

Portland State University

PDXScholar

Political Science Faculty Publications and
Presentations

Political Science

4-2020

Does It Matter What Observers Say? The Impact of International Election Monitoring on Legitimacy

Lindsay J. Benstead

Portland State University, benstead@pdx.edu

Kristen Kao

University of Gothenburg

Ellen Lust

University of Gothenburg

Follow this and additional works at: https://pdxscholar.library.pdx.edu/polisci_fac



Part of the [Political Science Commons](#)

Let us know how access to this document benefits you.

Citation Details

Benstead, L., Kao, K., & Lust, E. (2020). Does it matter what observers say? The impact of international election monitoring on legitimacy. *Mediterranean Politics*, 1-22.

This Post-Print is brought to you for free and open access. It has been accepted for inclusion in Political Science Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Does it Matter What Observers Say?

The Impact of International Election Monitoring on Legitimacy

Lindsay J. Benstead

Portland State University

(benstead@pdx.edu; Address: 506 SW Mill Street, Portland, OR 97201; Phone: 503-725-8278;

ORCID 0000-0003-2459-5857 @lindsaybenstead)

Kristen Kao

University of Gothenburg

(kristenkao@gmail.com; Address: C/O Elisabeth Elfving, Ljunggatan 12 LGH 1102, 41321

Gothenburg, Sweden, Phone: +46 70 938 7294, ORCID 0000-0001-5566-040X)

Ellen Lust

University of Gothenburg

(ellen.lust@gu.se; Address: Nordhemsgatan 45A, 4130 06 Gothenburg, Sweden; Phone: +46 76

618 6534, ORCID 0000-0002-7263-8201)

10,076 words (including abstract, references, and footnotes, but excluding optional 3,423 word online appendix, tables, and figures).

*The authors gratefully acknowledge the Program on Governance and Local Development (GLD), Yale University, the Moulay Hicham Foundation, and the World Bank for funding the 2014 Jordanian Governance and Local Development and Post-Election Survey and the 2015 Tunisian Local Governance Performance Index (LGPI) Survey. We thank Pierre Landry and Dhafer Malouche for collaborating on the Tunisian LGPI, and Sarah Bush, Lauren Prather,

Francesco Cavatorta, F. Michael Wuthrich, Alexander Dukalskis, and attendees at the Yale Program on Governance and Local Development Second Annual Conference, New Haven, 9-10 April 2015; Middle East Studies Association annual meeting, Washington, DC, 22-25 November 2014; Midwest Political Science Association annual meeting, Chicago, IL, 16-19 April 2015; the Project on Middle East Political Science (POMEPS), Sixth annual conference, Washington, DC, 21-22 May 2015; and, International Studies Association annual meeting, Baltimore, MD, 22-25 February 2017 for their comments on earlier versions of this paper. We are particularly grateful to Megan Reif and Sebastian Nickel for excellent assistance with the analysis. All remaining errors are the authors’.

Abstract

Election observation has spread dramatically, becoming a near rite-of-passage for entry into the international community. Scholars and democracy promoters often suggest that electoral observers' (EOs') assessments impact public opinion in a straightforward manner, yet, research on communication caution against these sanguine assumptions. We test the impact of EO statements on public opinion in two very different contexts using survey experiments conducted among 3,361 Jordanians and Tunisians. Our results demonstrate the need for democracy promoters to consider negative consequences when implementing democracy promotion programs, and for scholars to undertake further research regarding the impacts of election monitoring on domestic attitudes.

Keywords: Election monitoring, Political communication, Public opinion, Experiments, Middle East and North Africa

Introduction

International election observation has spread dramatically since the 1990s, spawning an industry and becoming a near rite-of-passage for entry into the global community. In 2004, observers witnessed close to 85% of the elections that took place around the world (Kelley, 2012). There is reason to believe that, as the Carter Center puts it, “Impartial, credible EOs play a key role in shaping perceptions about the quality and legitimacy of electoral processes” (Carter Center, 2017). Actors on all sides—incumbents and oppositions, winners and losers—use observers’ presence and their assessments as ammunition against opponents. Regimes trot out their invitations to observer missions as evidence that elections will be free, fair, and legitimate, and after the polls close, those on all sides use monitors’ statements as evidence either to convince domestic and international communities of the election’s legitimacy, or to draw the results into question (Marzouk, 2014).

But do such statements actually impact perceptions of domestic audiences? When residents of a polity hear that observers viewed their elections in a positive or negative light, does it alter how they themselves view the results of the polls? And, is the impact of election statements similar across individuals, and in different political contexts?

This paper uses the results of survey experiments conducted in Jordan and Tunisia to examine how election monitors’ statements may differentially impact opinions of groups within the country, and also to consider how consistent these impacts are across different political contexts. While the first questions have begun to be addressed in the literature (see Corstange & Marinov, 2012, Bush & Prather, 2017, as well as Robertson, 2017 as examples), the last remains largely untouched, and our cross-national exploration attempts to fill this gap. In our experimental approach, respondents are randomly placed into conditions describing EOs’

assessments and asked to what extent they believe the elections represented the will of the people. In one condition, the question is asked without first reporting an assessment from an international observer group. In a second condition, the question is preceded by a positive evaluation from international observers, and in the third condition, it is preceded by a negative evaluation. This allows us to test the assumption, while often not explicitly expressed, that international assessments should impact citizens' perceptions of the elections uniformly—and namely, that positive statements prompt more positive assessments among all respondents and negative statements lead to more negative assessments.

The results suggest that the impact of statements varies across citizens within a country, depending on individuals' *a priori* attitudes toward the government, and also across countries, depending on respondents' attitudes toward the messenger. These results are in line with insights derived from literature in political psychology and public opinion (e.g., Corstange & Marinov, 2012). For reasons we discuss below, it is thus not surprising that statements have a negative impact in contexts where the international community is viewed as playing an interventionist role in support of a non-democratic government (as in Jordan), and they are less influential where the international community is seen as less engaged in shaping elections and upholding authoritarian incumbents (as in Tunisia). Where the West—and by extension EOs—are viewed negatively, election observation may have unintended, negative consequences.

We proceed as follows. We begin in the next section by reviewing insights from the literature on how election observation impacts citizens. In the third section, we explore differences between Jordan and Tunisia and consider why this should increase our confidence in the generalizability of the findings. The fourth section presents the survey experiment. The final

section concludes with implications of these findings for both the study and practice of election observation, and for international interventions more broadly.

Impact of Election Observation: Election Processes, Election Day and Legitimacy

Scholars and practitioners have increasingly recognized that election observation may influence elections in the lead-up to Election Day, on the day itself, and in the process that follows; yet, they have much less to say about how statements affect public opinion within the monitored country. There is some evidence that EOs contribute to clean elections on Election Day, but their presence may have negative effects on the overall political environment. Looking at national legislative and executive elections from 1975 to 2004 across 182 countries, Kelley (2012) argues that election observation is associated with improved election quality in terms of being considered internationally acceptable, fewer reported electoral irregularities, and more frequent turnovers in power. Where observers are present, they appear to reduce voter registration fraud (Ichino & Schundeln, 2012), suspicious turnout rates (Sjoberg, 2012; Asunka, Brierley, Golden, Kramon & Ofosu, 2013), and the incumbent candidate's vote share (Hyde 2007; Enikolopov, Korovkin, Petrova, Sonin & Zacharov, 2013).

However, election observation is also associated with negative effects on domestic institutions, governance, and freedoms in the observed country (Simpser & Donno, 2012). Election monitoring may encourage strategic manipulation by parties who wish to thwart elections, and an increased probability of opposition party boycotts (Beaulieu & Hyde 2009). Even on Election Day, EOs may simply push fraud to the polling stations where they are not present (Ichino & Schundeln, 2012; Asunka et al., 2013). Yet, because this research is among the relatively few studies done on the effects of electoral observation, with some showing contradictory results, more research is needed.

Practitioners and scholars have largely overlooked questions of how observer statements affect domestic public opinion after elections. The act of releasing a statement implies a rather common sense notion that election monitoring statements shape domestic populations' views of their elections in a straightforward way: positive statements should increase electoral legitimacy while negative ones should decrease it.¹ For instance, a National Democratic Institute election monitoring guide declares that in order to have maximal impact, observer groups must get their results out quickly as “citizens want to know whether the election is ‘going well’ or ‘going badly’ on election day” (Estok, Nevitte, & Cowan, 2002, p. 84). The UN’s “Declaration of Principles for International Election Observation”—which most monitoring agencies have signed—links the reporting of observation findings with the promotion of public confidence in the elections (United Nations, 2005, p. 2). Much of the literature on election monitoring follows this lead. For example, Hyde highlights that while not all domestic actors will agree on the necessity of international election monitors, their statements are usually taken seriously. Governments that “receive negative reports fail to signal to both domestic and international audiences that they are holding plausibly democratic elections” (2017, p. 13). Similarly, Carothers cautions that, “A sharp condemnation by foreign observers of a flawed election could precipitate serious violence or political instability” and that observers may end up avoiding such

¹ Literature on the influence of cue sources suggests that political information has a greater impact when the recipient trusts the messenger or when the message goes against the sender’s interests and thus seems more credible (Dragojlovic, 2015; Weber, Dunaway & Johnson, 2012). Surveys conducted in Jordan do not allow us to test this hypothesis, although we plan to test it in a separate paper on Tunisia.

statements out of fear of their potential negative effect on public opinion (1997, 25). On the other hand, scholars assume that incumbents seek a positive statement from international monitors in order to gain legitimacy, presumably among both international and domestic audiences (e.g., Rich, 2001, pp. 26; Kelley, 2012, pp. 7-9; Hyde & Marinov, 2014, pp. 331).

In a rare but important step toward evaluating the impact of observers' statements on citizen perceptions, and one worth discussing in detail, Bush and Prather (2017) test the impact of positive and negative electoral monitoring statements on the public's perceived electoral credibility. In a survey experiment conducted shortly after the 2014 Tunisian election, they find that statements have, on average, a small but statistically significant effect in the expected direction: a positive statement yields a small improvement in perceived electoral credibility and the negative statement a small decrease. They also find evidence of heterogeneous effects; among those who supported Ennahda, the losing party, negative statements led to a decline in perceived electoral quality.

Two theories of public opinion formation may explain these results. First, Bayesian updating suggests that the impact of statements on respondents' evaluations depends on the *certainty* that citizens have over their assessments. Losers may be less certain than winners in their own evaluations of the elections, they argue, and alter their assessment to be in line with the observers' assessment. Second, motivated reasoning suggests that the negative evaluations of monitors may simply be more in line with losers' partisan goals, leading these respondents to see the elections as even less credible, due to their desire to absorb information with which they agree. This argument is consistent with findings from a study of Russia by Robertson (2017), who finds that opponents to the regime were more likely to view election monitors as credible and, presumably, to incorporate their information into their assessments of elections.

Bush and Prather’s study makes an important contribution to the literature, but it has limitations. Their study takes a creative approach to considering the certainty by which voters hold their beliefs, and thus attempts to test a theory of Bayesian updating. However, the assumption that Nidaa Tunis voters (i.e., the winning partisans) are more certain in their assessment than Ennahda voters (i.e., the losing party’s supporters)—the measure of certainty in the paper—does not necessarily hold. Unfortunately, neither they nor we have the ability to test the Bayesian theory adequately.

Our study complements and extends Bush and Prather’s study in several ways. Their analysis is focused only on Tunisia, which limits the generalizability of its findings. Experimental context matters, as Pritchett and Sandefur (2013) argue after evaluating experimental outcomes of development programs. In this study, we evaluate the impact of election observation statements in two very different electoral contexts to consider how context may mediate the impact of observers’ statements and citizens’ assessments. Our study also has a longer timeline, examining effects longer after the elections. Bush and Prather find significant effects of statements immediately following the election in Tunisia, but, in this study, we find no evidence for effects seven months later. Taken together, these studies suggest that more work needs to be done to understand the impact of observers’ statements over time. Our work also helps answer a call for “greater awareness of the domestic effects of international democracy promotion [in order to understand] how this type of foreign assistance can progress most effectively” (Beaulieu & Hyde, 2009, p. 410). Finally, our approach to theorizing about voters as engaged recipients disaggregates two possible effects of statements that Bush and Prather subsume under the category of motivated reasoning. Literature in the fields of political communications, media studies, and psychology identify two distinct ways that the engaged

recipient could be engaged with and their opinions shaped by positive and negative statements. We distinguish between and test for these two mechanisms—biased assimilation and backfire effects. Finally, our work complements studies of foreign influence on elections on whether citizens believe their country should have closer relations with those external powers (Corstange & Marinov, 2012).

