THE OBSERVER EFFECT IN INTERNATIONAL POLITICS Evidence from a Natural Experiment

Susan D. Hyde

World Politics, Volume 60, Number 1, October 2007, pp. 37-63 (Article)

Published by The Johns Hopkins University Press
DOI: 10.1353/wp.0.0001

For additional information about this article
http://muse.jhu.edu/journals/wp/summary/v060/60.1.hyde.html
THE OBSERVER EFFECT IN INTERNATIONAL POLITICS
Evidence from a Natural Experiment

By SUSAN D. HYDE*

INTERNATIONAL actors now play a prominent role in domestic elections and other democratic processes throughout the developing world. They pressure governments to hold democratic elections and they directly engage in the electoral process through provision of technical assistance and funding or by sending teams of observers to monitor elections. International influences on democratization are receiving a small but growing amount of scholarly attention, and international dimensions are increasingly accepted as relevant variables in theories of political transition.¹ Most recently, in a study of the international diffusion of democratization, Kristian Gleditsch and Michael Ward “firmly reject the idea that institutional change is driven entirely by domestic


World Politics 60 (October 2007), 37–63
processes and unaffected by regional and international events.” Although scholars have documented several macrolevel relationships between international-level variables and movement toward democracy, there has been little attention paid to the microlevel effects of international involvement in the democratization process. How do international actors influence democratization in developing countries?

In examining this question, I focus on one form of democracy promotion: international monitoring of elections. Until 1962 there had been no recorded cases of international election observation in sovereign states. By 2004 upwards of 80 percent of elections held in nonconsolidated democracies were monitored, and any leader of a developing country wishing to hold a legitimate election was expected to invite international election observers. Although the record of election observation demonstrates that observers grew willing to condemn fraudulent elections over the course of the 1990s, it remains unknown whether international monitors can actually bring about cleaner elections, as proponents of election monitoring assert.

Any cross-national study attempting to examine the domestic effects of international observers would be plagued by endogeneity problems. At the aggregate level it would be difficult to distinguish between an election that was clean because of the presence of international observers and an election that would have been clean regardless of their presence. One may use counterfactuals to make a persuasive argument, but demonstrating causality using cross-national evidence would be nearly impossible. This is similar to the problem faced by scholars attempting to assess the independent effect of political institutions, and it underscores the uncertainty of efforts to promote democracy abroad. As Adam Przeworski argues, the important empirical question underlying democracy promotion is “whether one can stick any institutions into some particular conditions and expect that they would function in the same way as they have functioned elsewhere.” It is not clear from existing evidence that democracy-promotion efforts encourage democratization independent of what would have occurred in the absence of international involvement.

This article offers a solution to this problem of causal inference by examining the effects of democracy-promotion efforts at the microlevel.

---

2 Gleditsch and Ward (fn. 1).
In order to examine the causal effect of international observers (who represent a prominent element of those efforts), I exploit a natural experiment in which observers were assigned to polling stations using a method that approximates randomization. The natural experiment provides a direct test of whether international observers influence the behavior of domestic political actors on election day. If the presence of international observers causes a reduction in election-day fraud, the effect of observers should be visible at the subnational level by comparing polling stations that were visited by observers with those that were not visited. More specifically, if international monitoring reduces election-day fraud directly, all else held equal, the cheating parties should gain less of their ill-gotten vote share in polling stations that were visited by international monitors. Using polling-station-level election results from the 2003 presidential elections in Armenia, the data show that in this case international observers caused a reduction in the vote share of the incumbent candidate who was widely expected to steal the election. Thus, by examining microlevel electoral data and a direct measure of international presence, this article is able to test for the effects of international scrutiny in a way that country-level studies are unable to do because of endogeneity problems.

Why might scholars of international relations and comparative politics care about the effects of international election monitors? Elections are obviously central in theories of democracy and democratic governance. As William Riker writes, “The essential democratic institution is the ballot box and all that goes with it.” When elections are corrupted, voting as an “instrument of democracy” fails, the “institutionalized uncertainty” of democracy is lessened, and in extreme forms, elections no longer function as an effective method of representative government.

Corrupt elections are not unique to the developing world or the third wave of democratization. However, as the formal institutions of democracy have spread throughout the world, issues of election quality have
become increasingly important for democracy activists and scholars. Fraudulent elections violate the principle of political equality argued to be fundamental to democratic governance, and they reduce representation and accountability, two of the most widely touted benefits of democracy. Thus, the efficacy of international efforts to bring about clean elections in countries that do not already experience them should be of interest to students of democratization.

Recent theoretical and empirical research also points to international observers as potentially important actors in electoral revolutions. James Fearon argues that evaluations of elections by international observers may help facilitate popular enforcement of democracy by improving the probability of popular coordination in defense of democracy. By providing a possible solution to coordination problems, the reports of international observers may encourage the collective uprising of voters following a fraudulent election and may have the potential to lead to alterations in power even when elections are stolen.

Additionally, findings about the consequences of international monitoring are relevant to debates about the effects of international institutions, particularly the call for research as to how international institutions matter, and about the consequences of international pressure within the broader research agendas on the second-image reversed and two-level games. Sovereign leaders of developing countries willingly invite international monitors to judge their elections. If election observation is costless to the state leaders who invite them, this action is easy to explain as a form of cheap talk. However, if election observers cause (cheating) incumbent leaders to gain fewer votes, the puzzle of why leaders invite observers is more interesting and speaks to the impact of international pressure on the behavior of state leaders.

Finally, policymakers also care about the effects of international observers. There is growing skepticism among practitioners that democracy-promotion activities, including election monitoring, are having the effects intended by donors and democracy activists. As Carlos Santiso

---

writes, “The initial enthusiasm within the international donor community is giving way to increasing skepticism and even frustration with the pace and depth of democratic transitions.” However, part of the perceived problem of democracy assistance is likely due to the fact that the field is, as Thomas Carothers has described it, “understudied and poorly understood.” Macrolevel analyses of the relationship between democracy scores and levels of democracy aid cannot reveal which programs work and which do not, nor whether efforts by international actors to encourage democratization have the intended effects. Evaluation of the effects of democracy-promotion activities, such as that presented here, can aid in designing more effective programs, while also contributing to academic understandings of international influences on domestic politics.

