracy, while only 18 percent prefer the police as their primary means of dispute adjudication. 20 percent of respondents believe that their Member of Parliament is responsible for providing services, while 93 percent respond that reporting an impending attack to the ANSF is important.[16] Sixty-one percent believe voting will improve their future, 84 percent believe that paying taxes is somewhat or very important, and 53 percent would trust the Afghan government to determine the guilt of a friend. Across these measures, attitudes toward government leave room for improvement.

Table 2 also reports high incidence of electoral malpractice at polling stations linked to survey respondents. At 13.5 percent of polling stations our staff recorded a report of candidate agents removing tallies (*Election Results Form*

[16] For ease of exposition, we restrict our sample in Tables 2 through 5 to 2,403 respondents who provide some response to the nine questions used across our two hypotheses. This keeps the number of observations fixed across outcomes. For results without this restriction see Appendix Tables 2 through 4 and 6. There are no meaningful differences. Furthermore, Appendix Table 1 reports that no differential attrition by treatment status into the restricted sample used in Tables 2 through 5.

Removed). The procedure for measuring who was responsible for tally sheets was performed identically in treatment and control polling centers. It involved sending an enumerator to the polling center the day after the election, checking whether the form was missing, and then visiting households in the vicinity of the polling center to enquire who had removed the form. A similar picture emerges from the baseline interviews, collected in August 2010, which we return to below.[17] Our data also include two important descriptors of the environment that the elections were held in: the number of local military events tracked as by International Security Assistance Force (ISAF) (from their Combined Information Data Network Exchange (CIDNE) database), with a mean of 2.5; and whether or not the polling station was visited by an international monitor on election day, which occurred in 16.3 percent of the sample (from Democracy International).

3.2 Baseline survey

We conducted a baseline survey in August 2010, one month before the election, to inform treatment assignment for the intervention. Here, we use data from the baseline survey to demonstrate randomization verification and support inferential claims regarding the effect of fraud reduction on attitudes reporting in the post-election survey discussed in section 3.2 and for which we had comparable measures at baseline. [18] Table 3 reports summary statistics and verifies balanced randomization of our anti-fraud intervention between treatment and control polling

[17] Similarly to the post-election survey procedure, in sampling respondents for the baseline enumerators were told to begin at the polling center and survey either 6 or 8 subjects. Surveys were conducted in individuals' homes. Enumerators adhered to the right hand rule random selection method and respondents within houses were selected according to a Kish grid (Kish, 1949).

[18] Similarly to the post-election survey procedure, in sampling respondents for the baseline enumerators were told to begin at the polling center and survey either 6 or 8 subjects. Surveys were conducted in individuals' homes. Enumerators adhered to the right hand rule random selection method and respondents within houses were selected according to a Kish grid (Kish, 1949).

stations using the baseline survey. Further, in Table 3, treatment status is balanced across baseline measures for all key outcomes used in the study, including our nine key outcomes (examined in Tables 4 and 5), which we expect given random assignment to treatment.[19] We also find no evidence of imbalance on other measures that might be relevant to attitudes, including military events in the vicinity and visits by international monitors (discussed in section 3.4).

[Table 3 about here]

Preserving respondent anonymity was a high priority. Consequently, we obtained only verbal (as opposed to written) consent and avoided questions that would allow subjects to be easily identified based on their responses (including specific location/address questions). This means we cannot know whether baseline and post-election respondents are the same. We did, however, design our survey protocols to try to encourage overlap between baseline and post-election surveys. It is therefore instructive to see how much overlap we observe matching on time-invariant demographics. To measure overlap, we perform a fuzzy match between the baseline and post-election surveys on polling center catchment, gender, years of education, ethnicity, language, and whether a respondent reports being born locally. We force matches to be exact on polling center and gender. Of the 3,048 interviews conducted in the post-election survey, 341 (11 percent) cannot be matched to the baseline, and so definitively are new respondents. 90 (3 percent) match perfectly on these measures, and so are very likely to be the same respondents. If we accept matches above a matching score of 0.80 (using Stata's reclink command), 1285 match (42 percent). The remaining 58 percent are all above a 0.5 matching score. Note that since treatment was at the polling center level

[19] The only exception is that we did not collect baseline data for the "Trust Afghan Government to Determine Guilt" question.

rather than the individual, it is not essential for inference that we have the same population post-election as baseline. Without a panel, we cannot rule out, however, that there was an imbalance on outcomes in the post-election population at baseline that we are interpreting as a treatment effect. We think this is unlikely, though, given that we observe no mean differences between treatment and control respondents at baseline, and the extent of overlap documented here.

3.3 Additional administrative data sources

In many of our main tests and robustness checks, we draw from administrative sources to create two additional variables that help characterize each polling center on election day: the number of local military events tracked as by International Security Assistance Force (ISAF) (from their Combined Information Data Network Exchange (CIDNE) database), with a mean of 2.5; and whether or not the polling station was visited by an international monitor on election day, which occurred in 16.3 percent of the sample (from Democracy International) (descriptives shown in Table 2). We include these as controls in main tests and randomization verification (Tables 1, 3, 4), and robustness checks in the Appendix.

We employ additional administrative data from the Free and Fair Elections Forum of Afghanistan (FEFA), a national and independent election monitoring organization, to explore mechanisms linking different types of fraud reduction with citizens' attitudes. FEFA sent Afghan monitors to a substantial share of polling centers across the country, of which 393 overlap with our 459 experimental sample. Their data report whether PCMs adhered to a range of official protocols. These data, therefore, allow us to investigate whether delivering treatment letters affect other dimensions of PCM performance and whether the mechanism linking our fraud reduction experiment with citizens' attitudes likely occurred related to the

posting of tallies. We attach the survey instrument filled out by the FEFA observers as Appendix B.

4. Estimation Strategy and Results

Assignment to treatment is random. So the following equation consistently estimates the effect of delivering the letter (which alerts the polling station manager of monitoring) on our measures of attitudes:

$$Attitude_{ic} = \gamma_1 + \gamma_2 LetterDelivered_c + \gamma_3 X_{ic} + \varepsilon_{ic}$$

where i denotes an individual respondent, c indexes a polling center (specifically, the neighborhood in the immediate vicinity of the polling center), attitudes are measured as described in the discussion of Table 2 above, $LetterDelivered_c$ is an indicator equal to one for polling centers that received the letter and X_{ic} is a vector of covariates described in Table 2. All specifications reflect our assignment strategy by including stratum dummies as suggested by Bruhn and McKenzie (2009).[20] All regressions cluster standard errors at the polling center level.

[Table 4 about here]

Table 4 reports our main results, testing whether notification letters improved (i) perceptions of government, (ii) compliant attitudes toward government, and (iii) an "*All Outcomes*" index of attitudes in general. Since assignment of the fraud-

[20] Alternatively, we have tried collapsing our data to polling center level averages to create a pseudo-panel of polling centers. That allows us to run a difference-in-difference version of the same estimating equation, but with polling center fixed effects, where the first difference is between treatment and control polling centers and the second difference is between baseline and post-election. We find very similar results taking this approach (results available on request). This is not surprising, given the high degree of balance we find on baseline outcomes in Table 3.

reducing treatment is randomized, we are not concerned with selection bias or other omitted variable biases affecting our results.

We answer both research questions in the affirmative. In column (1) we find that notification letters improved attitudes toward government by 0.054 standard deviations. That result is statistically significant. It is robust to the addition of both stratum fixed effects, and a broad set of control variables, as reported in columns (2) and (3) (as expected with random assignment of treatment, –though fixed effects and controls do improve precision). In column (4), we similarly find that notification letters increased compliant attitudes toward government by 0.068 standard deviations. That estimate is also robust to including stratum fixed effects (column 5) and additional covariates (column 6). It is not surprising then that we find a 0.062 standard deviation increase in general attitudes when using the *All Outcomes* index.

Table 5 reports the results of disaggregating the two indices into responses to each of the nine questions, using specifications including stratum fixed effects and additional covariates (as in columns (3), (6) and (9) in Table 4). In addition to reporting treatment effects, we also report multiple hypothesis-adjusted p-values for each hypothesis test. We adjust across the two indices to control the familywise error rate (FWER) computed following Westfall and Young (1993) and Anderson (2008); within each index group, we adjust to control the false discovery rate (FDR) computed following Benjamini, Krieger and Yekutieli (2006) and Anderson (2008). For all nine survey questions, the estimated treatment effect is positive. This effect remains significant or very close (adjusted p-values <=.11) in three cases---MP Provides Services,[21] Important to Report IED to ANSF, and Paying Taxes is

[21] We also estimated treatment effects on dummy variables set equal to one when respondents indicate supporting the Central Government, Provincial Government, religious or ethnic leaders, or local Community Development Council as the unit that should provide services. These results can

Somewhat or Very Important. We view these outcome-level results as exploratory and thus will not interpret them individually.

[Table 5 about here]

The largest standardized effects are on the variables MP Provides Services, Paying Taxes is Somewhat or Very Important, and Important to Report IEDs to the ANSF. Following on the discussion of these survey questions in Section 3 above, there is a strong argument that these three measures are among the most conceptually important. In Afghanistan, several authorities overlap in providing services, which we enumerated when asking the question. Respondents identify MPs, the group contesting office in this election, as being more important for providing services when the election was cleaner. Second, paying taxes is generally an important measure of support for the government, as it is critical for governments to operate, yet achieving compliance is challenging, so enforcement often depends on citizen attitudes. So it is indeed consequential if electoral fraud reduction improves attitudes to paying taxes. Last, we find that cleaner elections make citizens more willing to report IEDs. This relates specifically to `hearts and minds' theories of counterinsurgency, which posit that more effective governance should make citizens more willing to share information.