Elections in Jordan and Tunisia

Jordan and Tunisia offer unique leverage for testing hypotheses regarding the impact of monitoring statements on citizens' evaluations of their elections in two different contexts. Each country is demographically characterized by majority Arab, Muslim populations. However, the 2013 Jordanian and 2014 Tunisian elections differed in terms of each country's relationship with the international community, particularly the West. Because election monitoring is so closely associated with the West, and particularly the United States, it is important to consider how Jordanians and Tunisians express very different attitudes toward Western engagement in their respective countries. Moreover, the 2014 Tunisian elections were held in the context of transition, while the Jordanian elections took place amidst continued frustrations over stalled reforms in an authoritarian regime. These conditions may impact how EO statements shape citizens' perceptions of their elections.

Attitudes toward the International Community. At the time of the elections, Jordanians saw the US as supporting the monarchy and, relatedly, Israel—a political reality that arguably is more grating to Jordanians than Tunisians, given both the geographic proximity and the presence of a large population of Palestinian origin in Jordan. Tunisians, by contrast did not so strongly equate the international community's engagement with Tunisia as offering support for

authoritarian incumbents. Jordanians were thus more skeptical of the West's intentions than were Tunisians.

Public opinion polls also reflect these differences. In 2012, 45% of Jordanians polled in the Arab Barometer disagreed with the statement that Americans are good people, compared to 28% in Tunisia in 2013. A similar proportion of Jordanians (45%) disagreed that U.S. culture has positive attributes, compared 17% of Tunisians.² Moreover, 31% of Jordanians (2014)³ and 14% of Tunisians (2014)⁴ claim to have boycotted Western goods in response to international events (Benstead & Reif, 2017). However, a rather high proportion of both Tunisians (66%) and Jordanians (76%) agreed to some or a great extent that foreign interference is an impediment to reform in their country.⁵ When asked about the most important policy the U.S. should take toward the region, 36% of Jordanians and Tunisians, respectively, said the US should not interfere in their politics (Arab Barometer, "Wave III").⁶

² Arab Barometer Wave II (2010-2011) was conducted in Jordan in 2010 and Tunisia in 2011; Wave III was conducted in Jordan in 2012 and Tunisia in 2013.

³ GLD (2015), Yale University, poll conducted 2014 by Lust, Kao, and Benstead.

⁴ TPES (2014), poll conducted by Benstead, Lust, Malouche, and Wichmann.

⁵ "Foreign interference is an obstacle to reform in your country. Agree to a great extent, some extent, disagree, absolutely disagree."

⁶ "What is the most positive policy that the US can follow in our region? Promote democracy; Promote economic development; Contain Iran; Solve the Arab-Israeli Conflict; Promote women's rights; The US shouldn't interfere."

The Electoral Context. The electoral context also differed across the cases. Long periods of authoritarian rule meant that voters in Jordan and Tunisia generally had little attachment to political parties and had long believed that their ballots would have limited impact on the overall direction of the country. Yet, this began to change in 2011. The elections held in Tunisia in 2012 took place in a time of political transition and high uncertainty, while Jordan’s 2013 election occurred in a context of stasis. In both countries, however, party attachment remained weak, and “partisanship” had not fully developed.

Jordan implemented minor reforms to its electoral system in the lead-up to the 2013 elections, but they only heightened frustration. Jordanians had long experienced flawed elections, and the previous polls, held in 2010, were particularly problematic. Electoral observers and the international community had soundly criticized the electoral process and outcomes (NDI 2010), and in December 2010, nearly three quarters of Jordanians viewed the November 2010 elections as not free and fair (43%) or as having minor problems (30%; Arab Barometer, “Wave II”). In response, the king swore that the previous wrongs would not be repeated (Ibn Hussein 2012) and established a commission to revise the electoral law. Yet, reforms fell far short of serious change (NDI 2013). The Parliamentary Election Law, passed by royal decree in July 2012, established an Independent Electoral Commission and a national list system; however, it retained the SNTV system, which had been criticized heavily for decades,⁷ and did not remedy serious problems of malapportionment and gerrymandered districts (Kao 2015:47-51).

Frustration was evident in the lead-up to the 2013 elections, with a number of parties and political groups boycotting. And after the elections, about two-thirds of Jordanians polled

⁷ Kao and Lust served on the study mission that prepared The Carter Center (2013) report.

believed that politicians in their community tried to buy votes with gifts, money, or access to services.⁸

Unlike Jordanians, Tunisians believed that the 2014 elections would be free and fair, and they felt that the polls could determine the country's direction. Most had positive memories of the Constituent Assembly elections held in 2011, when campaigns were vibrant; voters flocked to the polls; and the process was heralded as remarkably successful.⁹ Three years later, the mood was more somber. The transition had been marred by economic decline, political instability, and terrorism, and citizens were frustrated. Yet, people continued to believe that the Constituent Assembly "mattered." In 2014, when asked if it was able to keep the government from taking too much power in its own hands, 38% of respondents said yes, 29% said no, and 33% did not know (TPES 2014). Tunisians also viewed vote-buying as much less prevalent. Only 13% said that they knew someone who was offered money or gifts in return for their vote in this election, and few said they personally had been offered money (2%) or gifts (2%). People were concerned about the outcome in the lead-up to the 2014 polls, but they were not resigned a flawed election.

Election Day went reasonably well in both countries. It appeared to leave Tunisians feeling more confident in the outcomes of their elections than it did Jordanians. Yet, importantly

⁸The GLD Jordan 2014 survey, within which the experiments are embedded, found that 62% of respondents agreed that many of the politicians in their community buy votes in exchange for gifts, money, or access to services (31% disagreed; 7% refused to answer). Similarly, 62% of respondents agreed that many people in their community sell votes in exchange for gifts, money, or access to services (while 32% disagreed; 6% abstained).

⁹ NDI (2011) and Council of Europe Parliamentary Assembly (2011).

for our study, elections were neither so perfect nor so flawed as to make positive or negative statements from election monitors seem ludicrous. Turnout was similar in the two elections: 57% in Jordan and 62% in Tunisia. Domestic and international observers covered both elections (Appendix 2), and while they reported some localized violence and disruptions (Al-Shibani 2014), there were no major disturbances or problems deemed to have altered outcomes significantly.

Finally, it is important to note that partisanship is weak in both Jordan as well as Tunisia. Jordanian political parties have historically been weak (Lust-Okar 2001; Kao 2015), with tribes often playing the role of mobilizing voters. The 2014 GLD survey in Jordan found that less than 3% of respondents state that they were members of political parties, and only 13% of respondents who contacted elected officials to seek assistance chose to contact the official because he or she was a member of a party they support. Similarly, the Arab Barometer finds that 0.03% of the Jordanian population claims membership in a political party (Arab Barometer, “Wave III”).

Parties are similarly weak in Tunisia. Only 3% of respondents in a 2014 TPES poll conducted in Tunisia reported being members of political parties in Tunisia; 2% reported membership in a Tunisian political party. Moreover, the poll revealed that only 40% thought they knew the party they would vote for if elections held tomorrow, with 20% leaning toward a political party and 40% undecided, and there was considerable shifting in party preference between the 2011 Constituent Assembly elections and the 2014 elections. Finally, citizens have low trust in political parties in both countries; the 2014 TPES poll in Tunisia found that 61% of respondents stated that they had very low trust in political parties and 22% had low trust. The Arab Barometer found in 2010 found that 55% of Jordanians do not trust parties or trust them to

a limited extent (Arab Barometer, “Wave II”). In short, partisanship should not be overstated in either Jordan or Tunisia.

Effects of Election Monitoring

We consider three hypotheses. The first is drawn from assumptions in the literature on election observation and from election monitoring agencies themselves, and it predicts homogeneous effects. The second two hypotheses are developed from literature in psychology and political communication, and they propose heterogeneous effects. Our initial assumption was that these hypotheses should operate similarly in different political contexts. Testing our hypotheses across two very different contexts, however, allowed us to test this assumption explicitly. This much-needed test of the generalizability of experimental findings drew these expectations into question. The study thus lays the foundation for a comparative framework to better understand how and why effects of monitoring statements vary both within and across cases.

Homogeneous Effects. As highlighted above, a widespread, implicit assumption underlying much of the literature and policy programming holds that when EOs report that elections were flawed, citizens view the elections as illegitimate; when international monitors report clean elections, the citizens evaluate the outcomes of elections as more legitimate. This assumption underpins much of the literature linking election observation to boycotts, protests, and other elite strategies. For instance, Hyde and Marinov (2014: 331) argue that particularly in political contexts where objective information is limited, opposition groups point to negative evaluations to mobilize demonstrations and protests. They recognize that some observers may be more reputable than others, and that international observers may have more impact in some contexts than others. However, they generally view EOs as providing credible information and

their impact to be positive or, at worst, disregarded. Based on their reading of hundreds of observer reports, they argue that even one negative statement from international monitors is likely both to “cast doubt” on the quality of the elections among domestic populations and potentially lead to post-election protests, whereas a positive statement from monitoring agencies will prevent such outcomes (2014: 337).

Thus, Hypothesis 1 argues that voters’ views of the election will be shaped in the direction of the EO’s statements. This would be consistent with Bush and Prather (2017), who find evidence for small average treatment effects when comparing the positive and negative treatments.

Hypothesis One. Voters accept and are swayed by positive and negative statements (Uniform effect). Positive assessments will make citizens perceive elections as being more representative of the will of the people, while negative assessments will reduce citizens’ perception of the elections’ legitimacy.

Heterogeneous Effects. Research on election observation generally predicts that statements affect all citizens equally, but research in political psychology and communication suggests that receivers are susceptible to several types of cognitive biases, and thus that the effects of statements may depend on the type of voter. As noted above, Bush and Prather (2017) find evidence for such heterogeneous effects in Tunisia, as does Robertson (2017) in a study of election monitors’ legitimacy in Russia, and Masoud, Jamal, and Nugent (2016) in a survey experiment on religious endorsements of women’s rights issues. Our own previous work has found evidence of similar effects on candidate electability (Benstead, Jamal, and Lust 2015) and attitudes about human rights issues (Muriaas, Wang, Benstead, Dulani, & Rakner 2018).

We examine two heterogeneous effects. The first is what psychologists call “biased assimilation,” or “confirmation bias.”¹⁰ Research in psychology finds that people tend to interpret evidence so as to maintain their prior beliefs (Allport, 1979; Lord, Ross, & Lepper, 1979). Thus, recipients are more likely to accept information that is consistent with their beliefs, to view ambiguous statements as being consistent with their views, and to ignore inconsistent information. This suggests that statements will impact recipients’ views when they are consistent with their views, making recipients believe more strongly in, or take more extreme positions toward, their initial assessment. Citizens discount or counter statements that are inconsistent with their prior beliefs and thus these statements will not impact their attitudes. Recipients of positive or negative observer assessments, when these are consistent with their initial beliefs, may view the elections as even more positive or negative, respectively. Statements that are contrary to their pre-established positions will have little or no impact on their views of the elections.

Other scholars have built upon this research to show that in processing information that is inconsistent with their beliefs, recipients may come to hold their original opinion even more strongly. This is referred to as a “backfire effect”. Information inconsistent with one’s existing views creates cognitive dissonance, forcing recipients to take more time and effort to process the information. In the process, they often return to their original convictions and hold them more strongly than before (Ditto and Lopez 1992; Nyhan and Reifler 2010). For instance, in a dynamic

¹⁰ Kunda (1990); Molden and Higgins (2005). Lord et al.’s (1979) shows how people are biased by prior attitudes. Multiple laboratory (Edwards & Smith, 1996; Miller, McHoskey, Bane & Dowd, 1993; Zuwerink & Devine, 1996) and field studies (Ahluwalia, 2000; Munro et al. 2002) confirm this bias.

process tracing experiment, Redlawsk (2002) finds that subjects take longer to process information that is inconsistent with their political views and seek out more information consistent with their views. This leads the respondents to develop stronger support for their preferred candidate, even after encountering negative information. These studies suggest that citizens' attitudes toward the government shape preconceptions about the electoral process and impact how observers' statements influence their assessments of the elections. Those who strongly support the government, when faced with statements that challenge its legitimacy, are more likely to express stronger beliefs consistent with their initial positions.

Thus, we consider two hypotheses:

Hypothesis Two. Voters accept consistent information and dismiss inconsistent information (Biased assimilation). Positive statements will lead those who already see the elections as legitimate to view them as more legitimate, but positive statements will not affect those with prior negative beliefs toward the elections. Negative statements will lead those who believed that the elections were illegitimate to see the elections in a more negative light, but they will not affect those who had positive prior beliefs.