This article first presents a theory of how international observers influence election fraud and then provides background information on politics in Armenia, details surrounding the 2003 presidential election, and a description of the election-observation mission. It then describes the data, the use of natural experimental research design, and the test of whether international observers reduce the rate of election-day fraud. Finally, the article presents the results indicating that, in the case of the 2003 presidential elections in Armenia, international observers reduced election-day fraud by about 6 percent in the polling stations they visited during the first round of the election and that the presence of international observers had statistically significant but nuanced effects in the second round.

Observations and Election Fraud

Understanding how international observers influence election fraud requires paying attention to the incentives of those leaders who invite international observers and cheat in front of them. Today’s autocrats who choose to hold elections face a conflict between pressure to hold internationally certified elections and a desire to guarantee their hold on power. Although some formerly autocratic countries now allow democratic elections with broad participation on a level playing field, a number of leaders still try to have it both ways, either betting that they will not be caught cheating or attempting to conceal manipulation.

of the election from international observers. Theoretically, observers should increase constraints on these leaders by increasing the costs of committing election fraud, thus making it marginally less likely that elections will be stolen outright. It is unlikely that election fraud can be eliminated. But by making outright theft of elections more costly to carry out and more likely to be criticized, international election observers may improve the quality of electoral governance and contribute to democratization.

How might observers influence election fraud? One possibility is that observers have no effect on fraud. There are leaders, such as Robert Mugabe of Zimbabwe, who invite international observers and cheat openly. Other leaders, who behave as though they fear being caught by observers, engage in types of electoral manipulation that seem less likely to be caught by international observers. An unlikely possibility is that inviting observers eliminates election fraud. Were this the case, international observers would never witness fraud, but the record of election observation clearly demonstrates that this is not so. Observers have encountered various types of electoral manipulation, including military intimidation of voters, ballot-box stuffing, improper attempts to influence voters inside the voting booth, vote-buying schemes, intentional inflation of the vote tallies, jailing of opposition voters, failure to distribute ballots to opposition strongholds, and manipulation of voter-registration lists.

The most intuitively plausible theory is that the effects of observers are not uniform across elections. In some cases, observers have no effect, while in others they may increase the risks associated with election manipulation, thus decreasing the likelihood that domestic actors will attempt manipulation. Andreas Schedler outlines a number of possible methods of electoral manipulation, ranging from direct election-day electoral fraud such as ballot-box stuffing, to longer-term forms of electoral bias such as gerrymandering. In theory, international election observers or other monitors could have a deterrent effect on any form of election fraud, but testing for this type of effect is difficult. In order to gain empirical traction, this article limits the focus to the manipulation that occurs in and around the polling station on election day. Even

---

17 Bjornlund (fn. 3); Carothers (fn. 4); Carothers (fn. 15).
18 The statement is made based on the reading of numerous international observer reports citing firsthand evidence of election fraud.
within this subcategory of electoral manipulation, election-day fraud takes many forms, all of which share the characteristic of increasing the vote share for the fraud-sponsoring candidate(s). It is my contention that the presence of international election observers may directly influence the behavior of actors inside and in the immediate vicinity of polling stations. If officials at polling stations and other domestic actors feared being caught manipulating the election, they would be less likely to engage in manipulation in the physical presence of international observers. I thus argue that if election-day fraud is occurring and if international observers reduce election fraud, then the presence of observers in a polling station should have a measurable effect on the average vote share of the fraud-sponsoring candidate(s).

The following section provides background on the political climate in Armenia leading up to the 2003 election; this should aid in evaluating the natural experimental research design.

POST-SOVIET POLITICS IN ARMENIA

Between independence in 1991 and the 2003 presidential election, Armenia held six elections. Political participation by voters and most candidates generally complied with democratic standards, but the executive office participated in elections in a manner that has been labeled “flagrantly undemocratic.” Following independence, the elected president Levon Ter-Petrossian and his supporters successfully consolidated power within the executive office while other nascent parties were still attempting to gain organizational strength. As a result, the president is generally the controlling force in Armenian politics, by virtue of his authority to dissolve parliament, appoint all judges, and declare martial law. Strong political parties did not develop as a challenge to executive power, in part because of Ter-Petrossian’s overt efforts to prevent any such opposition party from organizing. As of 2003 there were more than one hundred registered political parties. Because successful election-day manipulation requires at least minimal organizational capacity, the political structure in Armenia points to the incumbent executive as the political actor with the preponderant ability to commit widespread fraud.

21 Ibid.
23 Bremmer and Welt (fn. 20).
The two most prominent political figures in the postindependence period are Ter-Petrossian and Robert Kocharian. The former was president until 1998, when he resigned amid wide public dissatisfaction over his failure to increase the standard of living and his willingness to negotiate with Azerbaijan over the territorial conflict in Nagorno-Karabakh. The latter, who was elected to replace Ter-Petrossian in 1998, was the incumbent candidate in the 2003 presidential election.

The 2003 elections were viewed as a potential turning point for Armenian democracy. As an OSCE/ODIHR official report states:

The election provided an important test of the progress of democratic practices in Armenia, since the previous presidential elections were characterized by serious flaws and generally failed to meet international standards. Issues of concern at the two previous presidential elections . . . included inaccuracy of voter lists, shortcomings in the election administration, media bias, abuse of State resources, flawed voting by military personnel, the presence of unauthorized persons during polling and counting and discrepancies in the vote count.24

What variables correlate with political divisions in Armenian politics? There are few demographic or other variables, such as ethnic divisions, partisan registration, or support for particular issues, that can be used as predictors of political support. Unlike a number of other post-Soviet states, ethnic divisions play little role in internal Armenian politics. The parties do not have ethnic affiliations, as 96 percent of residents are ethnic Armenian, a homogeneity that simplifies the analysis of the patterns of vote share for the candidates.