To allow for better interpretation of our results, Appendix Table 5 provides non-standardized effects on each of the nine attitudes (and includes the standardized indices for ease). We can see that effect of treatment on MP Provides Services is

---

be found in Appendix Table 7. The only significant positive effect is on indicating Member of Parliament. There is also significant negative treatment effect on indicating the Provincial Government. This negative effect is not surprising since these choices are exclusive—there is a simple adding up constraint. We might be more concerned if the negative treatment effect on Central Government offsets the positive effect on MPs if people might think of the Central Government and MPs as interchangeable. However, if we combine these two indicators, the result in Table 5 on the Perceptions of Government Index weakens but remains significant at the 10 percent level.

4.6 percentage points, with 17.3 percent of respondents answering yes in the control group. This is a 27 percent impact. For Paying Taxes is Somewhat or Very Important, the treatment effect is 4 percentage points on top of a control mean of 82 percent, or a 4.9 percent increase. For Important to Report IEDs to the ANSF, the treatment effect is 2.2 percentage points on top of a control mean of 92.3 percent, or a 2.38 percent increase. While we are not aware of similar estimates in the literature to compare these to, they seem economically meaningful.

We report experimental evidence that the fraud-reduction intervention improved attitudes toward government. Taken together, these results indicate that even in Afghanistan—a nascent democracy with weak institutions, improving electoral fairness has consequential effects on attitudes.

How sensitive is *our main outcome index result* to *particular attitudes?*

It is natural to wonder whether the effects for the main indices reported in Table 5 are being being driven by a small number of component variables, namely MP Provides Services, Important to Report IED to ANSF, and Paying Taxes is Somewhat or Very Important. We check on robustness of the "*All Outcomes*" index by recalculating it several times, first removing each of these variables, one by one, then removing each possible pair of the three, and finally removing all three. When we remove MP Provides Services (=1) from the index, we estimate a treatment effect of 0.053 with a standard error of 0.017. When we remove Important to Report IED to ANSF from the index, we obtain a coefficient of 0.058 with a standard error of 0.018. When we remove Paying Taxes is Somewhat or Very Important, we obtain a coefficient of 0.056 with a standard error of 0.017. In all three cases, we obtain a result very similar in magnitude and still significant at the one percent level. When we remove pairs of these attitudes, we maintain one percent significance, with coefficients between 0.046 and 0.052. When we remove all three

attitudes simultaneously, we obtain a coefficient of 0.041 with a standard error of 0.019, which is still significant at the 5 percent level. We interpret the robustness of the "*All Outcomes*" index to exclusion of individual variables as evidence in support of a broader change in attitudes.

4.1 Does fraud reduction improve attitudes if perceived as an external intervention?

Last, we explore two concerns about interpreting these results, should respondents perceive that fraud reduction was an external intervention.

First, survey respondents might provide more favorable responses in the treatment group because of an experimenter demand effect, if they realized that the survey was fielded by the researchers who are responsible for the treatment.

Second, one might imagine that an intervention known to be external (and therefore perhaps temporary) should not change attitudes toward government. Why would voter attitudes toward their government change if they believed that a non-governmental actor, such as foreign election monitors or foreign donors, were the cause of improved procedural fairness?

To address both these concerns the post-election survey asked respondents if they were aware that international monitors visited their local polling center on election day. Practically, this is challenging for respondents to know. Recall that the intervention consisted of our enumerators (Afghan nationals, although accredited observers of an international organization) paying each polling center a short visit to hand-deliver a notification letter to the PCM. For a survey respondent to be aware that this happened, they would need to either observe the intervention directly, or be informed by polling center staff or other individuals who observed

the intervention. Indeed, only about 10% of respondents in the treatment group (and none in the control) reported that they were aware of the treatment.

Appendix Table 8 repeats the analysis of Table 4, estimating the same equation with an added indicator variable $Aware\ of\ Delivery_{ic}$, (which takes the value one if the respondent is in the treated sample and responded that they had knowledge of the treatment).[22] Estimated coefficients on the interaction of that variable with treatment are small and statistically insignificant, with a slightly negative point estimate on perceptions (1.1 percentage points) and a zero on compliance (0.00 percentage points). We do not find statistically significant evidence that respondents aware of delivery had a lower the treatment effect for either of the indices, though the point estimate suggests a smaller compliance effect for the *aware* sample (column 6).

Of course, these estimated interaction effects are not experimental, since awareness was not randomly assigned within the treatment group. They are subject to possible selection bias, since those aware of treatment might have *a priori* different outcomes. That would be true, for instance, if the *aware* were keen observers of local politics and were therefore more cynical about Afghan democracy. In addition, there are no means to identify a comparison group in the control sample who would have been aware of treatment had they been treated.

In summary, the small subsample who would be aware of external treatment if treated do not exhibit statistically significant evidence of smaller local average treatment effects relative to the remainder of the sample (i.e., that fraud-reduction improves their attitudes less than it does for others). So we find no evidence of

[22] This variable always takes the value of zero in the control sample. Thus we cannot separately identify the impact of awareness on outcomes in the control group.

experimenter demand effects or of differential response in attitudes to an intervention perceived as external.

More importantly, the local average treatment effects of the *unaware* show large and statistically significant improvements in attitudes due to fraud reduction, as we found in Table 4 for the pooled sample of aware and unaware respondents.[23]

4.2 How Did Treatment Affect Attitudes?

For electoral fraud reduction (i.e., delivery of the letter to PCMs) to affect attitudes (for those respondents unaware of the intervention) it must change some type of fraud which respondents notice. But there are many types of fraud, so which is the most plausible mechanism by which treatment affected attitudes?

In Section 3 above we emphasized one type of fraud which would be very noticeable to citizens, destruction of tally forms, and demonstrated treatment effects on tally form removal (including destruction) (Table 1). Communities learn how they voted by observing tallies pasted outside of polling centers. They are an object of great interest for many Afghans. Elections provide one of very few venues for Afghans to exert agency over a highly centralized government. Correspondingly, turnout is high (despite the threat of violence), and returns are an important topic of conversation. Appendix Figure 1 displays citizens reading a tally form.

Representatives of candidates illegally removed or destroyed tally forms at 43 out of 225 control polling centers but at only 19 out of 234 treated centers. Ensuring that the tally form was *not* torn down is one of the clearest ways a PCM can demonstrate careful management of the election to the community. Indeed, the

[23] A policy implication is that replication is best done by a local rather than an external agency, as treating the *unaware* sample shows unequivocally positive effects on attitudes.

letter specifically requests that they do so, but does not make reference to other measures of polling center management. We have argued that this is the primary mechanism linking the delivery of letters to improved perceptions of the government, as we can show a treatment effect, and it is clearly noticeable.

Additional data allow us to consider mechanisms by which *other* possible types of fraud could have affected attitudes. Recall that FEFA inspectors reported on 393 of our 459 experimental polling centers. We focus on ten additional proxies for fraud recorded by FEFA (campaign materials within 100m of polling station, intimidation, fraud complaints reported, unauthorized persons in polling center, threats during voting, unused and spoiled ballots, FEFA observers allowed, counted votes reflected exactly on tally sheet, tally posted at end of day, results list distributed to observers), and spoiled ballots, which are recorded separately by the IEC. We focus on those ten FEFA measures because they correspond to the types of PCM misbehavior that FEFA deemed important enough to require filing an incident report. While this provides an ex ante rationale for the outcomes we select, this analysis should be treated as exploratory. Importantly, many of these measures could have been recorded before letters announcing monitoring were delivered to polling centers, excluding a possible treatment effect.

First, we check whether the removal of tally forms by candidate agents is correlated with these 11 measures in the absence of treatment (i.e., in control polling centers (Appendix Table 9), and then we check if treatment affected any of these measures (Appendix Table 10).

Appendix Table 9 reports on the 207 of our control polling stations for which FEFA data are available. Note first that even in the absence of tally sheet removal, many types of irregularities are common: 27% of polling stations have campaign materials within 100m, 5.3% report intimidation, 9.9% had unauthorized persons in the polling stations, and only in 77.8% could FEFA staff observe without

difficulties. In that sample a removed tally sheet (as recorded by our election day enumerator) weakly predicts an increased incidence of three other measures of fraud: campaign materials within 100m of the polling center, spoiled ballots, and unused or spoiled ballots. It also predicts decreased incidence of two other measures: reported intimidation and official complaints. Estimated effects on the 6 other measures were statistically insignificant (at the 10 percent level). While many types of fraud are common, they do not all cluster statistically. These correlations are also hard to interpret, given that FEFA observers who encounter difficulties may be less able to report on intimidation or complaints.

Turning to the full experimental sample for which FEFA measurement is available (393 polling stations), we do not find any clear sizeable effects of treatment on 10 of these additional measures (Appendix Table 10). The exception is complaints reported by FEFA, which actually decline, but are difficult to interpret. Again, this may be, in part, because many of these activities could been taken and recorded before letters were delivered to PCMs.

Taken together, Appendix Table 9 reports on many varieties of electoral fraud that were of concern to FEFA and the IEC, which could have been observed by survey respondents and plausibly affected attitudes. Yet Appendix Table 10 fails to find statistically convincing treatment effects on any of them.