Hypothesis Three. Voters counter-argue inconsistent information, leading to strong effects in the opposite direction for those who disagree (Backfire effect). Recipients who receive information that is inconsistent with their views will develop even stronger views in the opposite direction of the statement.

Survey Experiment

We employ a framing experiment to examine the relationship between messages and citizens' attitudes toward elections. This allows us to examine how monitors' statements shape citizens' evaluations of electoral legitimacy. By doing so in the very different contexts described above,

we are able to consider how the impact of election statements may vary, depending on the context of the elections and attitudes toward the West, and to determine the extent to which findings are generalizable.

The survey experiments were embedded in the 2014 Jordanian GLD Survey¹¹ and the 2015 Tunisian Local Governance Performance Index (LGPI) survey.¹² These household surveys of 1,499 Jordanians and 3,659 Tunisians, respectively, were conducted face-to-face by local teams using tablet computers.¹³ We randomly assigned respondents to treatment groups in which they were given positive, negative, or no statements regarding EOs' assessments, and then were asked to rate the extent to which they believed the elections represented the will of the people.¹⁴

¹¹ GLD (2014), poll conducted 2014 by Lust, Kao, and Benstead.

¹² GLD (2015), poll conducted in 2015 by Benstead, Landry, Lust, and Malouche. Replication file for these two surveys will be made available through ICPSR and the project websites.

Studies approved by the appropriate institutional committee and performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its amendments.

¹³ The surveys took about 40 minutes. Survey experiments were embedded in sections on voting and placed near the beginning of the survey. There is no reason to expect differences in the surveys due to placement or contamination of questions surrounding the experiment. Further, the sampling designs differed slightly given different goals of the surveys, but there is no reason to believe that this affected results. See Appendix 2.

¹⁴ Due to a tablet survey application error in Jordan, the three experimental conditions were allocated only to voters. Voters are those who reported voting in the election under study (i.e., 2014 election). Thus, we limit our analyses to voters. In Jordan and Tunisia we opted for a ten-

The negative treatment group received a statement that said observers' reported that the elections were marred by problems. The control group received no statement. The positive treatment group received a statement that observers deemed that safeguards improved the transparency of the elections (Table 1). Random assignment of the experiments was effective, as indicated by insignificant chi-square tests of independence between the condition and all other independent variables in the analysis. (See Appendix 1 for further discussion).

We adapted all statements from actual Carter Center public statements regarding the Jordanian and Tunisian elections. We chose segments of reports such that statements included in the question were similar in length across the conditions and in severity of the statement. Question stems were identical across the treatments and countries. Yet we intentionally chose wording that was accurate but not identical across the countries.

Some may object to this approach because statements differed slightly across the countries, given the content of available reports; thus, it is worth discussing in more detail. First, in order to uphold the highest standard of integrity and ethical research practices, we did not want to engage in deception. Moreover, we wanted to use statements that conveyed enough information to appear credible. Thus, we did not simply want to make a general statement, for instance, that "observers criticized (applauded) the elections." Real statements allowed us to satisfy these concerns. Finally, we sought to simulate experiences that could actually occur in the real world, which is a necessary condition for a survey experiment to have merit (Gaines, Kuklinski & Quirk, 2006). We carefully considered different design options and chose an

point Likert scale in an attempt to capture more nuance in responses. To ease interpretation, we transformed the ten-point scales into four-point scales. See Appendix 3.

approach with high ethical standards and internal validity, even at the potential cost of a modest impact on our ability to compare across cases.

[Table 1]

Moreover, we chose this approach because of its limited drawbacks. It is common practice to standardize question wording in cross-national surveys, but consistent wording does not always ensure consistent meaning (Harkness et al., 2010). At the same time, the extent to which effects of the statements exist, they do so with regard to the cross-national comparability of the experiments, not the individual results within each country. If this assumed, though implicit, wisdom holds, we should still expect that individuals receiving positive treatments in all countries would be more likely to see the elections as legitimate, and those receiving negative statements would be less likely to do so.

It is also worth noting that due to logistical considerations, the time lag between the elections and surveys varied across the countries, with seven months in Tunisia and 15 months in Jordan. The difference in timing may at first appear to be a weakness of the study, and certainly it would be ideal to conduct the experiment at different periods following elections, in order to test how the impact of statements may change over time. However, timing does not explain the differences in the outcomes reported below. Indeed, to preview results, we find a significant impact of the statements on attitudes toward the elections in Jordan, where the time lag between the elections and survey was long, and an insignificant impact in Tunisia, where the time lag was short. If timing was a factor, we would expect to find larger effects in Tunisia than in Jordan.

Measurement

Testing the hypotheses requires measurement of the respondent's evaluation of the elections after receiving a statement—our dependent variable—and of their view of the election's legitimacy

before receiving a statement, which we expect to impact how respondents receive the observer's statements. Measuring a respondent's evaluation of the elections following the statement is straightforward; we use responses to the questions set forth in Table 1 to measure the individual's evaluation of the extent to which the elections represent the will of the people. We measure the individual's view of the elections before receiving a statement using the individual's assessments of the government. (See Table 2). People who are dissatisfied with the government are less likely to view the elections as representing the people's will, and vice versa for those who are satisfied with the performance of the government. We recognize that these views may have shifted in the period since the election, given more or less adept government performance after polling, and this measure may represent a retrospective view on election quality. However, because we aim to examine the impact of statements on the voters' assessments at the time of the survey, the measurement is appropriate. It provides insight into the evaluation of the election when the experiment is implemented.

Finally, it is worth noting that we do not examine the impact of partisanship on the responses to monitors' statements. In doing so, we differ from other scholars who have considered partisanship (Bush & Prather, 2017). However, given the low level and shifting nature of partisanship in Jordan and Tunisia, discussed above, we focus on satisfaction with government as the best indicator of support for the regime.

[Table 2]

Results

We examine the hypotheses derived from assumed wisdom in policy and scholarly literature, as well as social psychology and media studies. To interrogate the first hypothesis (i.e., Uniform effect), we examine the average treatment effects of monitoring statements in Jordan

and Tunisia. The second and third hypotheses (i.e., biased assimilation and backfire effects) require analyses of heterogeneous treatment effects.

Testing Homogeneous Treatment Effects. We use ordered logistic regression to estimate average treatment effects of international election monitoring statements.¹⁵ Table A4.1 (see Appendix Four) compares the effects of negative and positive statements on perception of election legitimacy to the control group that received no statement from an international monitoring agency. The first model for each country (1 and 3) presents the effect of the election monitoring statements alone, while the second model for each country (2 and 4) presents the effect of these statements including controls for respondent age, education level, and gender, as well as an additional independent variable measuring satisfaction with the government.

Hypothesis One holds that positive statements would be associated with more positive assessments, while negative statements would be associated with more negative assessments of the elections. Yet, in Jordan and Tunisia, no treatment group was significantly different than others in terms of their perceptions of the legitimacy of the elections, even when controls are included in the model. Our measure of satisfaction with the government is the only variable that significantly predicts perceptions of the legitimacy of the elections across the two cases.

Testing Heterogeneous Treatment Effects. To test for heterogeneous effects, we ran ordered logistic regression models for each country including interactions between measures of satisfaction with the government and the experimental treatments and we computed the predicted

¹⁵ For the sake of simplicity, we do not present models that controlled for age and education level, but we found neither to significantly affect the outcomes. See Appendix 4 for details.

probabilities. (See Tables A4.2 and A4.3 in Appendix 4). To ease interpretation, we present and discuss the results in graphical form.

The graphs show comparisons between a treatment group (receiving a positive or negative statement) and the control group (receiving no statement). Respondents are grouped according to their views of electoral legitimacy, as measured by their level of satisfaction with the current government, measured on a four-point scale: Very satisfied, somewhat satisfied, not very satisfied, and not at all satisfied. In the graphs, bars that fall below the horizontal axis show that those receiving the statement are less likely to select the corresponding level on the dependent variable—i.e., assessment of the elections statement—than those in the control group. When the bars fall above the horizontal axis, those receiving the statement are more likely than the control group to state the corresponding assessment.

We find that the extent to which observers' assessments influence citizens' evaluations of their elections varies according to the individual's satisfaction with the government and by country. We do not find strong evidence for the uniform influence or biased assimilation hypotheses. In Jordan, however, the effects are the strongest, and we find strong evidence for a backfire effect (Hypothesis Three), significant at the $p < .05$ level. Effects are weak and insignificant in Tunisia. We argue that the size of the impact is driven primarily by individuals' initial attitudes toward the elections and toward the messenger.

Uniform Influence (Hypothesis One) and Biased Assimilation (Hypothesis Two). We find no support for the biased assimilation hypothesis in either Jordan or Tunisia, or the uniform influence hypotheses. As shown in Figure 1, we find some evidence that the negative statement influences those who are somewhat satisfied with the government in the direction of the statement. Jordanians who are somewhat satisfied with the government and receive a negative

statement are nearly six percentage points less likely to say that the elections are very representative, and about four percentage points less likely to say that the elections are somewhat representative compared to the control group (depicted by the top two horizontal bars for the “somewhat satisfied” group in Figure 1). Moreover, they are nearly four percentage points more likely to say that the elections are not very representative and more than six percentage points more likely to say that they are not representative at all, when compared to the control group. (See the bottom two bars for the “somewhat satisfied” group in Figure 1). These results are significant at the $p \leq .05$ level. However, it should be noted that the uniform effect should hold consistently across respondents. That is, the effects should be homogeneous. Since we do not find this consistent effect across all types of voters, the experimental results in Jordan do not support this aspect of the hypothesis.

[Figure 1]

Backfire Effects (Hypothesis 3). Rather, the backfire effect hypothesis receives the strongest support. As shown in Figure 2, the effect is most pronounced in Jordan, where many citizens are likely to have a visceral reaction due to their perception of external actors’ impact on their political context. Respondents that received the positive statement about the elections and who are very satisfied with the government are twenty-six and seven percentage points less likely to see the elections as representing the will of the people to a very high degree or to some degree, respectively, compared to those who are in the control group. (See the first two bars in the ‘very satisfied’ respondent group). Similarly, very pro-government respondents who received a positive statement are seventeen and sixteen percentage points more likely to view the elections as not being representative of the will of the people to some degree or to a very high degree, respectively, compared to the control group. This is shown by the third and fourth horizontal bars

in Figure 2. (The effects are significant at the $p \leq .05$ level). Receiving a positive statement leads those who are inclined to view the elections as legitimate to be less likely to state that the elections are legitimate. This is consistent with the notion that citizens hold prior views and engage with observer statements.

[Figure 2]

There is limited evidence of an effect of observer statements on Tunisians' views of elections. Figure 3 shows no significant results, although the outcomes suggest a potential backfire effect. For instance, the negative statement increases the likelihood of people who are not at all satisfied with the government to see the elections as being very much or somewhat representative of the will of the people and *decreases* the likelihood by seven percentage points of them seeing the elections as not representative at all. Among Tunisian respondents who are very satisfied with the government, a negative statement is associated with a five percentage point increase in the likelihood that respondents see the elections as very legitimate. Yet, these results are not significant at conventional significance levels and we suspect they are weaker because Tunisians do not hold negative views of the West as strongly as Jordanians do. We find this result despite the fact that the experiment was conducted in Tunisia much closer to the election and with a larger number of respondents (1,985) than in the other cases. (See also Figure A1.1 in Appendix 1).

[Figure 3]

In sum, the survey experiment conducted in Jordan and Tunisia draws into question assumed wisdom about the uniform impact of EOs' statements and gives reason to expect that their assessments can have the opposite impact of that intended. As shown in Table 3, citizens hear and are sometimes convinced by statements, as shown by the homogenous treatment effects

in Jordan and Tunisia. Yet, statements cause a backfire effect, especially in Jordan, where western influence on domestic politics is often resented. Some citizens in Jordan hear observers' statements and adjust their evaluations in the opposite direction of the information that they are given.

[Table 3]

Discussion

The results are complex and differ significantly across the two countries. We find little support for the uniform influence and biased assimilation effects, but greater support for the backfire hypothesis. We also reveal important differences across the countries. This draws into question the extent to which findings from any single experimental study should be generalized to other contexts, and highlights the need to consider the factors that explain these differences.

One potential explanation, which we do not believe applies here, could relate to the time since the election. As noted above, one might reasonably argue that the impact of monitors' evaluations diminishes over time. As time goes by after elections, individuals have received and processed information from a number of sources, presumably leading them to develop a settled opinion regarding the polls' legitimacy. Given this, we may have expected stronger effects in Tunisia, where the survey was conducted more quickly after the election, than in Jordan.