The only issue during the 2003 elections was the presidential race, with nine candidates on the ballot. The incumbent president Robert Kocharian was the front-runner; he faced a serious challenge from Stepan Demirchian, the son of the late speaker Karin Demirchian, who had been killed in a 1999 attack on parliament.25 The other notable challenger was Artashes Geghamian, the last Soviet-era mayor of the capital city of Yeveran.

The ongoing conflict with Azerbaijan over the Nagorno-Karabakh region has been the single most important postindependence issue in Armenian politics. Kocharian, a native of Nagorno-Karabakh, is seen as a resolute supporter of independence for Nagorno-Karabakh. Ter-Petrossian’s willingness to negotiate with Azerbaijan over the territory

25 Armenian politics are characterized by violence, which overshadowed the 2003 elections. Most notably, in 1999, the parliament was attacked by gunmen, and eight prominent politicians were assassinated. The 2003 presidential elections were the first to be held after the attack.
in 1998 was partly responsible for his resignation from the presidency and Kocharian’s succession to his post through the 1998 special elections.

Kocharian, who did not have his own political party, officially ran as an independent. He had been supported by a shifting coalition, which in 2003 included the ruling Republican Party of Armenia and the Armenian Revolutionary Federation (also known as “Dashnak” or the Socialist Party). He also enjoyed the strong support of the military. Although his resolute unwillingness to negotiate on the Nagorno-Karabakh conflict was his most defining characteristic, he also campaigned in 2003 on the promise of economic stability, as did all of the candidates. Thus, the relative homogeneity of Armenian politics, the lack of other issues on the ballot, the fact that the incumbent ran without a political party, the presence of a dominant executive, and the central issue of Nagorno-Karabakh provide the background to the 2003 presidential election and the context of the natural experiment.

The first round of the 2003 presidential elections in Armenia took place on February 19 followed by a runoff on March 5. The Armenian constitution requires a second-round runoff if no presidential candidate garners more than 50 percent of the vote in the first round in the single-district national election. The official first-round vote share for Kocharian was 49.48 percent, thus triggering a runoff election.

Several months prior to the election, the Armenian Ministry of Foreign Affairs invited the OSCE/ODIHR to sponsor an international election-observation mission. The delegation included members of the Parliamentary Assembly Council of Europe. In the first round of the election, the OSCE deployed 233 observers from 35 OSCE participating states. The second round was observed by 193 short-term observers from 21 participating states.

The Natural Experiment Research Design in Armenia

The use of experimental methods has experienced renewed popularity in political science and economics. The distinguishing characteristic of experimental methods versus observational research is that the central independent variable, or “treatment” variable, is randomly assigned. In field experiments such as those conducted recently by Gerber and Green, Miguel and Kremer, Olken, and Wantchekon, the researcher supervises the random assignment of the treatment variable. In natural
experiments, the researcher does not manage the assignment of the treatment variable, but natural experiments can occur when the variable is assigned “as-if” the assignment were random.\(^{27}\) The burden in natural experiments rests on the researcher to provide evidence that the treatment can, in fact, be treated “as-if” it had been randomly assigned. Existing natural experiments vary in the degree to which the treatment approaches true randomization.\(^{28}\)

There are few published natural or field experiments within comparative politics and fewer within international relations. Exceptions within comparative politics include the work of Guan and Green; Humphreys, Masters, and Sandbu; and Wantchekon.\(^{29}\) Despite the relative scarcity of experimental work, the advantages are well established: properly conducted experimental work is one of the few means by which causal inference can be tested.\(^{30}\) Stated more forcefully, experiments have an “unrivaled capacity to demonstrate cause and effect.”\(^{31}\)

I present a natural experiment in which international observers were assigned to polling stations on election day using a method that I did not supervise but that comes very close to random assignment. This natural experimental setup allows a test of whether international observers reduced the rate of election-day fraud during the 2003 presidential elections in Armenia. If election-day fraud occurs in any election, it should have the observable implication of increasing the vote share of the fraud-sponsoring candidate. In the case of Armenia, the incumbent sponsored the majority of election-day fraud. Therefore, if international observers have no effect on election-day fraud, then

---


\(^{29}\) Mei Guan and Donald P. Green, “Noncoercive Mobilization in State-Controlled Elections: An Experimental Study in Beijing,” *Comparative Political Studies* 39 (December 2006); Macartan Humphreys, William A. Masters, and Martin E. Sandbu, “The Role of Leaders in Democratic Deliberations: Results from a Field Experiment in São Tomé e Príncipe,” *World Politics* 58 (July 2006); Wantchekon (fn. 26).


the incumbent should perform equally well in both groups of polling stations: those that were monitored and those that were not. If international observers reduce election fraud, the incumbent’s average vote share should be lower in monitored polling stations than in unmonitored polling stations.

There are three unique features of the 2003 Armenian elections that allow a test of whether the presence of international observers reduced election-day fraud. First, widespread and centrally orchestrated fraud occurred on election day. As the Economist described it, the 2003 election was “one of the dirtiest even Armenians can remember.” As stated, fraud (and therefore fraud deterrence) can occur at many points in the electoral process. However, it would be difficult to test for an election-day deterrent effect if no fraud occurred.

Eyewitness reports from international observers, domestic observers, and journalists documented many varieties of election-day fraud. The OSCE/ODIHR observed “significant irregularities” in more than 10 percent of the polling stations they visited, the most blatant of which were ballot-box stuffing, “carousel” voting, direct vote buying, individuals voting more than once, the intimidation of witnesses for political parties, the presence of government officials inside polling stations who attempted to intimidate officials and voters, and one isolated incident of the removal of more than fifty passports from a polling station by a policeman. During the counting process there were numerous attempts to change the vote totals by the polling-station officials, and observers recorded additional evidence of blatant ballot-box stuffing. In some cases the international observers were prevented from observing the counting process, which was interpreted as an attempt to conceal illicit behavior.

The second characteristic of the 2003 presidential elections is that the Armenian Central Election Commission made disaggregated election results available. The process of recording and making public polling-station-level election results requires a certain level of administrative competence and transparency that is not always present, even in developed democracies. In countries that experience significant amounts of electoral fraud, these data are often intentionally “lost” or kept private. The Armenian election data, disaggregated to the level of the polling station, were made public by the election commission on their Web site.