To conclude, the primary mechanism linking treatment to improved perceptions of government appears to be through PCMs properly posting tallies. That mechanism is consistent with our intuition and with that of our implementing partners. However, as we do not observe all dimensions of management/types of fraud in these data, it is certainly possible that polling center managers took other actions in response to treatment that were not recorded by FEFA or the IEC, but did affect attitudes.

5. Conclusion

Reducing electoral fraud causally improves attitudes toward government in general, and attitudes toward compliance with government authority in particular. Both suggest that fraud reduction enhances legitimacy. These findings are new to the literature and are potentially compelling given the setting: even in an extremely fragile context, with a raging insurgency and an ineffective government rife with corruption, enhancing electoral fairness seems to contribute to state legitimacy in Afghanistan.

These findings speak both to policy and to the study of legitimacy in nascent democracies. From a policy perspective, our results reinforce the notion that domestic attitudes toward government, and therefore presumably government capacity and stability, can be enhanced by reducing fraud in elections. That notion undergirds an emphasis the international community currently places on holding elections in fragile states and the considerable investments it makes to ensure electoral integrity.

Our results cannot provide guidance on *how fair* elections must be in order to legitimize a government, when compared to the counterfactual of no elections (Höglund et al 2009). Electoral processes in these contexts frequently suffer fraud (Bjornlund 2004; Hyde 2011; Kelley 2011), can incite violence (Horowitz 1985; Hyde and Marinov 2012; Snyder 2000; Wilkinson 2004), and may institutionalize former combatants into uncompromising political parties. In such circumstances, staging unfair elections in an attempt to increase state legitimacy may instead undermine it. In the context of a decision on when to hold elections for which electoral fairness is a consideration, our results contribute two insights: fraud reduction is both possible and legitimacy-enhancing.

So post-conflict elections need not be ruled out merely on the grounds that fraud is inevitable. Instead, fraud reduction might be seen as one "check and balance" on political authority, which complements other building blocks of democratic governance in fragile states (Commission on State Fragility, Growth and Development, 2018).

Enhancing policing, justice, health, education, security, or other basic services should also increase legitimacy, as would large infrastructure projects, according to theories of outcome legitimacy. Donors have spent billions of dollars on a variety of "democracy promotion" programs in Afghanistan, including massive technical and financial assistance to support elections. These include sponsorship of international election observers to monitor polling stations, and support to the Independence Election Commission (IEC) to improve its administrative functioning. Excluding election-specific security costs, international donors typically spend between 200-300 million USD per election round (Condra et al., 2018). Compared to those other governance-enhancing interventions in fragile states, electoral fraud reduction has not only proven to be effective, but is also cost-effective. We successfully visited 471 polling centers, with a budget of just over US$100,000. Relative to those interventions, fraud reduction in elections is a remarkably low cost approach.[24]

Legitimacy plays a key role in theories of political development. It is also relevant for understanding economic development: the government's ability to impose rules is a precondition for taxation, service provision, protecting human rights, enforcing property rights, correcting market failures, and implementing development programs. Assuming that this authority can be expressed without cost

[24] Our fraud-reduction intervention has been successfully replicated in two subsequent elections. Callen, Gibson, Jung, and Long, 2016 report results from replication in Uganda.

is unrealistic in a fragile state. Measuring attitudes regarding compliance with government authority, and exploring interventions that improve those attitudes is a first step toward a more realistic approach.

Why are attitudes affected by fraud reduction? We can only speculate. It may that procedural fairness affects attitudes directly, or it may induce an expectation of more responsive governance, or it may signal improved governance in other dimensions ---*outcome* legitimacy. Our evidence cannot adjudicate between those possibilities. Future experiments which enhance election integrity might attempt to do so.

References

Akerlof, G., Yellen, J.L., 1994. "Gang Behavior, Law Enforcement, and Community Values." in *Values and Public Policy*. Henry J. Aaron, Thomas E. Mann, and Timothy Taylor eds. Washington, D.C.: The Brookings Institution.

Alesina, A., Giuliano, P., Nunn, N., 2013. On the origins of gender roles: Women and the plough. The Quarterly Journal of Economics 128 (2), 469-530.

Anderson, M. L., 2008. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedaian, Perry Preschool, and Early Training Projects. Journal of the American Statistical Association 103 (484), 1481–1495.

Asunka, J., Brierley, S., Golden, M., Kramon, E., Ofosu, G., 2013. Protecting the Polls: The Effect of Observers on Election Fraud. Working Paper.

Bayley, D. H., 1994. Police for the Future. New York: Oxford University Press.

Beath, A., Christia, F., Enikolopov, R., 2012. Winning hearts and minds through development: evidence from a field experiment in Afghanistan. The World Bank.

Benjamini, Y., Krieger, A., Yekutieli, D., 2006. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93, 491–507.

Berman, E., Shapiro, J.N., Felter, J.H., 2011. Can hearts and minds be bought? the economics of counterinsurgency in Iraq. Journal of Political Economy 199 (4), 766-819.

Berman, E., Felter, J.H., Shapiro, J.N., 2018. Small Wars, Big Data: The Information Revolution in Modern Conflict. Princeton, NJ: Princeton University Press.

Bernstein, T., Lü, X., 2003. Taxation Without Representation in Rural China. Cambridge, UK: Cambridge University Press.

Bjornlund, E., 2004. Beyond Free and Fair: Monitoring Elections and Building Democracy. Washington, D.C.: Woodrow Wilson Center Press.

Brancati, D., Snyder, J.L., 2011. Rushing to the polls: the causes of premature postconflict elections. Journal of Conflict Resolution 55 (3), 469-92.

Bruhn, M., McKenzie, D., 2009. In pursuit of balance: randomization in practice in development field experiments. American Economic Journal: Applied Economics 1 (4), 200-232.

Burde, D., Middleton, J., Samii, C., 2016. The effect of social service provision on perceptions of government legitimacy: evidence from an impact evaluation of community based schooling in Afghanistan." Baseline Report.

Bursztyn, L., Callen, M., Ferman, B., Gulzar, S., Hasanain, A., Yuchtman, N., 2014. Identifying ideology: experimental evidence on anti-Americanism in Pakistan. National Bureau of Economic Research, No. w20153.

Callen, M., Long, J.D., 2015. Institutional corruption and election fraud: evidence from a field experiment in Afghanistan. American Economic Review 105 (1), 354-381.

Callen, M., Weidmann, N.B., 2013. Violence and election fraud: evidence from Afghanistan. British Journal of Political Science 43 (1), 53-75.

Callen, M., Gibson, C., Jung, D., Long, J.D., 2016. Improving electoral integrity with information and communications technology. Journal of Experimental Political Science 3 (1), 4 – 17.

Casey, K., Glennerster, R., Miguel, E., 2012. Reshaping institutions: evidence on aid impacts using a pre-analysis plan. Quarterly Journal of Economics 127 (4), 1755-1812.

Cassidy, R., 2008. Counterinsurgency and the Global War on Terror. Westport, CT: Praeger.

Commission on State Fragility, Growth and Development. 2018. Escaping the fragility trap. https://www.theigc.org/wp-content/uploads/2017/06/Escaping-the-fragility-trap_Aug-2018.pdf .

Condra, L. N., Callen, M., Iyengar, R. K., Long, J. D., Shapiro, J. N., 2019. Damaging democracy? security provision and turnout in Afghan elections. Economics & Politics 31 (2), 163-193.

Condra, L. N., Long, J. D., Shaver, A. C., Wright, A. L., 2018. The logic of insurgent electoral violence. American Economic Review 108 (11), 3199-3231.

Dal Bo, P., Foster, A., Putterman, L., 2010. Institutions and behavior: experimental evidence on the effects of democracy. American Economic Review 100 (5), 2205-2229.

de Zeeuw, J., ed. 2008. From Soldiers to Politicians: Transforming Rebel Movements After Civil War. Boulder: Lynne Rienner Publishers.

Desmet, K., Ortuño-Ortín, I., Wacziarg, R., 2017. Culture, ethnicity, and diversity. American Economic Review 107 (9), 2479-2513.

Enikolopov, R., Korovkin, V., Petrova, M., Sonin, K., Zakharov, A., 2013. Election fraud in Russian parliamentary elections in December 2011: evidence from a field experiment. Proceedings of the National Academy of Sciences 110 (2), 448-452.

Fearon, J. D., Humphreys, M., Weinstein, J.M., 2009. Can development aid contribute to social cohesion after civil war? evidence from a field experiment in post-conflict Liberia. American Economic Review 99 (2), 287-291.

Fearon, J. D., Humphreys, M., Weinstein, J.M., 2011. Democratic institutions and collective action capacity: results from a field experiment in post-conflict Liberia. Working Paper.

Ferree, K., Gibson, C., Jung, D., Long, J.D., McIntosh, C., 2018. How technology shapes the crowd: participation in the 2014 South African election. Center for Effective Global Action, Working Paper.

Fjeldstad, O., Semboja, J., 2000. Dilemmas of fiscal decentralisation: a study of local government taxation in Tanzania. Forum for Development Studies 27 (1), 7-41.

Fortin, N.M., 2005. Gender role attitudes and the labour-market outcomes of women across OECD countries. Oxford Review of Economic Policy 21 (3), 416-438.

Greif, A., Tadelis, S., 2010. A theory of moral persistence: crypto-morality and political legitimacy. Journal of Comparative Economics 38 (3), 229-244.

Grimes, M., 2006. Organizing willingness: the role of procedural fairness in political trust and compliance. European Journal of Political Research 45 (2), 285-315.

Grossman, G., Baldassarri, D., 2012. The impact of elections on cooperation: evidence from a lab-in-the-field experiment in Uganda. American Journal of Political Science 56 (4), 964-985.