The fact that this was not the case is important and suggests that election monitoring statements may have political relevance long after the elections are held. Timing not only fails to explain the differential effects of the statements revealed in the two countries, but it appears that the impact of monitoring statements may diminish over time much less than one originally expects. Monitoring statements, like other focal points, can be important as much for the

underlying political message that they send as they are for the direct information about elections that they contain.

While other rival hypotheses are possible, a second, and we believe more likely, explanation is found in the attitudes toward the West. The engaged recipient hypotheses are based on an understanding of the receivers' prior beliefs about electoral legitimacy, but also on their attitudes toward the messenger. As discussed above, Jordanians and Tunisians hold very different attitudes toward the West, and particularly the United States. Specifically, Tunisians generally hold positive attitudes toward the West, while most Jordanians are skeptical (Benstead, 2017). It is thus not surprising that backfire effects are most significant in Jordan. That is, the impact of observers' statements on assessments of electoral legitimacy may depend on the statement's content, the recipient's prior beliefs, and their attitudes toward the messenger.

Finally, it is important to consider the potential concern that the survey experiment measures attitudes, not behavior. Even if we understand the extent to which statements impact citizens' assessments of the elections, does it matter if it does not compel them to act—taking to the streets, supporting the regime, or otherwise *acting* in accordance with their belief? We argue it does. Statements are short-lived but public discourse that surrounds election outcomes shape subsequent political dynamics. If positive statements induce some citizens to make statements about the elections that are more negative than they otherwise would make—as they react to positive assessments from a disliked source—this becomes fodder for the opposition to challenge incumbents who gain their position through 'clean' elections.

Conclusion

The survey experiments conducted in Jordan and Tunisia suggest that EOs' assessments of elections may influence respondents, but their impact is neither as strong nor straightforward

as is often assumed by scholars and practitioners. Assessments do not uniformly influence respondents' evaluations of elections in the direction of the statement. Positive assessments do not always push citizens to see the elections in a better light, and negative assessments do not necessarily prompt them to view elections more negatively. This runs counter to the widespread assumption that citizens are easily swayed by these statements in predictable ways. Observers' statements have different impacts on respondents, depending on their initial attitudes toward the government and the political context of the elections.

This is not wholly surprising. Citizens are not blank slates. They have their own views of the elections and their own assessment of the sources of information about their elections. When they have strong beliefs about the government, there is no reason to believe that monitors' assessments change their subsequent evaluations. Similarly, when the information presented to them contradicts their views, they are not likely to be swayed by assessments, at least not in the direction of the message. Indeed, if anything, the strongest findings from these experiments demonstrate that those who dislike the message respond in a contrarian manner; they are more likely to give negative evaluations when presented with positive assessments.

Policymakers and development specialists, whose efforts are based on the principle of "Do no harm," need to take such backfire effects into account when formulating policies. Non-governmental organizations, international actors, and governments spend billions of dollars every year funding election monitoring. To a large extent, they decide where to invest based on their ability to monitor without facing significant political obstacles and the prospects for democratization. Our analyses, however, suggest that they should also take into consideration how the local population will respond to monitors. Even the most professional and well-intentioned monitoring organizations can be viewed by the local population as tainted by their

association with outside (often Western) forces. Practitioners should recognize the extent to which this affiliation may undermine their message, and consider the risk of backfire effects when determining how and where to invest time and resources. Until this danger taken into account, their programs may have the unintended effect of undermining the electoral legitimacy that they seek to support.

The findings are also consequential for opposition groups. They need to be aware that negative statements may inadvertently strengthen support for the regime. Where the West is particularly disliked, there is reason to believe that incumbents have little to fear from the presence of negative statements. Indeed, negative evaluations may actually help to strengthen their regime.

These findings have serious implications for the scholarly work on election observation as well. Many theories linking election observation to political protest, electoral manipulation, and other factors are based on the underlying assumption that positive statements enhance the legitimacy of elections and negative statements undermine it. Yet, at least in some cases and for some citizens, positive statements may make citizens more skeptical about the quality of their elections, while negative statements may enhance perceptions that elections were effective. In both theory building and empirical testing, scholars need to be careful about assuming that the relationship between monitors' evaluations and citizens' responses are positively correlated.

Given the importance of democratic transitions and the efforts put into monitoring, further study is warranted. Until we achieve improved understanding of the impact of EOs' assessments on citizens' assessments, we cannot fully understand how election observation relates to election protest, democratization, and other outcomes. We also cannot determine where and when to invest resources most effectively. And, despite best intentions, we are stymied in

our efforts to promote democratic processes that yield results more faithfully representing the will of the people.

References

- Ahluwalia, R. (2000). Examination of Psychological Processes Underlying Resistance to Persuasion. *Journal of Consumer Research*, 27(2), 217-232.
- Al-Shibani, H. (2014, October 24). Tunisia Rocked by Pre-Election Violence. *Independent Online (IOL)*. Retrieved from <http://www.iol.co.za/news/africa/tunisia-rocked-by-pre-election-violence-1770058>
- Allport, G. W. (1979). *The Nature of Prejudice*. Reading, MA: Addison-Wesley Publishing Co. (Reprinted from *The Nature of Prejudice* by Allport G. W., 1954, Cambridge, MA: Basic Books).
- Arab Barometer.Wave II. Conducted 2010 in Jordan and 2011 in Tunisia. Retrieved from: <http://www.arabbarometer.org/content/arab-barometer-ii>
- Arab BarometerWave III. Conducted 2012 in Jordan and 2013 in Tunisia, Retrieved from: <http://www.arabbarometer.org/content/arab-barometer-iii>
- Asunka, J., Brierley, S., Golden, M. A., Kramon, E., Ofosu, G. (2013, August). Election Observers and Electoral Fraud. *Annual Meetings of the American Political Science Association 2013*. Retrieved from <https://ssrn.com/abstract=2319651>
- Beaulieu, E. & Hyde, S. D. (2009). In the Shadow of Democracy Promotion: Strategic Manipulation, International Observers, and Election Boycotts. *Comparative Political Studies* 42(3), 392-415.
- Benstead, L. J. (2017, March). *Public Opinion Framework: Evolving Perceptions of Foreign Policy Issues*. Paper prepared for the conference on the Evolving Role of Non-State Actors on the Formulation and Implementation of North African Foreign Policies during the Arab Spring (2011-17), Loyola University, Chicago, IL.
- Benstead, L. J., Jamal, A. A., & Lust, E. (2015). Is It Gender, Religiosity or Both? A Role Congruity Theory of Candidate Electability in Transitional Tunisia. *Perspectives on Politics* 13(1), 74-94.
- Benstead, L. J., & Reif M. (2017). Coke, Pepsi or Mecca Cola? Why Product Characteristics Affect the Likelihood of Collective Action Problems and Boycott Success. *Politics, Groups, and Identities*, 5(2), 220-241.
- Benstead L. J., Lust, E., Malouche, D., & Wichmann, J. (2014). Tunisian Post-Election Survey (TPES). Retrieved from <http://transitionalgovernanceproject.org/>.
- Benstead, L., Lust, E., and Malouche, D. 2012. "Tunisian Post-Election Survey (TPES)." Retrieved from <http://transitionalgovernanceproject.org/>.

- Benstead, L., Landry, P., Lust, E., and Malouche, D. 2015. "Tunisian Local Governance Performance Index (LGPI)." Retrieved from <http://gld.gu.se/en/research-projects/lgpi>
- Bush, S. S., & Prather, L. (2017). The Promise and Limits of Election Observers in Building Election Credibility. *Journal of Politics*, 79(3), 921-935.
- The Carter Center. (2013). *The Carter Center releases study mission report on Jordan's 2013 parliamentary elections* [report]. Atlanta, GA: The Carter Center. Retrieved from: https://www.cartercenter.org/resources/pdfs/news/peace_publications/election_reports/jordan-2013-study-mission-eng.pdf
- The Carter Center. Democracy program. Retrieved from <https://www.cartercenter.org/peace/democracy/>
- Carothers, T. (1997). The Observers Observed. *Journal of Democracy*, 8(3), 17-31.
- Center for Global Development. (2013, August). *Context Matters for Size: Why External Validity Claims and Development Practice Do Not Mix* (Working Paper No. 336). Washington, DC: Pritchett, L. & Sandefur, J.
- Corstange, D. & Marinov, N. (2012). Taking Sides in Other People's Elections: The Polarizing Effect of Foreign Interventions. *American Journal of Political Science* 56(3), 655-670.
- Council of Europe Parliamentary Assembly. (2011, October 24). PACE Mission: Statement by the PACE Mission to observe the election of the Constituent National Assembly of Tunisia 2011. [election observation]. Retrieved from: http://assembly.coe.int/ASP/NewsManager/EMB_NewsManagerview.asp?ID=7091&L=2 (accessed 11 November 2014).
- Ditto, P. H. & Lopez, D. F. (1992). Motivated Skepticism: Use of Differential Decision Criteria for Preferred and Nonpreferred Conclusions. *Journal of Personality and Social Psychology*, 63 (4), 568-584.
- Dragojlovic, N. (2015). Listening to Outsiders: The Impact of Messenger Nationality on Transnational Persuasion in the United States. *International Studies Quarterly* 59(1), 73-85.
- Edwards, K & Smith, E. E. (1996). A Disconfirmation Bias in the Evaluation of Arguments. *Journal of Personality and Social Psychology* 71(1), 5-24.
- Enikolopov, R., Korovkin, V., Petrova, M., Sonin, K., & Zakharov, A. (2013). Field Experiment Estimate of Electoral Fraud in Russian Parliamentary Elections. *Proceedings of the National Academy of Sciences*, 110 (2), 448-452.
- Estok, M., Nevitte, N., & Cowan, G. (2002). *The Quick Count and Election Observation: An NDI Handbook for Civic Organizations and Political Parties*. Washington D.C.: National Democratic Institute for International Affairs. Retrieved from <https://ndi.org/node/24021>

- Gaines, B. J., Kuklinski, J. H., & Quirk, P. J. (2006). The Logic of the Survey Experiment Reexamined. *Political Analysis*. Advance online publication. doi:10.1093/pan/mpi008
- Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L. E., Mohler, P. Ph., Smith, T. W. (Eds.).(2010). *Survey Methods in Multicultural, Multinational, and Multiregional Contexts: Vol. 552*. Hoboken, NJ: John Wiley & Sons.
- Harriman Institute, Columbia University. (August, 2012). Making Voters Count: Evidence from Field Experiments about the Efficacy of Domestic Election Observation. (Working Paper No.1). New York, NY: Sjoberg, F. M.
- Hyde, S. D., (2007). The Observer Effect in International Politics: Evidence from a Natural Experiment. *World Politics* 60(1), 37-63.
- Hyde, S. D. (2017). *The Pseudo-Democrat's Dilemma: Why Election Observation Became an International Norm*. Ithaca, NY: Cornell University Press.
- Hyde, S. D. & Marinov, N. (2014). Information and Self-Enforcing Democracy: The Role of International Election Observation. *International Organization* 68(2), 329-359.
- Ibn Hussein, King Abdullah II (Speaker). (2012, October 23). *Gathering with national public figures* [Speech]. Retrieved from: <http://www.youtube.com/watch?v=e17kBQTs1Zk>
- Ichino, N. & Schundeln, M. (2012). Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana. *Journal of Politics* 74(1), 292-307.
- Kao, K. (2015). *Ethnicity, Electoral Institutions, and Clientelism: Authoritarianism in Jordan*. (Doctoral dissertation).University of California, Los Angeles, CA.
- Kelley, J. G. (2012). *Monitoring Democracy: When International Election Observation Works, and Why it Often Fails*. Princeton: Princeton University Press.
- Kunda, Z. (1990). The Case for Motivated Reasoning. *Psychological Bulletin*, 108(3), 480-498.
- Long, S. J. & Jeremy, F. (2006). *Regression Models for Categorical Dependent Variables using Stata*. College Station, TX: Stata Press.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence. *Journal of Personality and Social Psychology*, 37(11), 2098-2109.
- Lust, E., Kao, K., and Benstead, L. 2014. "Governance and Local Development and Post-Election Survey in Jordan."