The third and most important favorable feature of the 2003 Armenian elections is that the international observers were assigned in a way

---

33 Government-reported election results were made available online at http://www.elections.am by the Central Election Commission of Armenia.
that approximates random assignment. I learned from conversations with staff and participants in the OSCE observation mission to Armenia that the method used to assign observers to polling stations was functionally equivalent to random assignment. This permits the use of natural experimental design. Although the OSCE/ODIHR mission did not assign observers using a random numbers table or its equivalent, the method would have been highly unlikely to produce a list of assigned polling stations that were systematically different from the polling stations that observers were not assigned to visit. Each team’s assigned list was selected arbitrarily from a complete list of polling stations. Those making the lists did not possess information about polling-station attributes that would have allowed them to choose polling stations according to criteria that could have predicted voting patterns.

Since the validity of this natural experiment rests upon this point, I will take some time to support it. In this particular election the delegation leaders gave each team of short-term observers a preassigned list of polling stations to visit during election day. These lists were made with two objectives in mind: (1) to distribute the observers throughout the entire country (including rural and urban areas) and (2) to give each observer team a list of polling stations that did not overlap with that of other teams. Observers were encouraged to go only to those polling stations on their list and to travel between polling stations in a way that minimized travel time and still ensured coverage of their assigned polling stations. It is important for the validity of the natural experiment that the travel routes would not have been predictable by external observers, including government officials.

In the Armenian case, individuals who made these lists had little knowledge of polling-station characteristics other than their general geographic location. If the assignment of observers had been done with consideration to other variables that might be correlated with performance of the incumbent candidate, then the assignment could not be considered near random. In addition, the preceding discussion of Armenian politics indicated few observable characteristics of the population (such as socioeconomic status or ethnicity) that might be correlated with the incumbent’s popularity. In this case, the staff did not have access to disaggregated data on the demographic characteristics of the Armenian voting population. OSCE/ODIHR staff members have assured me that they had no desire to (and did not) choose polling stations on any basis other than the two criteria cited above. In addition, even if this were not true, it is highly unlikely that the mission’s office had the capability of choosing polling stations that were more or less
likely to favor the incumbent or the opposition candidates or that were more likely to experience election-day fraud.\textsuperscript{34} The fact that Armenian politics is not predictable along partisan or demographic lines underscores that this type of selection bias in the assignment of international observers would have required enormous effort, access to data that do not exist, and foresight about the trajectory of Armenian politics that would be unusual for foreigners to possess.

Additionally, assigning specific polling stations to each team cut out much of the agency on the part of individual observer teams that, in the absence of a directive, could choose to visit polling sites based on their own selection criteria within a defined geographic area. When observers are given leeway in choosing polling stations, the two most common alternative selection criteria (based on information from observation missions outside of Armenia) are to choose polling stations that are either “convenient” or “interesting.” Each of these decision criteria may create significant observation bias. This has been pointed out as a problem by several critics of election observation.\textsuperscript{35} Observer teams that select “interesting” polling stations typically go to areas in which problems are expected, so that teams using this decision rule may disproportionately observe and report irregularities. This is a common strategy among the more ambitious and enthusiastic international observers but was discouraged in this particular case.

Observer teams that go to “convenient” areas are criticized for being electoral tourists. This was more common in early election observation, although it seems to have continued as a favorite point to criticize in other observer groups. Other “convenient” selection methods may be observing near the observers’ hotel in the most comfortable urban areas or going to polling stations that are near tourist destinations. Clearly, these selection criteria are nonrandom and could lead to bias in both the observers’ reported observations and the natural experiment proposed here, particularly because a clever politician could recognize the tendency of observers to stick to certain areas and therefore concentrate any electoral manipulation in places where observers would be unlikely to go. For these reasons it was particularly important for this natural experiment that this type of observer agency was explicitly discouraged.

\textsuperscript{34} Even if this information were inaccurately communicated to me, if observers were more likely to visit stations they believe to be problematic, as suggested by one reviewer, then this would dampen an observed effect of observers on fraud. For the reasons cited, however, this is an unlikely scenario.

In sum, the assignment of international observers to polling stations for both rounds of the 2003 presidential elections in Armenia is very close to random. The selection was made arbitrarily from a list of all polling stations with only geographic logistics in mind, and the assignment was completed with no knowledge of variables that might be correlated with the incumbent’s likely vote share. Teams were instructed to visit only the polling stations assigned to them, and, due to the relatively small geographic area and limited number of polling stations assigned to each team, they had a high probability of reaching their assigned polling stations.

There are a few other points about election-observation methodology that are not unique to Armenia but that are important components of the validity of this natural experimental design. First, international observers do not preannounce which polling stations they will observe on election day. Keeping deployment plans confidential is standard practice for reputable international observer groups and is intended to enhance the safety of the observers by making it difficult for attackers to anticipate where their potential targets will be. It also makes it difficult for the candidates to anticipate the arrival of observers and thereby restrict their cheating to polling stations where international observers are not expected.

Second, international observers were mobile, moving from polling station to polling station throughout election day. During the course of one day, an observer team could visit ten to twenty polling stations based on the length of the election day and the distance between polling stations. Critics of election observation are fond of pointing out that it would be very difficult to catch any irregularities in twenty minutes of observation. However, if there are ongoing problems or “red flags” indicating that there might be problems, observer teams are instructed to stay for as long as they think is useful, which in some cases can be as long as several hours; however, they are not permitted to interfere in the electoral process. The questionnaires filled out by international observers include a number of observations related to the structure of the polling station, the available staff and materials, and the order of voting procedures that are immediately obvious. If the seal on the ballot box has been broken, international observers are most likely to see this and other evidence of fraud as soon as they enter a polling station. In addition, the partisan witnesses in each polling station remain in the same polling station throughout the day and are often able to report irregularities to the international observers.
Data and Results

The central measurable effect of observers on election-day fraud is to decrease the vote share for the incumbent. Holding all else constant, if international observers did in fact reduce fraud at the polling stations they visited, then the incumbent should perform worse in observed polling stations. Random assignment (or “as-if” random assignment) of the treatment of international observers is equivalent to holding all else constant.