Guyer, J. 1992. Representation without taxation: an essay on democracy in rural Nigeria, 1952-1990. African Studies Review 35 (1), 41-79.

Humphreys, M., Weinstein, J. M., 2012. Policing politicians: citizen empowerment and political accountability in Uganda preliminary analysis. Working Paper.

Hyde, S. D., 2007. The observer effect in international politics: evidence from a natural experiment. World Politics 60 (1), 37-63.

Hyde, S. D., 2009. Experimenting in democracy promotion: international observers and the 2004 presidential elections in Indonesia. Perspectives on Politics 8 (2), 511-527.

Hyde, S. D., 2011. The Pseudo-Democrat's Dilemma: Why Election Observation Became an International Norm. Cornell University Press.

Hyde, S. D., Marinov, N., 2012. Which elections can be lost? Political Analysis 20 (2), 191-210.

Jung, D., Long, J.D., 2018. Social sanctioning, trust, and voter turnout in emerging democracies. Working Paper.

Kelley, J., 2012. Monitoring Democracy: When International Election Observation Works, and Why it Often Fails. Princeton University Press. 1527-1556.

Kennedy, D. M., Braga, A. A., Piehl, A. M., 2001. Developing and implementing Operation Ceasefire," in Reducing Gun Violence: the Boston Gun Project's Operation Ceasefire, U.S. Department of Justice Research Report, NCJ 188741.

Kish, L., 1949. A Procedure for Objective Respondent Selection within the Household. Journal of the American Statistical Association 44 (247), 380–87.

Kling, J, R., Liebman, J. B., Katz, L. F., 2007. Experimental analysis of neighborhood effects. Econometrica 75 (1), 83–119.
Lake, D. A., 2009. Hierarchy in International Relations. Cornell University Press.
Levi, M., 2006. Why we need a new theory of government. Perspectives on Politics 4 (1), 5-19.
Levi, M., Sacks, A., Tyler, T. 2009. Conceptualizing legitimacy, measuring legitimating beliefs. American Behavioral Scientist 53 (3): 354-75.
Luttmer, E. F. P., Singhal, M., 2014. Tax morale. The Journal of Economics Perspectives 28 (4), 149-168.
Mansfield, E. D., Snyder, J. L., 2007. The sequencing "fallacy". Journal of Democracy 18 (3), 5-10.
Mao, 1937. On Guerrilla Warfare. Trans by Samuel Griffin. 2d edition. Chicago: Univ. Illinois.
McChrystal, S., 2009. Commanders Initial Assessment–30 August. *washingtonpost.com/wp-rv/politics/documents/Assessment_Redacted_092109.pdf*.
O'Brien, K. J. 2002. Collective action in the Chinese countryside. China Journal, 48, 139-154.
Paris, R., 2004. At War's End: Building Peace After Civil Conflict. New York: Cambridge University Press.
Paternoster, R., Brame, R., Bachman, R., Sherman, L. W., 1997. Do fair procedures matter? the effect of procedural justice on spouse assault. Law and Society Review 31 (1), 163-204.
Petraeus, D., 2006. Learning counterinsurgency: observations from soldiering in Iraq. Military Review (January–February): 45–55.
Sepp, K., 2005. Best practices in counterinsurgency. Military Review (May–June): 8–12.
Snyder, Jack L. 2000. From Voting to Violence: Democratization and Nationalist Conflict. New York: Norton.
Sunshine, J., Tyler, T.R., 2003. The role of procedural justice and legitimacy in shaping public support for policing. Law & Society Review 37 (3), 513-548.
Tyler, T. R., 2006. Why People Obey the Law. Princeton: Princeton University Press.
Tyler, T. R., Huo, Y. J., 2002. Trust in the Law: Encouraging Public Cooperation with the Police and Courts. Russell Sage Foundation.
Westfall, P. H., Young, S., 1993. Resampling-based Multiple Testing: Examples and Methods for P-Value Adjustment. Vol 279, John Wiley & Sons.
Wilkinson, Steven I. 2004. Votes and Violence: Electoral Competition and Communal Riots in India. Cambridge University Press.

Table 1: Effect of Treatment on Fraud - Three Measures

Dependent Variable:	Election Tally Removed (=1)			Votes (total)			Enough Votes to Win Station (=1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Letter Delivered (=1)	-0.110***	-0.109***	-0.111***	-0.039	0.008	0.026	0.003	0.003	0.004
(Treatment)	(0.032)	(0.031)	(0.032)	(0.192)	(0.046)	(0.048)	(0.004)	(0.002)	(0.002)
Provincial Aggregator Connection (=1)				23.318***	20.624***	20.622***	0.415***	0.408***	0.408***
				(2.680)	(2.491)	(2.492)	(0.027)	(0.027)	(0.027)
Treatment x Provincial Aggregator Connection				-6.919**	-6.887**	-6.883**	-0.112***	-0.114***	-0.114***
				(3.306)	(3.044)	(3.046)	(0.037)	(0.036)	(0.036)
Mean of DV in controls	0.191	0.191	0.191	1.417	1.417	1.417	0.085	0.085	0.085
R-squared	0.026	0.218	0.241	0.036	0.095	0.095	0.008	0.019	0.019
Stratum FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Additional Covariates	No	No	Yes	No	No	Yes	No	No	Yes
# Observations	459	459	459	375457	375457	375457	375457	375457	375457
# Clusters				451	451	451	451	451	451

Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

Notes: The level of analysis corresponds to the level at which we observe the dependent variable. Columns (1) - (3) report OLS specifications estimated at the polling center level. Columns (4) - (9) are estimated at the candidate - polling station level. Correspondingly, robust (White) standard errors are reported in parentheses for columns (1) - (3) (not clustered since data are already aggregated to the polling center level) and robust standard errors are clustered at the polling center level in columns (4) - (9). The "additional covariates" are the number of military events within 1KM of the polling center, whether the polling center was visited by international monitors, and the average response within the polling center catchment from our baseline survey fielded in August 2010 to whether the respondent is employed, years of education, general happiness (1-10), gender, marital status, and age. For descriptive statistics see Table 1 of (Callen and Long, 2015).

Table 2: Post-Election Summary Statistics

	Mean	Standard Dev.	Observations
Demographics (Survey):			
Employed (=1)	0.524	0.500	2403
Age (years)	32.500	12.221	2403
Female (=1)	0.469	0.499	2403
Married (=1)	0.690	0.463	2403
Education (years)	7.090	5.412	2403
General Happiness (1-10)	4.450	1.694	2403
Attitudes (Survey):			
MP Provides Services (=1)	0.196	0.397	2403
Afghanistan is a Democracy (=1)	0.674	0.469	2403
Voting Improves Future (=1)	0.610	0.488	2403
Gov. Exclt. or Good Job of Prov. Serv. (=1)	0.456	0.498	2403
Important to Report IED to ANSF (=1)	0.934	0.248	2403
Police Should Resolve Disp (=1)	0.183	0.387	2403
Courts Should Resolve Disputes (=1)	0.082	0.274	2403
Paying Taxes Somewhat. or Very Imp't (=1)	0.836	0.370	2403
Trust Afg. Gov. to Determine Guilt (=1)	0.529	0.499	2403
Elections and Violence:			
Military Events within 1KM	2.542	7.335	459
Visited by Int'l Monitor (=1)	0.163	0.369	459
Aware of Treatment (=1)	0.069	0.146	447
Election Tally Removed (=1)	0.135	0.342	459
Votes (total)	1.391	8.436	375507
Enough Votes to Win Station (=1)	0.087	0.281	375507
Votes for Candidate Connected to Provincial Aggregator	24.276	49.375	1846
Enough Votes to Win Station (Connected to Aggregator)	0.447	0.497	1846

Notes: Military event data are from International Security Assistance Force (ISAF) Combined Information Data Network Exchange (CIDNE) database. Data on international monitor visits are provided by Democracy International. Vote counts are from a web scrape performed on October 24, 2010 of the Independent Election Commission of Afghanistan website. Remaining data are from our post-election survey fielded in December 2010. The survey sample is restricted to the respondents who provide some response to the questions corresponding to all attitude variables. MP is a member of the national parliament. An IED is an improvised explosive device, generally a roadside bomb. ANSF are the Afghan National Security Forces, including police and military.