- Lust-Okar, E. (2001). The Decline of Jordanian Political Parties: Myth or Reality? *International Journal of Middle East Studies*, 33(4), 545-570.
- Marzouk Z. (2014, October 23). Election Watchdog Atide Highlights ISIE Irregularities. *Tunisia-live*, retrieved from: <https://www.tunisia-live.net/>
- Masoud, T., Jamal, A., Nugent, E. (2016). Using the Qur'ān to Empower Arab Women? Theory and Experimental Evidence From Egypt. *Comparative Political Studies*, 49(12), 1555-1598.
- Miller, A. G., McHoskey, J. W., Bane, C. M., & Dowd, T. G. (1993). The Attitude Polarization Phenomenon: Role of Response Measure, Attitude Extremity, and Behavioral Consequences of Reported Attitude Change. *Journal of Personality and Social Psychology*, 64(4), 561-574.
- Molden, D. C., & Higgins, T. E. (2005). Motivated Thinking. In K. Holyoak, & B. Morrison (Eds.), *Cambridge Handbook of Thinking and Reasoning*. (pp. 295-320). Cambridge: Cambridge University Press.
- Munro, G. D., Ditto, P. H., Lockhart, L. K., Fagerlin, A., Gready, M., Peterson, E. (2002). Biased Assimilation of Sociopolitical Arguments: Evaluating the 1996 US Presidential Debate. *Basic and Applied Social Psychology*, 24(1), 15-26.
- Muriaas, R., Wang, V., Benstead, L., Dulani, B., & Rakner, L. (2018). It Why the Gender of Traditional Authorities Matters: Intersectionality and Women's Rights Advocacy in Malawi. *Comparative Political Studies*. doi: [10.1177/0010414018774369](https://doi.org/10.1177/0010414018774369).
- NDI. (2010). Final International Election Observation Report on the Jordanian Parliamentary Elections. Retrieved from: <https://www.ndi.org/files/Jordan-International-Election-Report-2010-ENG.pdf>
- NDI. (2011). Final Report on the Tunisian National Constituent Assembly Elections. [report]. Washington, DC: NDI. Retrieved from: https://www.ndi.org/files/tunisia-final-election-report-021712_v2.pdf.
- NDI. (2013). International Observation Mission, Jordan. Retrieved from: <https://www.ndi.org/files/Jordan-International-Mission-Prelim-Statement-012413-ENG.pdf>.
- Nyhan, B. & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*, 32(2), 303-330.
- The Program on Governance and Local Development (GLD). (<https://gld.gu.se/>)
- Redlawsk, D. P. (2002). Hot Cognition or Cool Consideration? Testing the Effects of Motivated Reasoning on Political Decision Making. *Journal of Politics*, 64(4), 1021-1044.

- Robertson, G. (2017). Political Orientation, Information and Perceptions of Election Fraud: Evidence from Russia. *British Journal of Political Science*, 47(3), 589-608.
- Rich, R. (2001). Bringing Democracy into International Law. *Journal of Democracy*, 12(3), 20-34.
- Simpser, A. & Donno, D. (2012) Can International Election Monitoring Harm Governance? *Journal of Politics*, 74(2), 501-513.
- United Nations. (2005). *Declaration of Principles for International Election Observation and Code of Conduct for International Election Observers*. New York: United Nations.
Retrieved from: https://www.ndi.org/sites/default/files/1923_declaration_102705_0.pdf
- Weber, C., Dunaway, J., & Johnson, T. (2012). It's All in the Name: Source Cue Ambiguity and the Persuasive Appeal of Campaign Ads. *Political Behavior*, 34 (3), 561-584.
- Zuwerink, J. R. & Devine, P. G. (1996). Attitude Importance and Resistance to Persuasion: It's Not Just the Thought that Counts. *Journal of Personality and Social Psychology*, 70 (5), 931-944.

Tables

Table 1. Experimental conditions and question wording for the dependent variable

Jordan (Survey conducted April 2014 --Following Jan. 2013 elections)	Negative Treatment	An international organization that monitored the elections said that: “persistent concerns about vote buying, proxy registration, and other problems marred the Jordan’s 2013 elections.” To what extent do you believe that the 2013 parliamentary election results represented the will of the people? Please tell me on a scale of 1 to 10, where 1 is ‘not at all’ and 10 is ‘very much.’ (Scale: 1=Not at all/10=Very much)
	Control	To what extent do you believe that the 2013 parliamentary election results represented the will of the people? Please tell me on a scale of 1 to 10, where 1 is ‘not at all’ and 10 is ‘very much.’ (Scale: 1=Not at all/10=Very much)
	Positive Treatment	An international organization that monitored the elections said that: “the new Independent Election Commission...introduced several important procedural steps to safeguard ballot secrecy, improve electoral administration, and promote transparency in Jordan’s 2013 elections.” To what extent do you believe that the 2013 parliamentary election results represented the will of the people? Please tell me on a scale of 1 to 10, where 1 is ‘not at all’ and 10 is ‘very much.’ (Scale: 1=Not at all/10=Very much)
Tunisia (Survey conducted May 2015—Following Oct. 2014 elections)	Negative Treatment	An international organization that monitored the 2014 parliamentary elections in Tunisia said: “irregularities were reported in a limited number of polling stations, [including]... illegal campaigning and inadequate...polling staff.” To what extent do you believe that the 2014 parliamentary election results represented the will of the people? Please tell me on a scale of 1 to 10, where 1 is ‘not at all’ and 10 is ‘very much.’ (Scale: 1=Not at all/10=Very much)
	Control	To what extent do you believe that the 2014 parliamentary election results represented the will of the people? Please tell me on a scale of 1 to 10, where 1 is ‘not at all’ and 10 is ‘very much.’ (Scale: 1=Not at all/10=Very much)
	Positive Treatment	An international organization that monitored the 2014 parliamentary elections in Tunisia said: “ISIE electoral authorities succeeded in a relatively smooth and orderly implementation of the elections.” To what extent do you believe that the 2014 parliamentary election results represented the will of the people? Please tell me on a scale of 1 to 10, where 1 is ‘not at all’ and 10 is ‘very much.’ (Scale: 1=Not at all/10=Very much)

Table 2. Question wording for the independent variable

<p><i>Jordan (Survey conducted April 2014)</i></p>	<p>To what extent do you feel satisfied with the way the current government is handling the country's affairs? Are you: 1=Very satisfied, 2=satisfied, 3=dissatisfied, 4=very dissatisfied.</p>
<p><i>Tunisia (Survey conducted February 2015)</i></p>	<p>How confident are you that the new/current government in Tunisia will perform well? 1=Very confident, 2=confident, 3=not very confident, 4=not confident at all.</p>

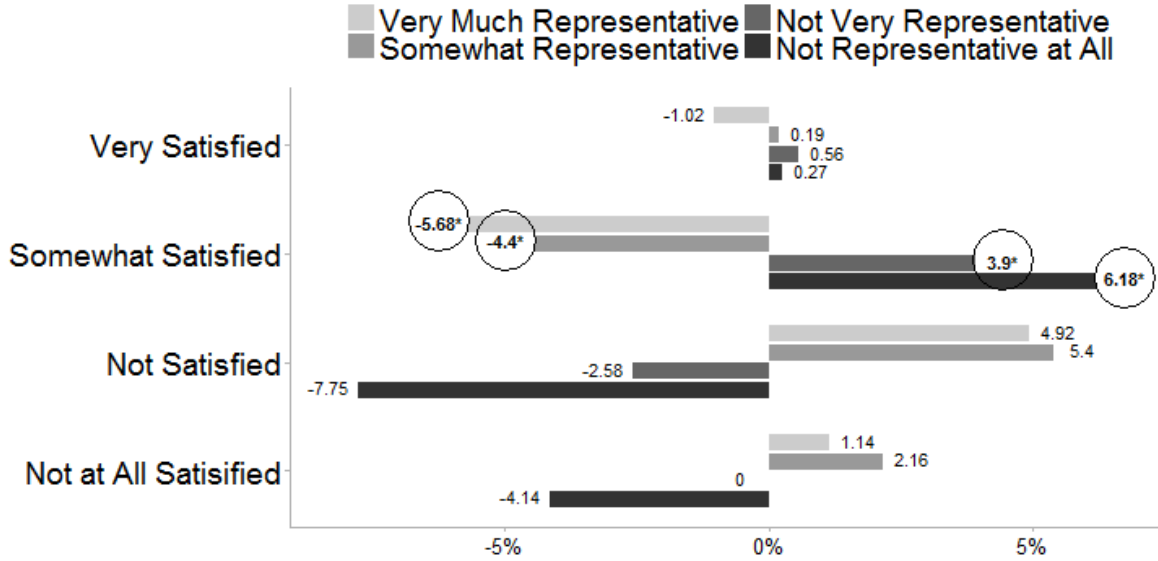
Table 3. Summary of findings

	<i>Homogeneous</i>	<i>Heterogeneous</i>	
	Uniform influence (Hypothesis 1)	Biased assimilation (Hypothesis 2)	Backfire effect (Hypothesis 3)
Jordan	No*	No (Insignificant)	Yes (p<.05)
Tunisia	No*	No (Insignificant)	No (Insignificant)

**Statistically significant in expected direction for some but not most population sub-groups. Theory anticipates more uniform and statistically significant effects.*

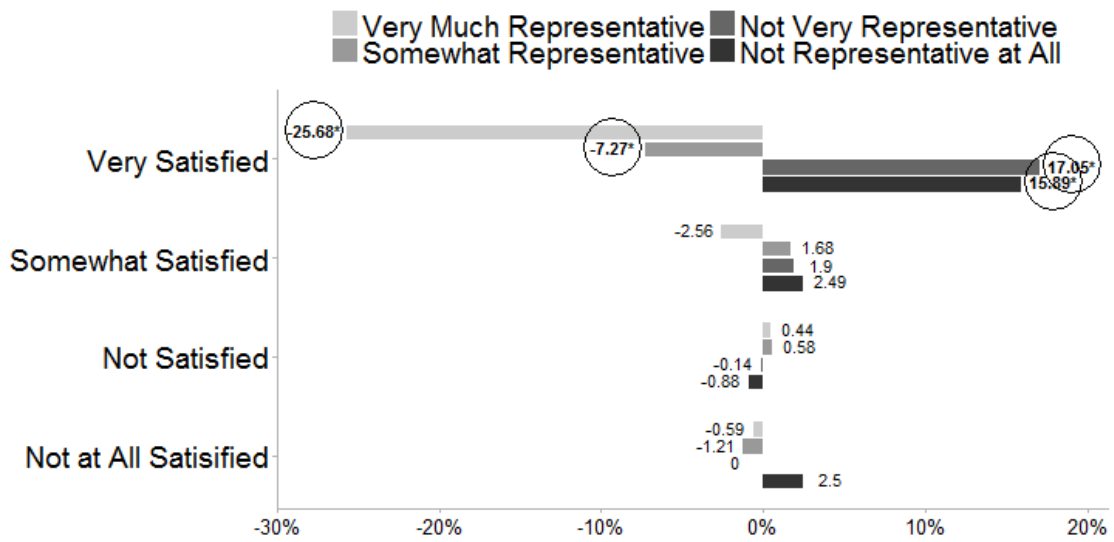
Figures

Figure 1. Jordanians receiving a negative statement, compared to control group



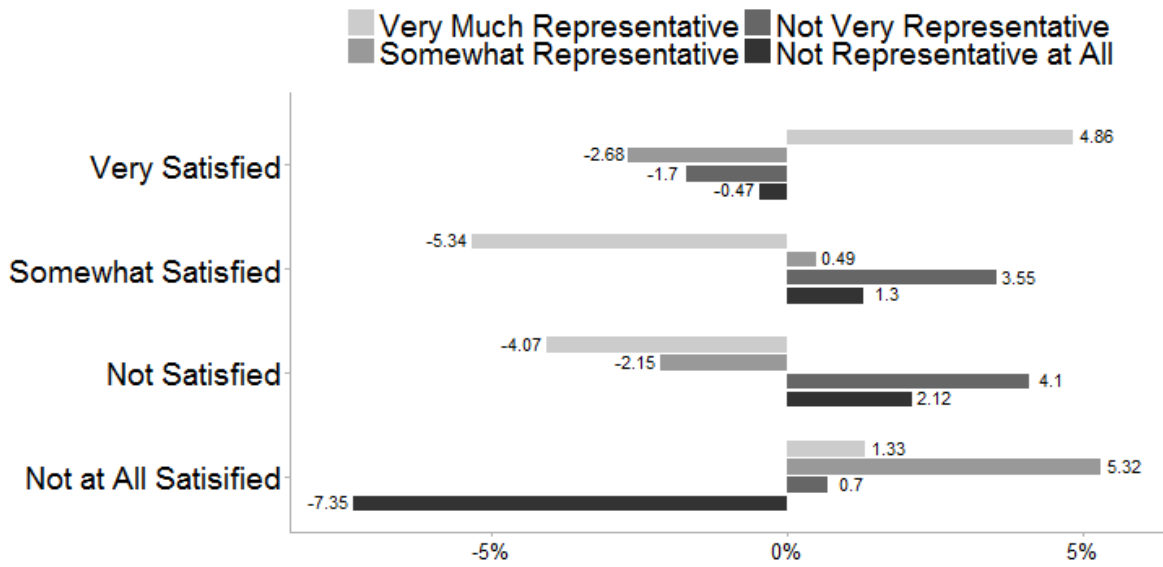
* $p \leq .05$. We find evidence that statements affect those who are somewhat satisfied with the government in the expected direction (see circles). However, since this effect is present for only one group of respondents, it is inconsistent with the uniform effect hypothesis, which expects statements to impact attitudes in the same direction for all population subgroups.