Because international observers can be considered randomly assigned to polling stations and because there were two rounds of the presidential election, the natural experimental design involves two rounds of “treatment” and a separate observation of vote share for each round. Data on vote share for Kocharian are reported. The international observers went to different polling stations in each round of the election but had some overlap between rounds. This divides the sample of polling-station-level election results into four experimental groups based on the treatment of international observation during the course of election day: one group of polling stations was never monitored (N=755), one group was monitored only in the first round (N=385), one was monitored only in the second round (N=260), and one group was monitored in both rounds (N=363).

Groups of polling stations received all possible combinations of the international observer treatment, including no treatment in either round of the election. Therefore the natural experiment also allows a test of whether first-round observation has any lasting effect in the second round. Approximately 43 percent of polling stations were not observed in either round of the election, and about 21 percent were observed in both rounds.

The Effect of Monitors on Vote Share

The dependent variable is the vote share for the incumbent presidential candidate, Kocharian. Depending on the comparison being made, three different forms of Kocharian’s vote share are used. For tests that compare groups separated by a difference in first-round treatment, the dependent variable is the mean vote percentage for Kocharian in the first round. For tests between groups separated by second-round treatment,
the dependent variable is the mean vote for Kocharian in the second round. When appropriate, I also make comparisons using the average vote for Kocharian across both rounds as the dependent variable.

A series of two-sample difference of means tests examines whether international observers reduce fraud and, if so, to what degree. A difference of means test, or *t*-test, compares two groups of observations and tests the hypothesis that the mean of the two groups is identical. The test assumes equal variances between the two groups, computes the result in terms of a specific confidence interval (in this case 95 percent). The results in Table 1 present the difference of means tests and show that international observers influenced election-day manipulation in the polling stations that they observed. To illustrate, the first comparison shown in Table 1 is the average vote share received in round 1 by Kocharian in unobserved polling stations compared with the average vote share for Kocharian in all observed polling stations in round 1. The parenthetical material in columns 1 and 2 refers to whether the percentage is in terms of Kocharian’s round 1 vote share, his round 2 vote share, or an average across the group in both rounds. The last column presents the difference in Kocharian’s vote share between the two groups and indicates whether the difference is statistically different from zero.

The results presented in Table 1 show clear evidence that during the 2003 presidential elections in Armenia, the presence of international observers reduced the vote share for the incumbent politician by about 5.9 percent in the first round and by more than 2 percent in the second round. Both results are statistically significant. This allows a rejection of the null hypothesis that there is no difference between observed and unobserved polling stations. Figure 1 illustrates this difference with a kernel density plot of Kocharian’s round 1 vote share in monitored and unmonitored polling stations. Note the unusual distribution of vote share in unmonitored polling stations.

In addition to testing whether international observers deterred election-day fraud (and the degree to which they did so), the difference of means tests also show whether the presence of international observers had a “transitory” or “persistent” effect. A transitory effect is similar to the effect of a painkiller on a headache. The treatment affects only one headache and has no effect on subsequent headaches. A persistent treatment has a lasting effect, such as an injection to relieve frequently occurring migraines. In the case presented here in which there were two rounds of the election, it appears that a visit by monitors in the first

---

38 Additional tests compare vote share for all candidates and show that the only candidate to perform better in unmonitored polling stations was Sargsian, who received only 0.04 percent more of the vote in unmonitored polling stations.
### Table 1
**Difference of Means Tests Comparing “Treatment” and “Control” Groups**

<table>
<thead>
<tr>
<th>Average Incumbent Vote Share among Polling Stations That Were...</th>
<th>vs.</th>
<th>Average Incumbent Vote Share among Polling Stations That Were...</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Not observed in R1 vs. observed in R1</td>
<td>54.2%</td>
<td>48.3%</td>
<td>5.9%</td>
</tr>
<tr>
<td>(R1 vote share)</td>
<td>(R1 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Not observed in R2 vs. observed in R2</td>
<td>69.3%</td>
<td>67.3%</td>
<td>2.0%</td>
</tr>
<tr>
<td>(R2 vote share)</td>
<td>(R2 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Never observed vs. observed in both R1 and R2</td>
<td>70.7%</td>
<td>66.2%</td>
<td>4.5%</td>
</tr>
<tr>
<td>(R2 vote share)</td>
<td>(R2 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Never observed vs. observed in both R1 and R2</td>
<td>62.8%</td>
<td>57%</td>
<td>5.8%</td>
</tr>
<tr>
<td>(Average of R1 and R2 vote share)</td>
<td>(Average of R1 and R2 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Never observed vs. observed in one or both rounds</td>
<td>62.7%</td>
<td>58.1%</td>
<td>4.6%</td>
</tr>
<tr>
<td>(Average of R1 and R2 vote share)</td>
<td>(Average of R1 and R2 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Never observed vs. observed only in R1</td>
<td>70.7%</td>
<td>66.3%</td>
<td>4.4%</td>
</tr>
<tr>
<td>(R2 vote share)</td>
<td>(R2 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Never observed vs. observed only in R2</td>
<td>70.7%</td>
<td>68.7%</td>
<td>2.0%</td>
</tr>
<tr>
<td>(R2 vote share)</td>
<td>(R2 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Observed only in R2 vs. observed in both R1 and R2</td>
<td>68.7%</td>
<td>66.2%</td>
<td>2.5%</td>
</tr>
<tr>
<td>(R2 vote share)</td>
<td>(R2 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Observed in both R1 and R2 vs. observed only in R1</td>
<td>66.3%</td>
<td>66.2%</td>
<td>.11%</td>
</tr>
<tr>
<td>(R2 vote share)</td>
<td>(R2 vote share)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Observed only in R1 vs. observed only in R2</td>
<td>68.7%</td>
<td>66.3%</td>
<td>2.4%</td>
</tr>
<tr>
<td>(R2 vote share)</td>
<td>(R2 vote share)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Reported results reflect two-sample t-tests with equal variances.*
round has a persistent effect on second-round behavior, indicating that international observers had both transitory and persistent effects.