Table 3: Randomization Verification at Baseline

	No Letter	Letter	*Difference*	P-value	# Control	# Treatment
Demographics (Survey):						
Employed (=1)	0.573	0.557	-0.017	0.379	1198	1194
	(0.014)	(0.013)	(0.019)			
Age (years)	33.303	33.560	0.257	0.616	1198	1194
	(0.356)	(0.368)	(0.512)			
Female (=1)	0.477	0.483	0.006	0.777	1198	1194
	(0.014)	(0.014)	(0.020)			
Married (=1)	0.708	0.705	-0.003	0.897	1198	1194
	(0.015)	(0.014)	(0.021)			
Education (years)	6.703	6.814	0.111	0.689	1198	1194
	(0.201)	(0.192)	(0.278)			
General Happiness (1-10)	4.992	4.956	-0.035	0.773	1198	1194
	(0.086)	(0.086)	(0.122)			
Attitudes (Survey):						
MP Provides Services (=1)	0.164	0.151	-0.014	0.501	1198	1194
	(0.015)	(0.013)	(0.020)			
Afghanistan is a Democracy (=1)	0.669	0.652	-0.017	0.499	1198	1194
	(0.019)	(0.017)	(0.025)			
Voting Improves Future (=1)	0.683	0.696	0.013	0.617	1198	1194
	(0.019)	(0.019)	(0.026)			
Gov. Exclt. or Good Job of Prov. Serv. (=1)	0.547	0.579	0.032	0.281	1198	1194
	(0.021)	(0.021)	(0.030)			
Important to Rept IED to ANSF (=1)	0.959	0.972	0.012	0.184	1198	1194
	(0.008)	(0.005)	(0.009)			
Police Should Resolve Disp (=1)	0.205	0.233	0.027	0.229	1198	1194
	(0.016)	(0.016)	(0.023)			
Courts Should Resolve Disputes (=1)	0.130	0.122	-0.008	0.657	1198	1194
	(0.013)	(0.012)	(0.018)			
Paying Taxes Somewhat or Very Imp't (=1)	0.851	0.859	0.009	0.664	1198	1194
	(0.014)	(0.014)	(0.020)			
Elections and Violence:						
Military Events within 1KM	2.759	2.618	-0.141	0.848	216	225
	(0.609)	(0.416)	(0.738)			
Visited by Int'l Monitor (=1)	0.153	0.186	0.033	0.354	216	225
	(0.025)	(0.026)	(0.036)			

Notes: Standard errors clustered at the polling center level reported in parentheses. Survey data are from the baseline survey fielded in August 2010. Military event data are from International Security Assistance Force (ISAF) Combined Information Data Network Exchange (CIDNE) database. Data on international monitor visits are provided by Democracy International. The survey sample is restricted to the respondents who provide some response to the questions corresponding to all Attitudes variables. MP is a member of the national parliament. An IED is an improvised explosive device, generally a roadside bomb. ANSF are the Afghan National Security Forces, including police and military.

Table 4: Effect of Treatment on Measures of Legitimacy---Primary Indices

Dependent Variable:	*Perceptions of Government Index*			*Compliant Attitudes Index*			*All Outcomes Index*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Letter Delivered (=1)	0.054*	0.059**	0.057**	0.068***	0.062***	0.064***	0.062***	0.061***	0.061***
	(0.031)	(0.025)	(0.025)	(0.024)	(0.020)	(0.021)	(0.020)	(0.017)	(0.017)
Mean of DV in controls	0.018	0.018	0.018	0.002	0.002	0.002	0.009	0.009	0.009
R-squared	0.002	0.125	0.152	0.006	0.099	0.119	0.007	0.090	0.118
Stratum FEs	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Additional Covariates	No	No	Yes	No	No	Yes	No	No	Yes
# Observations	2403	2403	2403	2403	2403	2403	2403	2403	2403
# Clusters	459	459	459	459	459	459	459	459	459

Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Notes: Standard errors clustered at the polling center level in parentheses. "Perceptions of Government Index" is a z-score index of four dummy variables: MP Provides Services, Afghanistan is a Democracy, Voting Improves Future, and Gov. Ext. or Good Job of Prov. Serv. "Compliant Attitude Index" is a z-score index of five dummy variables: Impt to Rept IED to ANSF, Police Should Resolve Disp, Courts Should Resolve Disputes, Paying Taxes is Some. or Very Imp't, and Trust Afg. Gov. to Determine Guilt. "All Outcome Index" is a z-score index of all nine of these variables. The "additional covariates" are the number of military events within 1KM of the polling center, whether the polling center was visited by international monitors, and the average response within the polling center catchment from our baseline survey fielded in August 2010 to whether the respondent is employed, years of education, general happiness (1-10), gender, marital status, and age. See also Table 2. The survey sample is restricted to the respondents who provide some response to the questions corresponding to all nine variables.

Table 5: Standardized Treatment Effects for All Variables Measuring Legitimacy

	Control Mean	Treatment Effect	Naïve P-Value	Adjusted P-value
Perceptions of Government Index	0.015	0.059**	0.019	0.024
	(0.019)	(0.025)		
MP Provides Services (=1)	0.000	0.120**	0.010	0.043
	(0.031)	(0.047)		
Afghanistan is a Democracy (=1)	0.025	0.047	0.283	0.396
	(0.033)	(0.044)		
Voting Improves Future (=1)	0.006	0.009	0.822	0.608
	(0.029)	(0.041)		
Gov. Ext. or Good Job of Prov. Serv. (=1)	0.030	0.059	0.222	0.396
	(0.035)	(0.049)		
Compliant Attitudes Index	0.004	0.062***	0.002	0.009
	(0.015)	(0.020)		
Important to Rept IED to ANSF (=1)	0.020	0.08**	0.040	0.110
	(0.030)	(0.039)		
Police Should Resolve Disp (=1)	0.018	0.048	0.306	0.299
	(0.032)	(0.047)		
Courts Should Resolve Disputes (=1)	-0.035	0.014	0.693	0.403
	(0.025)	(0.036)		
Paying Taxes is Somewhat or Very Imp't (=1)	-0.004	0.103**	0.027	0.110
	(0.035)	(0.046)		
Trust Afg. Gov. to Determine Guilt (=1)	0.022	0.066	0.172	0.209
	(0.035)	(0.049)		
All Outcomes Index	0.009	0.061***	0.000	
	(0.013)	(0.017)		

Significance levels (naive p-value) indicated by *p < .10, **p < .05, ***p < .01.

Notes: Standard errors clustered at polling center level reported in parentheses. Treatment effects are standardized regression coefficients from a regression of the dependent variable, normalized by subtracting the mean and dividing by the standard deviation, on an indicator for treatment and stratum fixed effects. Indices take an average of all of the variables listed within the given hypothesis group, or across all nine variables in the case of the All Outcomes Index. P-values are corrected for multiple hypothesis testing as follows---we adjust across the two primary H1 and H2 indices to control the familywise error rate (FWER) computed following Westfall and Young (1993) and Anderson (2008); within each hypothesis group, we adjust to control the false discovery rate (FDR) computed following Benjamini, Krieger and Yekutieli (2006) and Anderson (2008). The survey sample is restricted to the respondents who provide some response to the questions corresponding to all nine variables.

Figure 1: Experimental Sample in Afghanistan

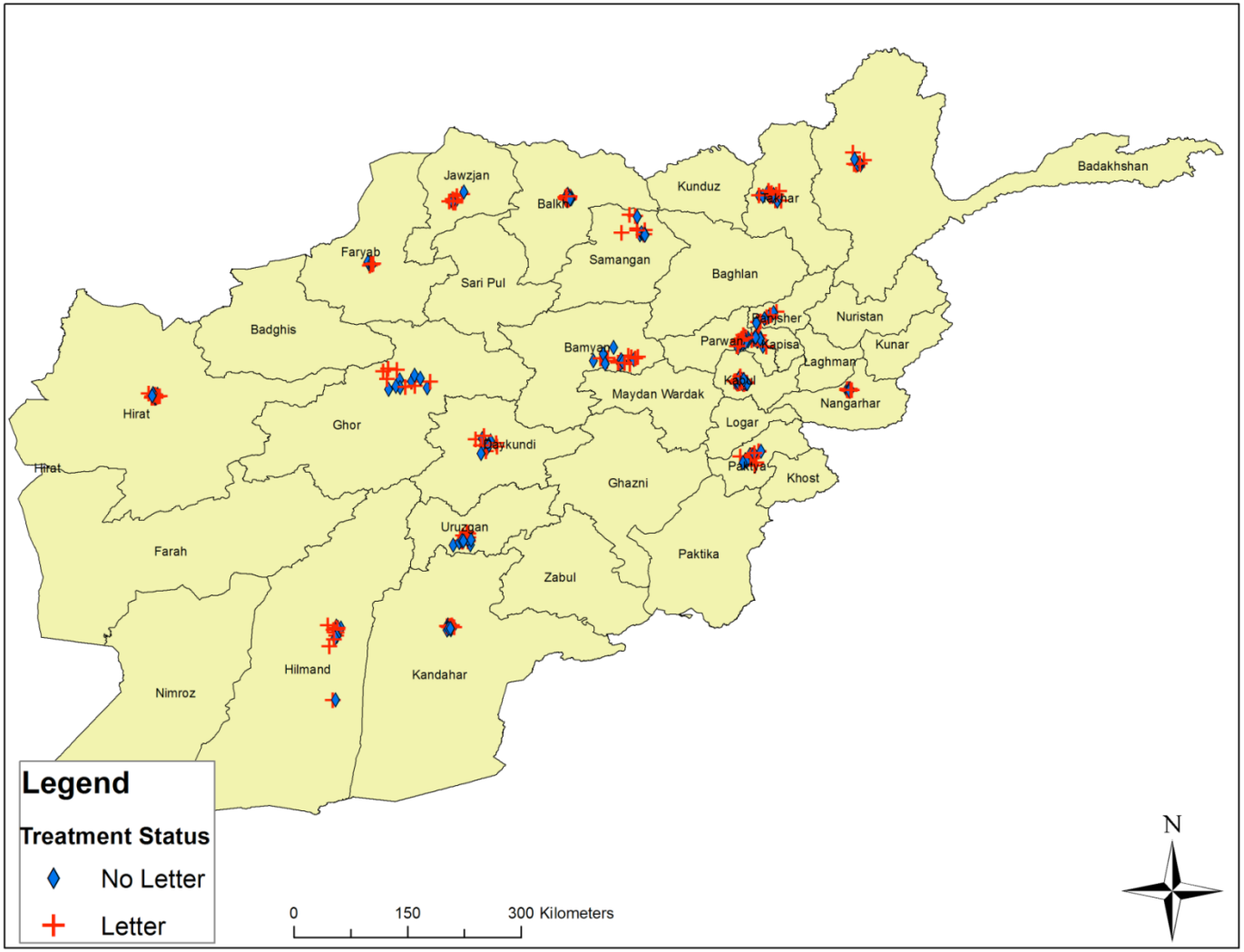

Figure 2: Announcement of Monitoring

Polling Center Name: ……..