Figure 2. Jordanians receiving positive statement, compared to control group



* $p \leq .05$. Evidence for a backfire effect shown by the circles among those who are very satisfied with the government.

Figure 3. Tunisians receiving negative statement, compared to control group



Online Appendices
Appendix 1: Randomization Checks

A randomized block design was used. All tests showed that randomization was effective. Table A1.1 shows that the 1,260 Tunisian respondents (voters, in the three conditions) were not randomly assigned to the negative, control, and positive conditions ($p < .038$),¹⁶ as a result of an error in the randomization function in the SurveyToGo software that was used in the tablet administration of the survey. However, as shown in Table A1.2, the conditions were randomly distributed across the independent variable and all key demographic variables.

Table A1.1.

Randomized block design: Assignment of respondents to control and experimental conditions (Tunisia)

	Negative	Control	Positive	Total
Bizerte Nord	29(6.9%)	29(6.9%)	28(6.7%)	86(6.8%)
Bizerte Sud	23(5.5%)	19(4.5%)	19(4.5%)	61(4.8%)
Menzel Jemil	18(4.3%)	23(5.4%)	19(4.5%)	60(4.8%)
Siliana Ville	34(8.1%)	16(3.8%)	18(4.3%)	68(5.4%)
Bouarada	25(6.0%)	23(5.4%)	24(5.7%)	72(5.7%)
Gaafour	14(3.4%)	8(1.9%)	8(1.9%)	30(2.4%)
Monastir	31(7.4%)	28(6.6%)	34(8.1%)	93(7.4%)
Tebolba	27(6.5%)	24(5.7%)	33(7.9%)	84(6.7%)
Moknine	21(5.0%)	18(4.3%)	20(4.8%)	59(4.7%)
Souassi	14(3.4%)	35(8.3%)	26(6.2%)	75(6.0%)
Ksour Essaf	19(5.7%)	30(7.1%)	27(6.4%)	76(6.0%)

¹⁶ 3,659 Tunisians participated in the study, however, only 2,090 were assigned to the three conditions analyzed in this paper.

Mahdia	24(5.7%)	23(5.4%)	20(4.8%)	67(5.3%)
Cité Khadhra	34(8.1%)	23(5.4%)	20(4.8%)	77(6.1%)
Jbel Jloud	22(5.3%)	20(4.7%)	31(7.4%)	73(5.8%)
La Marsa	31(7.4%)	19(4.5%)	24(5.7%)	74(5.9%)
Sekiet Eddayer	20(4.8%)	26(6.2%)	21(5.0%)	67(5.3%)
Sfax Sud	21(5.0%)	25(5.9%)	26(6.2%)	72(5.7%)
Sfax Ville	11(2.6%)	34(8.0%)	21(5.0%)	66(5.2%)
Total	418(100.0%)	423(100.0%)	419(100.0%)	1260(100.0%)

Two-tailed χ^2 test show treatments were not randomly distributed across electoral districts ($p < .038$).

Table A1.2 shows that the conditions were randomly distributed across the independent variable (i.e., satisfaction with the government) and key demographic variables, as shown by insignificant chi-square tests.

Table A1.2.

Randomization of treatment and demographic variables (Tunisia)

	Negative	Control	Positive
Gender			
Male	55.5%	56.3%	52.5%
Female	44.5%	43.7%	47.5%
($N=1260/$ Mean=.45/ $Sd=.50$)	$\chi^2(2)=1.339(p<.512)$		
Interviewer gender			
Male	56.5%	56.7%	54.7%
Female	43.5%	43.3%	45.4%
($N=1260/$ Mean=.45/ $Sd=.50$)	$\chi^2(2)=.4360(p<.804)$		

Gender stereotypes: Works harder¹

No difference	68.8%	63.3%	64.2%
Woman	17.5%	20.1%	19.7%
Man	13.7%	16.7%	16.1%
(N=1219/Mean=2.50/Sd=.75)	$\chi^2(2)=3.12(p<.539)$		

Class

Lower	4.1%	5.0%	2.6%
Lower middle	7.7%	7.1%	9.4%
Middle	57.1%	55.8%	59.5%
Upper middle	23.9%	24.1%	19.7%
Upper	7.2%	8.0%	8.9%
(N=1255/Mean=3.23/Sd=.86)	$\chi^2(2)=7.91(p<.442)$		

Education

No formal education	4.5%	4.7%	4.8%
Primary school	16.9%	14.4%	12.9%
Secondary school	29.2%	36.9%	38.0%
Higher education	49.1%	43.7%	44.3%
(N=1238/Mean=3.22/Sd=.87)	$\chi^2(2)=13.18(p<.214)$		

Evaluation of government²

Very good	11.8%	13.4%	13.0%
Good	35.3%	43.8%	41.9%
Bad	39.8%	33.6%	33.3%
Very bad	13.3%	9.3%	11.8%
(N=1219/Mean=2.46/Sd=.86)	$\chi^2(2)=10.416(p<.108)$		

Residence

Rural	4.4%	4.0%	5.0%
Small town	19.6%	22.9%	23.4%
Large town	42.6%	43.0%	39.5%
Suburb or large city	33.5%	30.0%	32.1%
(<i>N</i> =1259/ <i>Mean</i> =.301/ <i>Sd</i> =.85)	$\chi^2(2)=3.704(p<.717)$		
Age			
18-30 years	24.4%	25.3%	24.3%
31-40 years	19.1%	19.9%	20.1%
41-50 years	24.2%	17.5%	21.7%
51-60 years	16.5%	20.8%	16.2%
61-70 years	10.8%	11.4%	11.0%
71 years or more	4.8%	5.0%	6.7%
(<i>N</i> =1260/ <i>Mean</i> =2.87/ <i>Sd</i> =1.51)	$\chi^2(10)=9.920(p<.448)$		

Two-tailed χ^2 test show treatments are randomly distributed across groups.

¹In general, would you say that male or female members work harder to provide services and represent citizens, or would you say there is no difference?

²How confident are you that the new/current government in Tunisia will perform well?

Table A1.3 shows that the 954 Jordanian respondents were randomly assigned to the negative, control, and positive conditions ($p<.100$).

Table A1.3.

Randomized block design: Assignment of respondents to control and experimental conditions (Jordan)

	Negative	No statement	Positive	Total
Amman 1	14(4.6%)	21(5.8%)	24(7.9%)	59(6.2%)
Amman 4	18 (5.9%)	22(6.0%)	16(5.2%)	56(5.9%)
Balqa 2	31(10.2%)	34(9.3%)	25(8.2%)	90(9.4%)
Zarqa 3	23(7.5%)	36(9.9%)	30(9.8%)	89(9.3%)
Irbid 1	31(10.2%)	23(6.3%)	26(8.5%)	80(8.4%)
Irbid 8	29(9.5%)	35(9.6%)	27(8.9%)	91(9.5%)
Jerash	33(10.8%)	40(11.0%)	22(7.2%)	95(10.0%)
Ajloun 2	25(8.2%)	38(10.4%)	24(7.9%)	87(9.1%)
Tafileh 1	27(8.9%)	22(6.0%)	26(8.5%)	75(7.9%)
Tafileh 2	27(8.9%)	31(8.5%)	22(7.2%)	75(7.9%)
Ma'an 2	22(7.2%)	33(9.0%)	33(10.8%)	88(9.2%)
Ma'an 3	25(8.2%)	30(8.2%)	30(9.8%)	85(8.9%)
Total	305(100.0%)	365(100.0%)	305(100.0%)	954(100.0%)

Two-tailed χ^2 test show treatments are randomly distributed across administrative districts ($p < 1.00$).

Table A1.4 shows that the conditions were randomly distributed across the independent variable (i.e., satisfaction with the government) and key demographic variables, as shown by insignificant chi-square tests.

Table A1.4.

Randomization of treatment and demographic variables (Jordan)

	Negative	Control	Positive
Gender			
Male	49.7%	50.7%	47.5%

Female	50.3%	49.4%	52.5%
(N=971/Mean=.51/Sd=.50)	$\chi^2(1)=.637(p<.727)$		

Interviewer gender

Male	90.6%	90.6%	91.4%
Female	9.4%	9.4%	8.6%
(N=971/Mean=.10/Sd=.29)	$\chi^2(1)=.146(p<.930)$		

Better at providing services¹

No difference	47.8%	48.1%	47.5%
Woman	39.0%	34.4%	37.2%
Man	11.6%	12.3%	11.6%
(N=938/Mean=.63/Sd=.70)	$\chi^2(2)=9.580(p<.296)$		

Class

Lower	14.1%	13.3%	11.6%
Lower middle	19.6%	17.5%	23.3%
Middle	54.1%	52.6%	54.2%
<i>Upper middle</i>	2.5%	5.5%	3.3%
Upper	9.7%	10.4%	7.3%
(N=968/Mean=2.76/Sd=1.04)	$\chi^2(1)=15.991(p<.192)$		

Evaluation of government²

Very satisfied	4.8%	6.0%	3.5%
Satisfied	43.9%	38.6%	41.0%
Dissatisfied	29.1%	32.6%	32.4%
Not satisfied at all	22.2%	22.8%	23.1%

(N=939/Mean=2.72/Sd=1.0) $\chi^2(1)=4.045$ (p<.671)

Residence (N/A)

Age

18-25 years	33.1%	29.3%	30.3%
26-39 years	28.3%	30.9%	31.3%
40-59 years	18.4%	17.4%	17.4%
60 years or more	20.1%	22.4%	21.1%

(N=951/Mean=2.29/Sd=1.1) $\chi^2(1)=1.965$ (p<.923)

Two-tailed χ^2 test show treatments are randomly distributed across groups.

¹In general, would you say that male or female deputies are more capable of providing services and representing citizens, or would you say there is no difference?

²To what extent do you feel satisfied with the way the current government is handling the country's affairs? Are you: 1=Very satisfied, 2=satisfied, 3=dissatisfied, 4=very dissatisfied.

Appendix 2. Survey Design and Electoral System

2015 Tunisian Local Governance and Performance Index Survey (LGPI)

The 2015 Local Governance Performance Index (LGPI) survey was developed by Lindsay Benstead (Associate Professor of Political Science at Portland State University), Pierre Landry (Professor of Political Science at NYU-Shanghai) and Ellen Lust (Professor of Political Science at the University of Gothenburg). The survey was implemented in May 2015 as part of the Program on Governance and Local Development at Yale University, with funding from the Moulay Hicham Foundation and Yale University. Dhafer Malouche (*Ecole Supérieure de la Statistique et de l'Analyse de l'Information*) served as the survey manager, working in conjunction with MAZAM.

The purpose of the study was to assess governance and service delivery at the local level. It targeted Tunisian citizens of voting age (18 or older). The survey team was composed of six supervisors and 42 interviewers. After the interviewer training program, the face-to-face multi-topic survey was administered in the field in May 2015 to a nationally representative sample.

A multi-stage sampling design was used to first select six governorates (*Wilayat*) by probability of selection proportional to size (PPS). Within governorates, the “delegation” that is the seat of the governorate was selected as a self-representing unit, while two other delegations were selected at random, also by PPS. We thus obtained a set of 18 municipalities/delegations as secondary sampling units (SSUs). Given the lack of updated census information below the SSU level, the selection of TSUs (defined as square half-arc minutes from a spatial grid) was conducted by gridding each municipality with the lasted nightlight-data from the DMSP-OLS series. The light intensity of each pixel on the remote-sensing image was used as proxy for the relative population density within the municipality. Ten TSUs (and a backup unit) were drawn within each SSU. Finally, the enumerators were sent to 20 randomly selected coordinates within each TSU. Taking those as starting points and using instructions to conduct a random walk, enumerators reached and contacted the corresponding households. Within each household, computer tablets were used to select a final respondent through a “Kish grid” in order to select eligible individuals randomly within households. Interviewers of either gender proceeded to interview the randomly selected respondent, whether male or female and then administered the entire questionnaire using the tablet. This process resulted in 3,559 completed interviews.