Transitory Effects

The first two tests presented in Table 1 reveal the transitory effect of observers on electoral fraud committed by supporters of the incumbent. The first test compares the incumbent’s performance in the first round between unmonitored and monitored polling stations. The incumbent did about 6 percent better in polling stations without international observers, suggesting that the presence of international observers in round 1 reduced election-day fraud by about 6 percent in visited polling stations. The level of statistical significance is high. In the second round test 2 shows that the deterrent effect of observers continued. The incumbent performed about 2 percent better in polling stations where observers were not present, suggesting a reduction in the size of the effect. However, test 2 does not control for any persistent effects that first-round observation may have had on second-round behavior.

Tests 3 and 4 compare the two groups of polling stations that received the same treatment in both rounds (never monitored versus monitored in both rounds). Test 3 uses the round 2 incumbent vote
share to compare polling stations that were never monitored and polling stations that were monitored in both rounds. In this comparison, the incumbent received 4.5 percent more of the round 2 vote in polling stations that were never observed relative to those observed in both rounds.

Test 4 makes the same comparison using the two-round average vote share and shows a similar result. Across both rounds, there is a 5.8 percent difference in average vote share for the incumbent. Together, tests 3 and 4 provide more nuanced empirical support for the finding that international observers had a strong deterrent effect on the incidence of election-day fraud.

Test 5 adds one additional piece of support for the hypothesis that observers reduced election-day fraud. Test 5 uses the average vote share for both rounds and shows that over the course of the election, the incumbent received 4.6 percent more of the vote in polling stations that were not monitored in either round as compared with the average vote share in polling stations visited in one or both rounds.

**Persistent Effects**

The second set of tests examines the possibility that the presence of international observers in the first round had a persistent effect in the second round. If the effect of international observers had lasting effects on fraudulent behavior, the polling stations that were monitored in the first round should be less likely to experience fraud in the second round.

Tests 2 and 3 demonstrate a persistent effect. If first-round observation had no effect on second-round fraud, then the difference in the incumbent vote share should be about the same in the two tests. When polling stations that were monitored in round 1 but not in round 2 are included in the control group and the round 2 vote share is used as the dependent variable, as in test 2, the observed deterrent effect shrinks by more than half, from 4.5 percent to 1.98 percent.

In order to test whether the treatment of first-round monitoring had a persistent deterrent effect on fraud in the second round of the election, test 6 compares the round 2 vote share between two groups that were not monitored in the second round. One group was monitored only in the first round; the other group was not monitored in either round. The result of this test provides additional evidence that observers had a persistent deterrent effect on electoral fraud. Monitoring in the first round is associated with a decrease in the incumbent’s second round vote share by 4.4 percent. This implies that polling station officials who were visited by international observers in the first round were less likely to commit fraud in the second round.
Comparing the round 2 vote share between polling stations that were monitored in round 2 and those that were never monitored, test 7 shows that the incumbent performs about 2 percent worse in monitored polling stations, but the difference is not quite statistically significant. Similarly, test 8 compares the round 2 vote between polling stations that were monitored in both rounds with those that were monitored only in round 2. The difference of means test shows a small difference between the two groups that nearly achieves statistical significance. The relatively smaller size of the round 2 effect is not explained by the results but could stem from the fact that Kocharian earned close to 50 percent of the round 1 vote and around 70 percent of the round 2 vote. These results suggest that if first-round monitoring took place then second-round monitoring had only a marginal additional deterrent effect.

Within first-round monitored polling stations test 9, like test 8, provides additional support that if international observers were present in round 1, round 2 monitoring does not have a statistically significant effect. Test 10 indicates that there is little difference in overall fraud reduction between the two groups that were monitored only once, regardless of whether the monitoring took place in the first round or the second round.

**Checking As-If Randomization**

Ideally, in any experimental research design the assignment of the treatment could be examined in relation to a background covariate in order to test for balance between the treatment and control groups. In this case, the ideal covariate would be an independent measure of the candidates’ likely vote share, such as public opinion polling. These data are not available for Armenia at the polling-station level for the first round, but as Table 2 shows, observer distribution does not appear to follow a clear pattern that would predict Kocharian vote share. Coverage varies by region from a low of 28 percent of polling stations monitored in Aragatsotn to 69.59 percent of polling stations monitored in the capitol of Yerevan. The last column of Table 2 suggests that much of this difference is due to voter density, as there is relative balance in voters per monitored polling station within each region. Additionally, the OSCE mission observed extensively outside of urban areas where there are fewer voters and travel is more time consuming. To illustrate, an urban polling station is defined as one that is in the region of Yerevan, is a regional capital, or is one of the seven biggest cities (population > 40,000). All other polling stations are nonurban, which includes rural and periurban polling stations. Using these criteria, 45 percent of all polling stations are nonurban. International observers visited 38 percent nonurban polling
stations in the first round of the election and 35 percent nonurban in the second round. Given that there are more voters in each urban polling station, observers covered nonurban areas extensively.

As a further examination of the “as-if” randomization, round 2 treatment of observers is compared with vote share and turnout in round 1. These round 1 outcomes should be equal between polling stations that were monitored in the second round and those that were not. The results presented in Table 3 support the assertion that the method used by the OSCE/ODIHR to assign observers approximates randomization. In an OLS regression, the presence of observers in the second round is unrelated to voter turnout or to vote share for Demirchian.39