Polling Center Code:…….

Date: ……

Dear Sir or Madam-

Greetings! I am an official election observer with the Opinion Research Center of Afghanistan (ORCA). My organization is providing this letter to collect some important information about your polling center and share it with our main office. Your polling center has been randomly selected from among polling centers in this province.

In our attempts to help Afghanistan have free and fair elections, I will return to this polling center tomorrow morning in order to take pictures of the results for every candidate in every station on the tally sheets after they have been posted.

The information will be posted on a website that belongs to local and international election observers so that it will be used by the people of Afghanistan, the international community, and local and international media. We will also compare the photos taken with the tally certified by the IEC in Kabul.

As recognition that you have read and understood this letter, please sign here: _____

Thank you kindly for your help and cooperation.

Sincerely,

Haj Abdul Nabi Barakzai

Deputy Head of ORCA

Name and Signature of manager of polling station:……

Figure 3: Announcement of Monitoring (Dari)

نام مرکز رای دهی: _____ تاریخ: _____
_____ مرکز رای دهی: _____کود

به حضور محترم آقای / خانم

مسئولیت نظارت 472 مراکز رای بر حسب توافقنامه کمیسیون مستقل انتخابات دفتر اورکا
دهی را بر عهده دارد.

مسئولیت و برای او (ORCA) دفتر به مربوطیک تن از نظارت کنندگان رسمیدارنده مکتوب
معلومات تا بتواند مرکز رای دهی تسلیم نمودهراین تا این مکتوب را وظیفه سپرده شده است
. این مرکز دفتر مرکزی شریک بسازدجمع آوری نموده و با مرکز رای دهی این و دقیق را ازموثق
این ولایت تمام مراکز رای دهی میانبه صورت تصادفی از گر به شمول چندین مراکز دیرای دهی
انتخاب شده است.

فردا صبح . نظر ما یک انتخابات آزاد و مشروع در افغانستان کمک خواهیم کرد تقویتبرایما
. نصب میگردد اخذ نمایندهمرکز رای دهیاین که در را نتایج کاندیدان لست آمد تا تصاویر از دخواه

گذاشته مربوط به نظرین انتخاباتی داخلی و خارجی این نتایج در سایت انترنتی تصاویر
از این نتایج ، موسسات خارجی، و مطبوعات داخلی و خارجیخواهد شد تا تمام مردم افغانستان
نتایج را با نتایج که از طرف ایننتصاویر حاصله از نظرحیث منلستفاده کنند. و همچنان ما
انتخابات در کابل نشر میشود مقایسه خواهیم کرد. همتقلکمیسیون

در پایین ایدبرای تایید اینکه این مکتوب بدستترس شما قرار گرفته و شما انرا مطالعه نموده
مضا نمایید. لطف نموده ا

از همکاری شما قبلا اظهار سپاس.

بااحترام

حاجی عبدالنبی بارکزی

معاون دفتر اورکا

آمر محترم مرکزرای دهی: _____ یامضااسم و

Appendix Table 1: Ensuring There is No Differential Attrition into Consistent Sample

Dependent Variable:	In Consistent Sample (=1)		
	(4)	(5)	(6)
Letter Delivered (=1)	-0.002	0.003	0.001
	(0.022)	(0.016)	(0.016)
Mean of DV in controls	0.800	0.800	0.800
R-squared	0.000	0.159	0.199
Stratum FEs	No	Yes	Yes
Additional Covariates	No	No	Yes
# Observations	3010	3010	3009
# Clusters	462	462	462

Notes: Standard errors clustered at the polling center level are reported in parentheses. Data is from our post-election survey fielded in December 2010. "In Consistent Sample" is equal to one for respondents who provide some response to the questions corresponding to all attitudes variables reported in Table 2. The "additional covariates" are the number of military events within 1KM of the polling center, whether the polling center was visited by international monitors, and the average response within the polling center catchment from our baseline survey fielded in August 2010 to whether the respondent is employed, years of education, general happiness (1-10), gender, marital status, and age.

Appendix Figure 1: Voters viewing results on the polling center's tally form

Appendix Table 2: Post-Election Summary Statistics for Unrestricted Sample

	Mean	Standard Dev.	Observations
Demographics:			
Employed (=1)	0.492	0.500	3010
Age (years)	32.654	12.367	3009
Female (=1)	0.500	0.500	3010
Married (=1)	0.696	0.460	3010
Education (years)	6.593	5.470	3009
General Happiness (1-10)	4.382	1.724	3010
Attitudes:			
MP Provides Services (=1)	0.187	0.390	2965
Afghanistan is a Democracy (=1)	0.666	0.472	2706
Voting Improves Future (=1)	0.600	0.490	2763
Gov. Ext. or Good Job of Prov. Serv. (=1)	0.434	0.496	2900
Impt to Rept IED to ANSF (=1)	0.925	0.263	2930
Police Should Resolve Disp (=1)	0.173	0.378	2994
Courts Should Resolve Disputes (=1)	0.091	0.288	2994
Paying Taxes is Some. or Very Imp't (=1)	0.831	0.375	3010
Trust Afg. Gov. to Determine Guilt (=1)	0.514	0.500	2907
Elections and Violence:			
Military Events within 1KM	2.619	7.517	462
Visited by Int'l Monitor (=1)	0.162	0.368	462
Aware of Treatment (=1)	0.066	0.135	460
Election Tally Removed (=1)	0.134	0.341	462
Votes	1.402	8.445	376893
Enough Votes to Win Station (=1)	0.087	0.282	376893
Votes for Candidate Connected to Provincial Aggregator	24.230	49.331	1850
Enough Votes to Win Station (Connected to Aggregator)	0.446	0.497	1850

Notes: Military event data are from International Security Assistance Force (ISAF) Combined Information Data Network Exchange (CIDNE) database. Data on international monitor visits are provided by Democracy International. Vote counts are from a web scrape performed on October 24, 2010 of the Independent Election Commission of Afghanistan website. Remaining data are from our post-election survey fielded in December 2010. The survey sample is restricted to the respondents who provide some response to the questions corresponding to all Attitudes variables. MP is a member of the national parliament. An IED is an improvised explosive device, generally a roadside bomb. ANSF are the Afghan National Security Forces, including police and military.

Appendix Table 3: Baseline Randomization Verification for Unrestricted Sample

	No Letter	Letter	Difference	P-value	# Control	# Treatment
Demographics:						
Employed (=1)	0.566	0.556	-0.01	0.575	1410	1456
	(0.012)	(0.012)	(0.017)			
Age (years)	33.291	33.577	0.285	0.547	1410	1456
	(0.335)	(0.336)	(0.474)			
Female (=1)	0.5	0.5	0	1.000	1410	1456
	(0.013)	(0.013)	(0.019)			
Married (=1)	0.706	0.71	0.004	0.815	1410	1456
	(0.014)	(0.013)	(0.019)			
Education (years)	6.462	6.565	0.103	0.699	1410	1456
	(0.193)	(0.182)	(0.266)			
General Happiness (1-10)	4.949	4.913	-0.035	0.768	1410	1456
	(0.084)	(0.086)	(0.120)			
Attitudes:						
MP Provides Services (=1)	0.163	0.142	-0.021	0.259	1396	1440
	(0.014)	(0.012)	(0.019)			
Afghanistan is a Democracy (=1)	0.655	0.643	-0.011	0.654	1286	1307
	(0.019)	(0.017)	(0.025)			
Voting Improves Future (=1)	0.68	0.69	0.01	0.687	1339	1367
	(0.018)	(0.018)	(0.025)			
Gov. Ext. or Good Job of Prov. Serv. (=1)	0.54	0.563	0.024	0.406	1384	1413
	(0.021)	(0.020)	(0.028)			
Impt to Rept IED to ANSF (=1)	0.956	0.961	0.005	0.592	1390	1418
	(0.007)	(0.006)	0.01			
Police Should Resolve Disp (=1)	0.202	0.217	0.015	0.480	1410	1456
	(0.015)	(0.015)	(0.021)			
Courts Should Resolve Disputes (=1)	0.14	0.133	-0.008	0.654	1410	1456
	(0.013)	(0.012)	(0.018)			
Paying Taxes is Some. or Very Imp't (=1)	0.826	0.836	0.01	0.611	1410	1456
	(0.015)	(0.014)	0.02			
Elections and Violence:						
Military Events within 1KM	2.747	2.617	-0.13	0.860	217	227
	(0.606)	(0.413)	(0.733)			
Visited by Int'l Monitor (=1)	0.152	0.184	0.032	0.365	217	227
	(0.024)	(0.026)	(0.035)			

Notes: Standard errors clustered at the polling center level reported in parentheses. Survey data are from the baseline survey fielded in August 2010. Military event data are from International Security Assistance Force (ISAF) Combined Information Data Network Exchange (CIDNE) database. Data on international monitor visits are provided by Democracy International. MP is a member of the national parliament. An IED is an improvised explosive device, generally a roadside bomb. ANSF are the Afghan National Security Forces, including police and military.