Given the larger sample size in Tunisia, the survey experiment included five conditions, varying both the message and the source of information: (1) a negative treatment group, which received a statement made by international observers about the elections having been marred by

problems, (2) a negative, domestic treatment group, which received a similar statement made by domestic observers, (3) a control group, which received no statement, (4) a positive, international treatment group, in which an international organization deemed that safeguards improved the transparency of the elections; and (5) a positive, domestic treatment group, in which a similar statement was made by a domestic organization. This will allow us to examine the impact of domestic vs. international observers, although we will do so in a subsequent paper.

Jordan 2014

The National Dialogue Committee (NDC), established as part of a palace-initiated reform effort in March 2011, recommended eliminating the SNTV majoritarian system and introducing a two-tiered, open-list, proportional representation (PR) system, dividing the kingdom into electoral districts based on governorates and increasing the women's quota. The proposal was to combine an open-list, proportional representation (PR) system at the provincial level (115 seats) with an open-list system at the national level (15 seats), and to raise the number of parliamentarians from 120 to 130.

Jordan remained divided into 45 electoral districts for the 2013 elections, but added a single, national-level district for PR lists. The 45 districts comprise a mix of First-Past-the Post and SNTV rules in single and multi-member constituencies with a total of 108 seats. There are quota seats for Chechens and Circassians (three), Christians (nine), and Bedouins (three in each of three Bedouin areas, for nine total). There were 27 seats in the single, nationwide district. Seats in this district were filled through a largest remainder, closed-list PR system. Women are elected to an additional 15 seats through a quota system at the governorate/Bedouin area level.

There were approximately 7,800 domestic and 400 international observers for the Jordanian elections (Carter Center 2013) and 13,422 domestic and 661 international accredited observers in Tunisia (NDI 2011, p. 17).

The 2014 Jordanian Post-Election Survey followed the country's 2013 parliamentary elections and was developed through collaboration between Ellen Lust (University of Gothenburg), Lindsay Benstead (Portland State University), and Kristen Kao (University of Gothenburg). Local partnership with an experienced survey implementation firm, Middle East Marketing Consultants led by Tony Sabbagh, facilitated the translation of the questionnaire into the local dialect, the creation of a complex sample design suitable for the needs of the study, the recruitment and training of fifty enumerators and supervisors, and the careful implementation of the survey in the field. Data collection was carried out 21 April 2014–28 April 28 2014 employing tablet computers in face-to-face household interviews. The enumerators and supervisors¹⁷ were trained for two days prior to the implementation of the survey, after which they were sent out into the field in teams of five (four enumerators to each supervisor). Every attempt was made to have data uploaded to the main database in Amman every evening over Internet, and Kristen Kao analyzed the results each evening to check for errors or inconsistencies.

The purpose of the study was to assess political attitudes and behaviors following parliamentary elections in 2013 among the target population of Jordanian citizens of voting age (18) or older. To ensure sampling of adequate numbers of respondents eligible to vote in either

¹⁷ In the dataset, the first ten surveyors (variable svyr) are actually supervisors whose tablets were only employed for practice or in times of emergency should another tablet fail in the field.

multi-member or single-member electoral districts as well as adequate representation of rural,¹⁸ tribal populations under-counted in conventional probability-proportional-to-size sampling, a purposive multistage stratified sampling design was used with electoral district as the primary sampling unit. Jordan's 45 electoral districts were stratified by region (north, central, and south),¹⁹ size (small and large), and district type (multi-member single non-transferable vote (SNTV) or single member plurality).²⁰ Twelve districts were selected, within which 25

¹⁸ Defining the terms "rural" versus "urban" is a subjective process. Population density statistics at the electoral district level for Jordan were unobtainable, so two measures were constructed from available eligible voter data try to capture the differences between urban and rural electoral districts.

¹⁹ Some scholars note that the culture in the south of Jordan is more akin to that of the Arab Gulf region, while the culture in the north is more akin to that of the rest of the Levantine region. In some instances, this means that tribes might be either more or less sedentary and/or reliant on agriculture versus livestock for their livelihoods. In the modern era, this distinction should not greatly affect the results of this survey or conclusions drawn from it concerning the current state of politics in Jordan. There are other notable differences between the different areas of Jordan. The sample is stratified by region to make sure that these differences are represented.

²⁰ The variable "Sntv" captures the dividing line between multi-member versus single member districts, the former of which are run under an SNTV electoral system (coded as a "1" in the data) versus the latter, which are run under a single-member district plurality system (coded as a "0" in the data).

households were randomly selected from blocks enumerated in the 2004 census. (See Table A2.1).

Once in the field, interviewers were instructed to stratify selection of respondents to obtain approximately equal numbers of male and female respondents, and to select newer buildings in all replacement interviews in an attempt to gain representation of respondents who live in buildings constructed after the 2004 census. Kish tables were used to select one eligible individual within each household at random. Due to the sampling design, results are likely to show clustering in responses. Attempting to weight the data based on inferences implying national representativeness are not recommended given the purposive sampling design, however data analyses should take stratification variables into account.

Interviewers recorded detailed sampling and refusal information on cover sheets completed for up to two visits to each residence. Incomplete surveys in the dataset are the result of a variety of issues. Refusals, in which either the person answering the door or the participant selected by the Kish table refused to participate, make up one type of incomplete survey in the dataset. A survey in which the participant decided to stop participating halfway through the questionnaire constitutes another example of an incomplete interview. Towards the end of fielding, it became apparent that some of the surveys took 20 minutes or less for enumerators to complete. The researchers conducting the project and the local partner concluded that these

surveys were too short to be considered to be realistic, thus part of the final two days of the survey was spent replacing these surveys.²¹

A response rate of 79.8% is based on a total of 1,879 residences visited and 1,499 completed interviews. Among these visits, 7.3% interviews are incomplete due to empty, closed, or non-residential units; ineligible respondents; or other reasons. Another 13.6% are incomplete because respondents refused to participate in the survey. Finally, 12% of the interviews were marked incomplete because they fell short of the 20-minute threshold established to verify that the interviewer did not rush through the interview, so as to call the validity of the interview into question.

Table A2.1.

Descriptive statistics for the electoral districts in the survey (Jordan)

District	Region	Sntv	Seats	N Sample	% Sample
Ma'an 2	South	1	1	119	7.9
Tafileh 2	South	1	1	118	7.9

²¹ If the interview took less than 20 minutes, it is marked as incomplete under the variable “complete.” Analyses of this dataset should be carried out keeping these issues in mind, dropping these surveys from analyses where appropriate.

Ma'an 3	South	1	1	127	8.5
Tafileh 1	South	0	3	136	9.1
Zarqa 3	Central	1	1	127	8.5
Balqa 2	Central	1	1	124	8.3
Amman 4	Central	0	3	132	8.8
Amman 1	Central	0	5	114	7.6
Irbid 1	North	0	5	124	8.3
Ajloun 2	North	1	1	124	8.3
Irbid 8	North	1	1	128	8.5
Jerash	North	0	4	125	8.3

Appendix 3: Inspection of Data to Develop Coding for Dependent Variable

Following the procedure recommended by Long and Freese (2004), a visual inspection of the distributions, informed by the logic of voters' possible thought process, was combined with

modeling the outcomes as a multinomial logistic model. Each outcome was tested on a pairwise

basis for differences with only treatment to inform recoding of the 10-point Likert scale into a 4-point scale. This simplification eases interpretation. The multinomial logit results suggest that the

extreme ends of the scale are different from the each of the next three levels, and that there is a

break in the middle of the scale, with not much difference between the 5 and 6 category. Long

and Freese suggest collapsing categories informed both by the multinomial logit results as well

as theory and practical considerations, so while separating out the extreme ends of the scale and

grouping all levels (2-5, and 6-9), there would be insufficient numbers of cases in the extreme

categories, so levels 1 and 2 in the 10-point scale are coded as "1" while 9 and 10 are coded as

"4." This choice of coding makes it less likely that differences will be found between categories

and thus biases the results against our theoretical expectations, so the evidentiary burden is higher with this conservative coding. Histograms for the recoded scales follow.

Distributions by treatment group (Tunisia)

	CONTROL	NEGATIVE	POSITIVE																																																																								
VOTERS	<p>T2015 CTRL Voters' Evaluation of 2014 Parliamentary Elections</p> <table border="1"> <caption>Distribution for control condition</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>20</td></tr> <tr><td>1</td><td>22</td></tr> <tr><td>2</td><td>14</td></tr> <tr><td>3</td><td>15</td></tr> <tr><td>4</td><td>71</td></tr> <tr><td>5</td><td>53</td></tr> <tr><td>6</td><td>55</td></tr> <tr><td>7</td><td>57</td></tr> <tr><td>8</td><td>43</td></tr> <tr><td>9</td><td>0</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	20	1	22	2	14	3	15	4	71	5	53	6	55	7	57	8	43	9	0	10	0	<p>T2015 NEG Voters' Evaluation of 2014 Parliamentary Elections</p> <table border="1"> <caption>Distribution for negative frame</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>29</td></tr> <tr><td>1</td><td>18</td></tr> <tr><td>2</td><td>21</td></tr> <tr><td>3</td><td>21</td></tr> <tr><td>4</td><td>57</td></tr> <tr><td>5</td><td>71</td></tr> <tr><td>6</td><td>44</td></tr> <tr><td>7</td><td>61</td></tr> <tr><td>8</td><td>46</td></tr> <tr><td>9</td><td>55</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	29	1	18	2	21	3	21	4	57	5	71	6	44	7	61	8	46	9	55	10	0	<p>T2015 POS Voters' Evaluation of 2014 Parliamentary Elections</p> <table border="1"> <caption>Distribution for positive frame</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>28</td></tr> <tr><td>1</td><td>19</td></tr> <tr><td>2</td><td>20</td></tr> <tr><td>3</td><td>17</td></tr> <tr><td>4</td><td>60</td></tr> <tr><td>5</td><td>53</td></tr> <tr><td>6</td><td>70</td></tr> <tr><td>7</td><td>62</td></tr> <tr><td>8</td><td>33</td></tr> <tr><td>9</td><td>50</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	28	1	19	2	20	3	17	4	60	5	53	6	70	7	62	8	33	9	50	10	0
Rating	Unweighted Count																																																																										
0	20																																																																										
1	22																																																																										
2	14																																																																										
3	15																																																																										
4	71																																																																										
5	53																																																																										
6	55																																																																										
7	57																																																																										
8	43																																																																										
9	0																																																																										
10	0																																																																										
Rating	Unweighted Count																																																																										
0	29																																																																										
1	18																																																																										
2	21																																																																										
3	21																																																																										
4	57																																																																										
5	71																																																																										
6	44																																																																										
7	61																																																																										
8	46																																																																										
9	55																																																																										
10	0																																																																										
Rating	Unweighted Count																																																																										
0	28																																																																										
1	19																																																																										
2	20																																																																										
3	17																																																																										
4	60																																																																										
5	53																																																																										
6	70																																																																										
7	62																																																																										
8	33																																																																										
9	50																																																																										
10	0																																																																										
NON VOTERS	<p>T2015 CTRL Nonvoters' Evaluation of 2014 Parliamentary Elections</p> <table border="1"> <caption>Distribution for control condition</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>43</td></tr> <tr><td>1</td><td>22</td></tr> <tr><td>2</td><td>16</td></tr> <tr><td>3</td><td>22</td></tr> <tr><td>4</td><td>47</td></tr> <tr><td>5</td><td>37</td></tr> <tr><td>6</td><td>22</td></tr> <tr><td>7</td><td>15</td></tr> <tr><td>8</td><td>17</td></tr> <tr><td>9</td><td>22</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	43	1	22	2	16	3	22	4	47	5	37	6	22	7	15	8	17	9	22	10	0	<p>T2015 NEG Nonvoters' Evaluation of 2014 Parliamentary Elections</p> <table border="1"> <caption>Distribution for negative frame</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>27</td></tr> <tr><td>1</td><td>19</td></tr> <tr><td>2</td><td>18</td></tr> <tr><td>3</td><td>20</td></tr> <tr><td>4</td><td>55</td></tr> <tr><td>5</td><td>30</td></tr> <tr><td>6</td><td>31</td></tr> <tr><td>7</td><td>18</td></tr> <tr><td>8</td><td>16</td></tr> <tr><td>9</td><td>18</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	27	1	19	2	18	3	20	4	55	5	30	6	31	7	18	8	16	9	18	10	0	<p>T2015 POS Nonvoters' Evaluation of 2014 Parliamentary Elections</p> <table border="1"> <caption>Distribution for positive frame</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>49</td></tr> <tr><td>1</td><td>19</td></tr> <tr><td>2</td><td>18</td></tr> <tr><td>3</td><td>20</td></tr> <tr><td>4</td><td>31</td></tr> <tr><td>5</td><td>23</td></tr> <tr><td>6</td><td>36</td></tr> <tr><td>7</td><td>26</td></tr> <tr><td>8</td><td>12</td></tr> <tr><td>9</td><td>14</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	49	1	19	2	18	3	20	4	31	5	23	6	36	7	26	8	12	9	14	10	0
Rating	Unweighted Count																																																																										
0	43																																																																										
1	22																																																																										
2	16																																																																										
3	22																																																																										
4	47																																																																										
5	37																																																																										
6	22																																																																										
7	15																																																																										
8	17																																																																										
9	22																																																																										
10	0																																																																										
Rating	Unweighted Count																																																																										
0	27																																																																										
1	19																																																																										
2	18																																																																										
3	20																																																																										
4	55																																																																										
5	30																																																																										
6	31																																																																										
7	18																																																																										
8	16																																																																										
9	18																																																																										
10	0																																																																										
Rating	Unweighted Count																																																																										
0	49																																																																										
1	19																																																																										
2	18																																																																										
3	20																																																																										
4	31																																																																										
5	23																																																																										
6	36																																																																										
7	26																																																																										
8	12																																																																										
9	14																																																																										
10	0																																																																										