**Table 2**

**Round 1 Observer Coverage by Region**

<table>
<thead>
<tr>
<th>Region</th>
<th>Total Polling Stations</th>
<th>Percent Monitored</th>
<th>Average Voters per Polling Station</th>
<th>Voters/Total Monitored Polling Stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aragatsotn</td>
<td>133</td>
<td>27.82</td>
<td>701</td>
<td>2520</td>
</tr>
<tr>
<td>Ararat</td>
<td>137</td>
<td>53.28</td>
<td>1355</td>
<td>2543</td>
</tr>
<tr>
<td>Armvir</td>
<td>153</td>
<td>38.56</td>
<td>1319</td>
<td>3420</td>
</tr>
<tr>
<td>Geglharkunik</td>
<td>148</td>
<td>32.43</td>
<td>1140</td>
<td>3515</td>
</tr>
<tr>
<td>Kotayk</td>
<td>132</td>
<td>43.18</td>
<td>1429</td>
<td>3309</td>
</tr>
<tr>
<td>Lori</td>
<td>226</td>
<td>33.63</td>
<td>1059</td>
<td>3148</td>
</tr>
<tr>
<td>Shirak</td>
<td>273</td>
<td>25.64</td>
<td>907</td>
<td>3537</td>
</tr>
<tr>
<td>Syunik</td>
<td>54</td>
<td>37.04</td>
<td>859</td>
<td>2319</td>
</tr>
<tr>
<td>Tavush</td>
<td>80</td>
<td>28.75</td>
<td>1152</td>
<td>4007</td>
</tr>
<tr>
<td>Vayots Dzor</td>
<td>40</td>
<td>37.50</td>
<td>1017</td>
<td>2712</td>
</tr>
<tr>
<td>Yerevan</td>
<td>388</td>
<td>69.59</td>
<td>1751</td>
<td>2516</td>
</tr>
</tbody>
</table>

Even though the assignment of international observers approximated randomization and observers were instructed to visit only the polling stations assigned to them, and in theory random assignment controls for possible covariates, it is still possible to test whether other variables that could influence the incumbent’s vote share bias the results. The three most likely alternative explanations for which data are available are observer difficulty in reaching small polling stations, the nature of

39 Demirchian vote share is used to make the comparison clearer. If there were systematic biases in the polling stations monitored in round 2, they should show as a statistically significant difference between these two groups. Comparisons using round 1 Kocharian vote share are available from the author, but for reasons that are likely related to polling-station accessibility, round 1–monitored polling stations are more likely to be monitored in round 2, thus making it unlikely that observed round 1 Kocharian vote share would be statistically independent of round 2 monitoring.
the urban-rural split, and whether a given polling station is located in one of the regions that border Nagorno-Karabakh.

First, note that the variables recorded in the analysis are actual visits by international observers rather than assignment of the polling station to the treatment group. If observers had systematic difficulty in reaching polling stations they were asked to visit due to, for example, factors such as muddy roads, the observed difference between monitored and unmonitored polling stations could be the result of systematic differences between polling stations that are easy to find and those that are not. This would lead to a biased test if and only if polling stations that were difficult for observers to find were systematically related to support for a particular candidate. Table 4 presents three alternative specifications of an OLS model examining the effects of monitoring and proxies for polling-station accessibility on Kocharian vote share in round 1. Two measures of polling-station size are used: registered voters and total votes cast, as both could be influenced by fraudulent practices and as the total number of votes cast could also be a reasonable proxy for the difficulty in reaching polling stations (that is, both voters and observers have a difficult time reaching the polling station). The effect of monitoring on Kocharian vote share remains statistically significant in models 1–3, although the size of the monitoring effect varies somewhat with the specification and is smaller than the effect presented in Table 1. This suggests that polling-station size is related to monitoring (most likely because smaller stations are harder to find),

---

The table below presents the results of the OLS model examining the effects of monitoring and proxies for polling-station accessibility on Kocharian vote share in round 1.

<table>
<thead>
<tr>
<th>DV: Voter Turnout Round 1</th>
<th>DV: Demirchian Vote Share Round 1</th>
<th>DV: Demirchian Vote Share Round 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coefficient</strong> (S.E.)</td>
<td><strong>Coefficient</strong> (S.E.)</td>
<td><strong>Coefficient</strong> (Robust S.E.)</td>
</tr>
<tr>
<td>Round 2 Monitored</td>
<td>0.018 (0.010)</td>
<td>0.002 (0.007)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.637** (0.006)</td>
<td>0.276** (0.004)</td>
</tr>
<tr>
<td>N</td>
<td>1763</td>
<td>1763</td>
</tr>
<tr>
<td>F</td>
<td>(1, 1761) = 2.96</td>
<td>(1,1761) = 0.10</td>
</tr>
<tr>
<td>Prob &gt; F</td>
<td>0.0853</td>
<td>0.7471</td>
</tr>
</tbody>
</table>

* significant at 5%; ** significant at 1%

Robust standard errors are clustered by region.

---

I am grateful to an anonymous reviewer for suggesting this be included as an alternative explanation.
but the effect of monitoring is robust to the inclusion of measures of polling-station size.

The second alternative explanation is that if Kocharian performed disproportionately well in rural areas and if the selection of polling stations included a disproportionate number of urban polling stations, then it is possible that Kocharian’s support in rural areas could be driving the difference between monitored and unmonitored polling stations. Table 5 presents the results of difference of means tests between experimental groups using round 1 vote share. The other dependent variables (round 2 vote share and average vote share between rounds) are consistent with the results presented below.

Table 5 shows that the reported differences in average Kocharian vote share between monitored and unmonitored polling stations are not driven by urban-rural bias or by voters in the Nagorno-Karabakh region. The first two rows of Table 5 show that the deterrent effect of observers is present in urban areas. If Kocharian’s high performance in rural areas were driving the main results, then the difference in Kocharian’s performance between monitored and unmonitored polling stations should disappear when rural areas are excluded.

The final alternative explanation is that because Kocharian is a native of Nagorno-Karabakh and is overwhelmingly popular there, the results could be driven by disproportionately high support near Nagorno-Karabakh if observers avoided the area. Rows 3 and 4 of Table 5 show...
that the observed deterrent effect of observers is not a result of unusually high support in the regions most affected by the conflict in Nagorno-Karabakh. If observation were biased away from the Kocharian stronghold near Nagorno-Karabakh, the nationwide results could be biased. Table 5 shows that the results hold in the regions outside of the region of Kocharian’s highest support.

In summary, the first-round deterrent effect is consistently present in comparisons between monitored and unmonitored groups, and three alternative explanations do not account for the difference in the incumbent’s vote share between monitored and unmonitored polling stations.