Appendix Table 4: Effect of Treatment on Measures of Legitimacy---Primary Indices, Unrestricted Sample

Dependent Variable:	*Perceptions of Government Index*			*Compliant Attitudes Index*			*All Outcome Index*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Letter Delivered (=1)	0.049	0.056**	0.053**	0.046**	0.045**	0.048**	0.062***	0.061***	0.061***
	(0.030)	(0.024)	(0.024)	(0.023)	(0.019)	(0.019)	(0.020)	(0.017)	(0.017)
Mean of DV in controls	0.020	0.020	0.020	0.001	0.001	0.001	0.009	0.009	0.009
R-squared	0.002	0.126	0.156	0.003	0.101	0.125	0.007	0.090	0.118
Stratum FEs	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Additional Covariates	No	No	Yes	No	No	Yes	No	No	Yes
# Observations	2488	2488	2488	2841	2841	2841	2403	2403	2403
# Clusters	459	459	459	462	462	462	459	459	459

Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Notes: Standard errors clustered at the polling center level in parentheses. Perceptions of Government Index is a z-score index of four dummy variables: MP Provides Services, Afghanistan is a Democracy, Voting Improves Future, and Gov. Ext. or Good Job of Prov. Serv. Compliant Attitude Toward Government Index is a z-score index of five dummy variables: Impt to Rept IED to ANSF, Police Should Resolve Disp, Courts Should Resolve Disputes, Paying Taxes is Some. or Very Imp't, and Trust Afg. Gov. to Determine Guilt. All Outcome Index is a z-score index of all nine of these variables. The "additional covariates" are the number of military events within 1KM of the polling center, whether the polling center was visited by international monitors, and the average response within the polling center catchment from our baseline survey fielded in August 2010 to whether the respondent is employed, years of education, general happiness (1-10), gender, marital status, and age. See Table 2 for an explanation of variables.

Appendix Table 5: Non-standardized Treatment Effects for All Variables Measuring Legitimacy

	Mean in Controls	Treatment Effect	Naïve P-Value	Adjusted P-value
Perceptions of Government Index	0.015	0.059**	0.019	0.024
	(0.019)	(0.025)		
MP Provides Services (=1)	0.173	0.046**	0.010	0.043
	(0.012)	(0.018)		
Afghanistan is a Democracy (=1)	0.663	0.023	0.283	0.396
	(0.016)	(0.021)		
Voting Improves Future (=1)	0.608	0.004	0.822	0.608
	(0.014)	(0.020)		
Gov. Ext. or Good Job of Prov. Serv. (=1)	0.441	0.029	0.222	0.396
	(0.018)	(0.024)		
Compliant Attitudes Index	0.004	0.062***	0.002	0.009
	(0.015)	(0.020)		
Impt to Rept IED to ANSF (=1)	0.923	0.022**	0.040	0.110
	(0.008)	(0.011)		
Police Should Resolve Disp (=1)	0.174	0.018	0.306	0.299
	(0.012)	(0.017)		
Courts Should Resolve Disputes (=1)	0.080	0.004	0.693	0.403
	(0.007)	(0.010)		
Paying Taxes is Some. or Very Imp't (=1)	0.816	0.040**	0.027	0.110
	(0.013)	(0.018)		
Trust Afg. Gov. to Determine Guilt (=1)	0.513	0.033	0.172	0.209
	(0.017)	(0.024)		
All Outcomes Index	0.009	0.061***	0.000	
	(0.013)	(0.017)		

Significance levels (naive p-value) indicated by *p < .10, **p < .05, ***p < .01.

Notes: Standard errors clustered at polling center level reported in parentheses. Treatment effects are standardized regression coefficients from a regression of the dependent variable, normalized by subtracting the mean and dividing by the standard deviation, on an indicator for treatment and stratum fixed effects. Indices take an average of all of the variables listed within the given hypothesis group, or across all nine variables in the case of the All Outcomes Index. P-values are corrected for multiple hypothesis testing as follows---we adjust across the two primary H1 and H2 indices to control the familywise error rate (FWER) computed following Westfall and Young (1993) and Anderson (2008); within each hypothesis group, we adjust to control the false discovery rate (FDR) computed following Benjamini, Krieger and Yekutieli (2006) and Anderson (2008).

Appendix Table 6: Standardized Treatment Effects for All Variables Measuring Legitimacy, Unrestricted Sample

	Mean in Controls	Treatment Effect	Naïve P-Value	Adjusted P-value
Perceptions of Government Index	0.016	0.056**	0.023	0.032
	(0.018)	(0.024)		
MP Provides Services (=1)	-0.006	0.083*	0.050	0.252
	(0.029)	(0.042)		
Afghanistan is a Democracy (=1)	0.007	0.049	0.242	0.321
	(0.030)	(0.042)		
Voting Improves Future (=1)	-0.010	0.003	0.936	0.478
	(0.028)	(0.039)		
Gov. Ext. or Good Job of Prov. Serv. (=1)	-0.012	0.058	0.193	0.321
	(0.033)	(0.045)		
Compliant Attitudes Index	0.001	0.045**	0.020	0.032
	(0.014)	(0.019)		
Impt to Rept IED to ANSF (=1)	-0.003	0.062	0.100	0.332
	(0.027)	(0.037)		
Police Should Resolve Disp (=1)	-0.001	0.029	0.490	0.581
	(0.028)	(0.042)		
Courts Should Resolve Disputes (=1)	0.004	0.001	0.988	0.737
	(0.024)	(0.034)		
Paying Taxes is Some. or Very Imp't (=1)	-0.002	0.071*	0.086	0.332
	(0.031)	(0.041)		
Trust Afg. Gov. to Determine Guilt (=1)	-0.001	0.050	0.254	0.342
	(0.031)	(0.044)		
All Outcomes Index	0.009	0.061***	0.000	
	(0.013)	(0.017)		

Significance levels (naive p-value) indicated by *p < .10, **p < .05, ***p < .01.

Notes: Standard errors clustered at polling center level reported in parentheses. Treatment effects are standardized regression coefficients from a regression of the dependent variable, normalized by subtracting the mean and dividing by the standard deviation, on an indicator for treatment and stratum fixed effects. Indices take an average of all of the variables listed within the given hypothesis group, or across all nine variables in the case of the All Outcomes Index. P-values are corrected for multiple hypothesis testing as follows---we adjust across the two primary H1 and H2 indices to control the familywise error rate (FWER) computed following Westfall and Young (1993) and Anderson (2008); within each hypothesis group, we adjust to control the false discovery rate (FDR) computed following Benjamini, Krieger and Yekutieli (2006) and Anderson (2008).

Appendix Table 7: Treatment Effects on Who is Mainly Responsible for Delivering Services

Main Provider Selected:	Central Government	Member of Parliament	Religious or Ethnic Leaders	Provincial Government	Community Driven Council
	(1)	(2)	(3)	(4)	(5)
Letter Delivered (=1)	-0.050	0.120**	0.043	-0.076**	-0.022
	(0.044)	(0.047)	(0.046)	(0.038)	(0.045)
Mean of DV in controls	0.019	0.007	-0.033	0.008	-0.010
R-squared	0.171	0.071	0.067	0.118	0.066
Stratum FEs	Yes	Yes	Yes	Yes	Yes
Additional Covariates	Yes	Yes	Yes	Yes	Yes
# Observations	2403	2403	2403	2403	2403
# Clusters	459	459	459	459	459

Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Notes: Standard errors clustered at the polling center level in parentheses. The "additional covariates" are the number of military events within 1KM of the polling center, whether the polling center was visited by international monitors, and the average response within the polling center catchment from our baseline survey fielded in August 2010 to whether the respondent is employed, years of education, general happiness (1-10), gender, marital status, and age. See Table 2 for an explanation of variables. The survey sample is restricted to the respondents who provide some response to the questions corresponding to all nine legitimacy variables in Table 5.

Appendix Table 8: Impact of Awareness of International Involvement

Dependent Variable:	*Perceptions of Government Index*			*Compliant Attitudes Index*			*All Outcomes Index*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Letter Delivered (=1)	0.061*	0.065**	0.065**	0.084***	0.080***	0.080***	0.074***	0.073***	0.073***
	(0.034)	(0.029)	(0.029)	(0.026)	(0.022)	(0.022)	(0.021)	(0.019)	(0.019)
Treat X Aware of Delivery	-0.004	-0.011	-0.011	-0.027	0.000	0.000	-0.017	-0.005	-0.005
	(0.052)	(0.049)	(0.049)	(0.037)	(0.036)	(0.036)	(0.031)	(0.031)	(0.031)
Mean of DV in controls	0.034	0.034	0.034	0.005	0.005	0.005	0.018	0.018	0.018
R-squared	0.003	0.135	0.135	0.008	0.107	0.107	0.010	0.097	0.097
Stratum FEs	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Additional Covariates	No	No	Yes	No	No	Yes	No	No	Yes
# Observations	2136	2136	2136	2136	2136	2136	2136	2136	2136
# Clusters	447	447	447	447	447	447	447	447	447

Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Notes: Standard errors clustered at the polling center level in parentheses. Perceptions of Government Index is a z-score index of four dummy variables: MP Provides Services, Afghanistan is a Democracy, Voting Improves Future, and Gov. Ext. or Good Job of Prov. Serv. Compliant Attitude Toward Government Index is a z-score index of five dummy variables: Impt to Rept IED to ANSF, Police Should Resolve Disp, Courts Should Resolve Disputes, Paying Taxes is Some. or Very Imp't, and Trust Afg. Gov. to Determine Guilt. All Outcome Index is a z-score index of all nine of these variables. The survey sample is restricted to the respondents who provide some response to the questions corresponding to all nine variables. The "additional covariates" are the number of military events within 1KM of the polling center, whether the polling center was visited by international monitors, and the average response within the polling center catchment from our baseline survey fielded in August 2010 to whether the respondent is employed, years of education, general happiness (1-10), gender, marital status, and age. See Table 2 for an explanation of variables.