Distributions by treatment group (Jordan)

	CONTROL	NEGATIVE	POSITIVE																																																																								
VOTERS	<p>J2014 CTRL Voters' Evaluation of 2013 Parliamentary Election</p> <table border="1"> <caption>Distribution for control condition</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>65</td></tr> <tr><td>1</td><td>22</td></tr> <tr><td>2</td><td>28</td></tr> <tr><td>3</td><td>33</td></tr> <tr><td>4</td><td>52</td></tr> <tr><td>5</td><td>41</td></tr> <tr><td>6</td><td>27</td></tr> <tr><td>7</td><td>37</td></tr> <tr><td>8</td><td>12</td></tr> <tr><td>9</td><td>39</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	65	1	22	2	28	3	33	4	52	5	41	6	27	7	37	8	12	9	39	10	0	<p>J2014 NEG Voters' Eval of 2013 Parl Elect</p> <table border="1"> <caption>Distribution for negative frame</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>39</td></tr> <tr><td>1</td><td>22</td></tr> <tr><td>2</td><td>32</td></tr> <tr><td>3</td><td>41</td></tr> <tr><td>4</td><td>48</td></tr> <tr><td>5</td><td>35</td></tr> <tr><td>6</td><td>20</td></tr> <tr><td>7</td><td>21</td></tr> <tr><td>8</td><td>23</td></tr> <tr><td>9</td><td>18</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	39	1	22	2	32	3	41	4	48	5	35	6	20	7	21	8	23	9	18	10	0	<p>J2014 POS Voters' Eval of 2013 Parl Elect</p> <table border="1"> <caption>Distribution for positive frame</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>51</td></tr> <tr><td>1</td><td>37</td></tr> <tr><td>2</td><td>21</td></tr> <tr><td>3</td><td>29</td></tr> <tr><td>4</td><td>47</td></tr> <tr><td>5</td><td>22</td></tr> <tr><td>6</td><td>27</td></tr> <tr><td>7</td><td>28</td></tr> <tr><td>8</td><td>18</td></tr> <tr><td>9</td><td>24</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	51	1	37	2	21	3	29	4	47	5	22	6	27	7	28	8	18	9	24	10	0
Rating	Unweighted Count																																																																										
0	65																																																																										
1	22																																																																										
2	28																																																																										
3	33																																																																										
4	52																																																																										
5	41																																																																										
6	27																																																																										
7	37																																																																										
8	12																																																																										
9	39																																																																										
10	0																																																																										
Rating	Unweighted Count																																																																										
0	39																																																																										
1	22																																																																										
2	32																																																																										
3	41																																																																										
4	48																																																																										
5	35																																																																										
6	20																																																																										
7	21																																																																										
8	23																																																																										
9	18																																																																										
10	0																																																																										
Rating	Unweighted Count																																																																										
0	51																																																																										
1	37																																																																										
2	21																																																																										
3	29																																																																										
4	47																																																																										
5	22																																																																										
6	27																																																																										
7	28																																																																										
8	18																																																																										
9	24																																																																										
10	0																																																																										
NON VOTERS	<p>NOT ASKED</p>	<p>J2014 NEG Nonvoters' Eval of 2013 Parl Elect</p> <table border="1"> <caption>Distribution for negative frame</caption> <thead> <tr> <th>Rating</th> <th>Unweighted Count</th> </tr> </thead> <tbody> <tr><td>0</td><td>128</td></tr> <tr><td>1</td><td>51</td></tr> <tr><td>2</td><td>46</td></tr> <tr><td>3</td><td>50</td></tr> <tr><td>4</td><td>67</td></tr> <tr><td>5</td><td>42</td></tr> <tr><td>6</td><td>31</td></tr> <tr><td>7</td><td>36</td></tr> <tr><td>8</td><td>22</td></tr> <tr><td>9</td><td>29</td></tr> <tr><td>10</td><td>0</td></tr> </tbody> </table>	Rating	Unweighted Count	0	128	1	51	2	46	3	50	4	67	5	42	6	31	7	36	8	22	9	29	10	0	<p>NOT ASKED</p>																																																
Rating	Unweighted Count																																																																										
0	128																																																																										
1	51																																																																										
2	46																																																																										
3	50																																																																										
4	67																																																																										
5	42																																																																										
6	31																																																																										
7	36																																																																										
8	22																																																																										
9	29																																																																										
10	0																																																																										

Appendix 4. Homogenous and Heterogeneous Treatment Effects

Table A4.1.

Homogeneous effects

	Tunisia		Jordan	
	(1)	(2)	(3)	(4)
Negative Frame	-0.396	-0.170	-0.0263	-0.00245
	(0.293)	(0.220)	(0.151)	(0.126)
Positive Frame	-.0588	-0.0342	-0.190	-0.189
	(0.194)	(0.211)	(0.130)	(0.153)
Not Satisfied		1.387***		0.686***
		(0.200)		(0.192)
Satisfied		1.935***		1.040**
		(0.292)		(0.243)
Very Satisfied		2.802***		1.773***
		(0.377)		(0.274)
Gender		0.199		0.381*
		(0.157)		(0.131)
Age		-0.0053		-0.00237
		(0.0061)		(0.00545)
Secondary		-0.196		0.063
		(0.514)		(0.126)
Technical/ Professional				-0.218
				(0.160)
University +		-0.130		0.020
		(0.330)		(0.135)
Cut 1	-2.303***	-1.283**	-1.187***	-0.631
	(0.257)	(0.402)	(0.0997)	(0.415)
Cut 2	-1.039***	0.471	0.301*	0.962
	(0.124)	(0.422)	(0.111)	(0.455)

Cut 3	1.177** (0.321)	2.630*** (0.450)	1.751*** (0.158)	2.460*** (0.465)
Observations	2045	1985	1478	1455

Standard errors in parentheses; *p < .05, **p < .01, ***p < 0.001

Table A4.2 shows comparisons of confidence in the government across treatment groups, the base category is the interaction between the respondent being very unsatisfied with the government and the frame.

Table A4.3.

Ordered logistic regressions showing the effect of the treatments across groups with different a priori attitudes about the government

	Tunisia	Jordan
Control x Not Satisfied	1.928** (0.609)	0.533** (0.166)
Control x Satisfied	2.463*** (0.502)	1.251*** (0.232)
Control x Very Satisfied	3.031*** (0.592)	2.319*** (0.420)
Negative x Not Satisfied	1.652** (0.539)	0.970*** (0.217)
Negative x Satisfied	2.193*** (0.525)	0.810** (0.206)
Negative x Very Satisfied	3.228*** (0.536)	2.273*** (0.452)
Positive x Not Satisfied	1.759** (0.517)	0.578† (0.280)
Positive x Satisfied	2.335** (0.600)	1.079*** (0.226)

Positive x Very Satisfied	3.287*** (0.706)	0.894† (0.497)
Gender	0.206 (0.185)	0.410** (0.122)
Age	-0.00493 (0.0062)	-0.000098 (0.0048)
Secondary School	-0.158 (0.301)	0.0341 (0.175)
Technical School		0.245 (0.243)
University & Higher	-0.100 (0.334)	0.0248 (0.207)
<hr/>		
Cut 1	0.760 (0.669)	0.243 (0.240)
Cut 2	1.000 (0.740)	1.362*** (0.267)
Cut 3	3.161** (0.805)	2.875*** (0.279)
Observations	1985	1455

Standard Errors in Parenthesis

†p<0.10 *p<0.05 **p<0.01 ***p<0.001

Table A4.3.

Predicted probabilities for the effect of the treatments across groups with different a priori attitudes about the government

	Predicted Probabilities of Election Quality Given Interaction of Government Satisfaction and Election Monitor Frame <i>(To what extent does most recent parliamentary election reflect will of the people?)</i>											
Perception of Current	Not at all			Not very much			Somewhat			Very much		
	NEG	CTRL	POS	NEG	CTRL	POS	NEG	CTRL	POS	NEG	CTRL	POS
Very Negative												
Jordan 2014	0.360	0.402	0.427	0.373	0.365	0.358	0.192	0.170	0.158	0.075	0.063	0.058
Tunisia 2015	0.284	0.358	0.228	0.412	0.405	0.402	0.256	0.202	0.306	0.048	0.035	0.064
Negative												
Jordan 2014	0.197	0.275	0.266	0.349	0.375	0.373	0.297	0.243	0.249	0.157	0.107	0.112
Tunisia 2015	0.095	0.074	0.086	0.283	0.242	0.267	0.462	0.483	0.472	0.161	0.201	0.175
Positive												
Jordan 2014	0.219	0.157	0.181	0.360	0.320	0.339	0.282	0.326	0.309	0.140	0.197	0.171
Tunisia 2015	0.057	0.044	0.050	0.203	0.168	0.184	0.492	0.487	0.491	0.248	0.301	0.275
Very Positive												
Jordan 2014	0.068	0.065	0.224	0.197	0.192	0.362	0.352	0.350	0.278	0.383	0.393	0.136
Tunisia 2015	0.022	0.027	0.021	0.093	0.111	0.089	0.414	0.441	0.405	0.471	0.422	0.485
Note: Probabilities are calculated following ordered logistic regression using SPost13 package over subgroups of respondents falling into each category of satisfaction with current government, with means of other control variables allowed to vary by subgroup (Long and Freese 2004). Chart does not indicate statistical significance.												

Figure A1.1.

Tunisians receiving positive statement, compared to control group

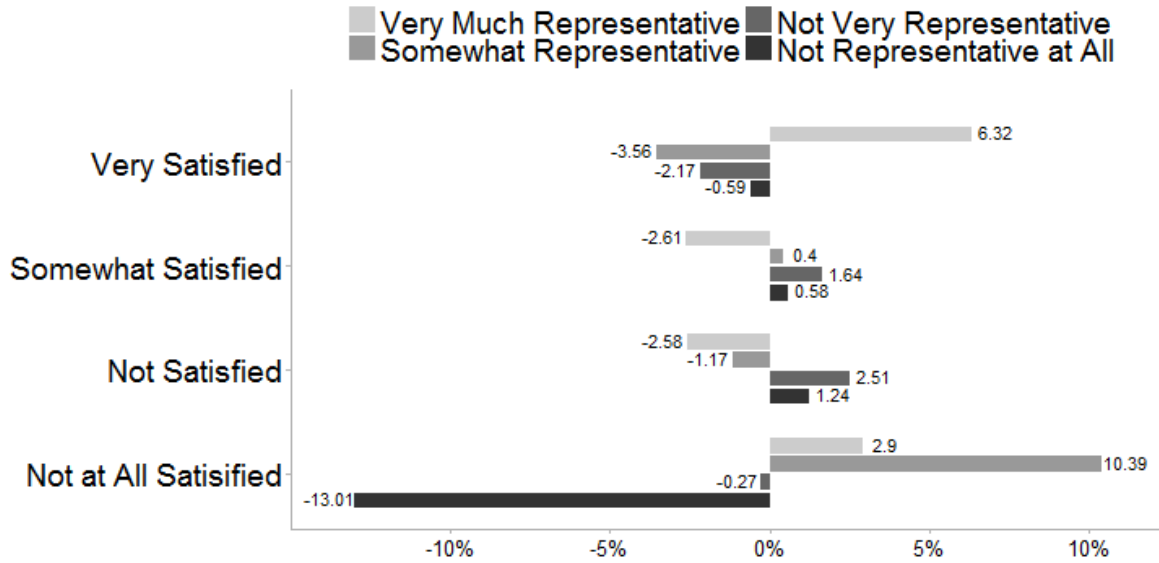


Figure A1.1 shows no significant effects for any population sub-group in Tunisia.