Appendix Table 9: Tally Sheet Removal Predicts Other Measures of Electoral Fraud?

	Campaign Materials Within 100M (=1)	Number of Spoiled Ballots	Intimidation Reported (=1)	Official Complaints Reported (=1)	Un-authorized People in Polling Center (=1)	Threats Made During Voting (=1)	Unused and Spoiled Ballots Recorded (=1)	FEFA Allowed to Observe Without Difficulties	Counted Votes Reflected Exactly on the Tally (=1)	Tally Posted Outside at End of Day (=1)	Results List Distributed to Observers (=1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
ERF Removed	0.128*	17.333*	-0.031*	-0.012**	-0.009	-0.002	0.009*	0.037	-0.002	-0.001	-0.023
	(0.071)	(9.711)	(0.018)	(0.006)	(0.035)	(0.015)	(0.005)	(0.070)	(0.007)	(0.007)	(0.018)
Constant	0.270***	7.036***	0.053***	0.012**	0.099***	0.015*	0.991***	0.778***	0.995***	0.994***	0.991***
	(0.028)	(1.727)	(0.013)	(0.006)	(0.019)	(0.008)	(0.005)	(0.031)	(0.004)	(0.004)	(0.005)
Sample	Control	Control	Control	Control	Control	Control	Control	Control	Control	Control	Control
Obs.	207	207	206	207	206	207	207	207	207	207	207
R-squared	0.017	0.042	0.006	0.005	0.000	0.000	0.003	0.001	0.000	0.000	0.015

Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Notes: This table reports on whether the primary measure of election fraud using in this paper, whether elections tally were removed, correlates with additional measures of electoral misconduct collected by election observers working for the Free and Fair Elections Foundation of Afghanistan (FEFA). ERF removed corresponds to the Election Tally being removed (ERF-Elections Return Form =1). The instrument used by observers to collect these data is reproduced in Appendix B. White heteroscedasticity robust standard errors are reported in parentheses.

Appendix Table 10: Does Treatment Affect Other Measures of Election Fraud?

	Campaign Materials Within 100M (=1)	Number of Spoiled Ballots	Intimidation Reported (=1)	Official Complaints Reported (=1)	Unauthorized People in Polling Center (=1)	Threats Made During Voting (=1)	Unused and Spoiled Ballots Recorded (=1)	FEFA Allowed to Observe Without Difficulties	Counted Votes Reflected Exactly on the Tally (=1)	Tally Posted Outside at End of Day (=1)	Results List Distributed to Observers (=1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Letter Delivered (=1)	0.005	-2.612	0.000	-0.009*	0.027	-0.011	-0.010	-0.006	-0.000	-0.000	0.005
	(0.034)	(2.252)	(0.017)	(0.005)	(0.024)	(0.008)	(0.008)	(0.013)	(0.006)	(0.007)	(0.008)
Stratum FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean in Controls	0.292	10.561	0.05	0.01	0.096	0.015	0.992	0.775	0.995	0.994	0.986
Obs.	393	393	393	393	393	393	393	393	393	393	393
R-squared	0.357	0.165	0.288	0.140	0.229	0.155	0.195	0.897	0.267	0.187	0.113

Significance: *** p<0.01, ** p<0.05, * p<0.1.

Notes: This table reports on whether treatment impacted additional measures of electoral misconduct collected by election observers working for the Free and Fair Elections Foundation of Afghanistan (FEFA). The instrument used by observers to collect these data is reproduced in Appendix B. White heteroscedasticity robust standard errors are reported in parentheses.

Appendix B

Free Fair Election Foundation of Afghanistan (FEFA)

Wolesi Jirga 2010

Election Observation Form

	A. Details About Polling Station and Observer									
1	Polling Center Code									
2	Polling Station Name									
3	Province									
5	Observer Name									
6	Observer Mobile Number	0	7	9	9	3	1	0	6	6	4
7	Starting of observation (Time)	H		M		AM		PM			
8	Ending of Observation (Time)	H		M		AM		PM			
9	District Coordinator name									
10	District Coordinator Telephone Number	0	7	9	9	3	2	6	3	6	0
11	Provincial Coordinator Name									
12	Provincial Coordinator Telephone Number	0	7	7	2	3	2	9	7	4	4

13	Where there campaign materials within 100 M of Polling Center (Complete incident Form)	Yes	No
14	Was there any intimidation outside the polling Center (Complete incident Form)	Yes	No
15	Were there any unauthorized people in the polling Center Complete incident Form)	Yes	No
16	Did you observer any prevention while your observation Period? Complete incident Form)	Yes	No

	B. Management assets of polling Center					
17	Polling Station Code	*	*	*	*	*

					Male		Female								
18	Number of Polling Station Staff				Male		Female								
19	Did Polling Center open at Exact time (7:00 am)?												Yes	No	
20	Were all the IEC polling staff Present?												Yes	No	
20	if no, Please specify the person														
21	Was the polling station faced lack with necessary materials?	Journal		Implementation Form		Punch		Color		Complain Form		TEB			
22	Were the ballot boxes empty before polling started?												Yes	No	
23	Where the ballot papers recorded in the journal before polling start?												Yes	No	
24	Where the numbers of the ballot box seals announced loudly and recorded in the journal?												Yes	No	
25	Box lock No	*	*	*	*	*	26	Box Lock No	*	*	*	*	*		
27	Box lock No	*	*	*	*	*	28	Box Lock No	*	*	*	*	*		
29	Were any official complaints registered about the polling process?												Yes	No	
30	If yes who?	Complain Person						Subject of Complain							
31	Presence of observers and representative			Representative of Candidates /parties				Interior				Foreigner			

C. Voting Process (until the end of process)				
32	Do the IEC staffs check the polling cards?			Yes No
33	Did the IEC stamp the ballot before giving it to the voters?			Yes No
34	Were voter's fingers being inked?			Yes No
35	Was it possible the inks to be cleaned form the fingers of the voters?			Yes No
36	Were the voter registration cards being entered in the journal?			Yes No
37	Do they save the speed of voting speed?			Yes No
37	If No, Please fill the incident Form			
38	Does voting took place in absence of person (Voting more than once by same person)	NO	Sometimes =1-10	More then 10
39	Do the voting took place with under age (a child under age of 18)	No	Sometimes =1-10	More then 10

40	Do they prevent a person from voting since he/she Had voting card in his/her hand?	No	Sometimes =1-10	More then 10
41	Do voting took place as group or family	No	Sometimes=1-10	More then 10
42	Was the voting process disturbed because of any problem?	No	Sometimes=1-10	More then 10
43	Did you observe a person to help several voters?	No	Sometimes=1-10	More then 10
44	Did you observe a person to help several voters?	No	Sometimes=1-10	More then 10
45	Were any official complain registered regarding to the voting process?			Yes / No
46	If yes, Please fill the incident Form	Complain Person:	Complain Them:	
47	Did you observe any threat during the voting process?			Yes / No
47	If yes, (Please fill the incident form)			

D. Closing process (in the same Polling Station observed at opening)												
48	Polling station Number	*	*	*	*	*	Closing Time	*	*	*	*	*
49	Were all voters waiting in the cue at 16:00 allowed to vote										Yes	No
50	Did they lock the ballot boxes after the ending of voting process?										Yes	No
51	The presence of observers/ candidate or parties representatives	Domestic	NO	Foreign	NO	Candidates and party agents						NO
52	Were unused and spoiled ballot papers counted and recorded on the reconciliation form?										Yes	No
52	If no, Please fill the incident form											
53	a) The number of distributed at the beginning of process							*	*	*	*	*
54	b) (Minus) the number of unused ballot papers							*	*	*	*	*
55	c) (minus) Number of spoiled ballots							*	*	*	*	*
56	Total number of voters issued with ballots (a-(b+c))							*	*	*	*	*

E. Counting process:			
57	Was the counting process happen at the courting place?	Yes	No
58	Were there unauthorized people in the polling center while counting?	Yes	No
59	Did they allow you to observe process without any difficulties?	Yes	No
59	If No, Please fill the incident Form		
60	Were the seals of the ballots box intact and undamaged?	Yes	No
61	Did the IEC staff read the number of the seals loudly (Enter the number below)	Yes	No

62	Number of seal	*	*	*	*	*	63	Number of seal	*	*	*	*	*
64	Number of seal	*	*	*	*	*	65	Number of seal	*	*	*	*	*

66	Did the counting officer turn the ballots face down to display the valid ballots paper stamp?											Yes	No	
68	Did they put the documents which are without stamp at different Column?											Yes	No	

Unused ballot papers														
69	Run Votes	*	*	*	*	*	70	White votes	*	*	*	*	*
71	Invalid Votes	*	*	*	*	*	72	Total	*	*	*	*	*
72	The number of voter who have received the ballot papers.								*	*	*	*	*	
73	Differences (the difference should be by Zero)								*	*	*	*	*	
74	Were the counted votes reflected exactly in the final result sheet?											Yes	No	
75	If No, Please fill the incident Form													
75	Did they post a copy of the mentioned list outside of the polling center											Yes	No	
77	If No, Please fill the incident Form													
78	Did they distribute the result list to the observer											Yes	No	
79	If No, Please fill the incident Form													

	Name of Observer	
	Observer signature	

THE END