### DISCUSSION OF FRAUD REDUCTION

This natural experimental design has provided a unique test of whether international observers can deter election-day fraud, providing causal evidence of how international actors can influence domestic politics. The evidence presented above shows that the presence of international observers depressed the incumbent’s round 1 average vote share by 6 percent in polling stations that were observed. Since the incumbent was only 0.52 percent away from avoiding a round 2 runoff, international observers may have been responsible for triggering a second round of competition.41 It is also possible that amidst the already charged at-

---

41 In the first round, observers visited 42 percent of polling stations. The 5.8 percent average reduction in Kocharian vote share, reflected nationally, was 2.44 percent (aggregate observer effect = 42% * 5.8%).
mosphere surrounding the election and Kocharian’s desire to play the election game (if not abide by all the rules), Kocharian engineered a last-minute reduction in his vote share in order to look more like a democrat. This possibility does not detract from the demonstrated deterrent effect of international observers on election-day fraud; it simply moderates the possible claim that international observers caused the second-round runoff election.

The evidence does not suggest that international observers deterred fraud to the point that they could not observe it. As already detailed, observer teams witnessed a significant number of electoral irregularities, including evidence of ballot-box stuffing, political violence, and intimidation of polling officials. The official OSCE/ODIHR report lists twenty-two polling stations where ballot-box stuffing was witnessed in the first round and more than seventy polling stations in the second round. In the case of Armenia, election-day fraud deterrence meant that the international observers slowed down the rate of fraud in the polling stations that they visited. The precise mechanisms by which this took place are unknown, but a plausible explanation is that the presence of international observers caused polling-station officials to reduce the rate of intended election-day fraud, either because they were instructed to stop fraudulent activities in front of international observers or because many were worried about being caught.

What explains the persistent deterrent effect of observers on fraud in round 2? It is possible that in the first round of the election, the polling-station workers were uncertain regarding the reach and quality of the international observers. Prior to the election the commission charged with administering the election had undergone significant restructuring. The restructuring was such that most of the polling-station administrators were selected just before the 2003 presidential elections, making them relatively inexperienced with international observers. After being observed in the first round and witnessing the mission’s negative assessment of the election, those officials monitored in the first round may have updated their beliefs regarding the quality and reach of international observers and decided they were unlikely to get away with election-day fraud. Those that were not observed in the first round were less likely to update their beliefs on the reach of international observers but may have still been somewhat deterred in the second round by the OSCE’s negative assessment of the first round of elections.

Election observation in the first round decreased round 1 fraud and to a lesser degree also decreased fraud in the second round. The second-round reduction of fraud likely occurred in three ways: those that were monitored
in the first round committed less round 2 fraud, those that were monitored in the second round committed less fraud in the second round, and because the observers demonstrated that they were likely to give a negative evaluation and were capable of catching fraudulent activities, all polling stations were less likely to experience second-round fraud.

**Conclusion**

Reducing electoral fraud on election day is one way that international observers can influence politics in democratizing countries, and it provides relatively unique causal evidence that international actors can affect the democratization process. This article has measured the magnitude and the persistence of this effect in the 2003 presidential elections in Armenia. By examining a localized effect rather than cross-national correlations, this method was able to sidestep problems with endogeneity. The finding of a statistically significant and robust difference between monitored and unmonitored polling stations is somewhat surprising, in part because international monitors spent only a short time in each polling station, and reports from observers suggest that some voting officials committed fraudulent activities quite blatantly in front of them. However, the results indicate that international monitors can reduce the corrupt behavior of election officials and point to an increased role for nonpartisan observation of elections and other political processes with potential for corruption. The experiment could be replicated using field experimental methods in order to test whether domestic nonpartisan observers have more or less of a deterrent effect than foreign observers. As other experiments have shown, most recently Olken, the potential for documentation and exposure of corruption can have positive downstream effects on individual behavior.\(^\text{42}\)

For IR theory, the results represent empirical evidence of how international institutions matter. The international expectation that leaders will invite international election monitors means that leaders engaging in undesirable behavior choose to invite international scrutiny even when doing so may be costly. In the Armenian case, I have demonstrated the increased costs in terms of the incumbent’s vote share, but one can imagine other scenarios in which international pressure and international scrutiny combine to limit the choices for state leaders and influence the trajectory of domestic politics in important ways.

\(^{42}\text{Olken (fn. 26).}\)
Deterring fraud is not the only effect that international observers could have on domestic politics. Their presence may also provide the incentive for electoral autocrats to use methods of cheating that are less likely to be detected by international observers, such as manipulating the election in advance of election day. Knowledge that observers would be present did not prevent supporters of President Kocharian from engaging in election-day manipulation, but this remains a distinct possibility elsewhere.

On a more practical note, these results suggest that observer groups should be wary of the fact that they may be less likely to observe electoral fraud when they enter a polling station. With true random assignment of international observers and the availability of polling-station-level results, it is possible to measure the magnitude of this form of an “observer effect” and utilize this information as a component of fraud detection.

This article does not hypothesize about whether the results can be extended outside of the context of the experiment. Do international observers always deter election-day fraud when it exists? Are elections in which no election-day fraud was witnessed to be considered cases of fraud deterrence or of elections that would have been free and fair regardless of international observers? Alternatively, could fraud-free election days simply be cases in which the incumbent was able to manipulate the election before election day? It is likely that a wide spectrum of cases exists, although to date, none of these questions has been answered with systematic empirical evidence. The results presented here represent one step toward an improved understanding of how international actors influence domestic politics and provide a test of the effectiveness of one of the most widespread democracy-promotion activities carried out by international organizations and international NGOs.

Finally, this study shows that reducing election fraud does not translate inevitably into democratization. The 2007 parliamentary elections in Armenia were also criticized, and it is unclear at this time whether Armenia will successfully democratize. Even so, a proven method by which fraud can be reduced may help bring about a democratic election in which voter preferences are accurately translated into political representation. More observers can make it more difficult for elections to be stolen, by reducing the rate of fraud and increasing the likelihood that fraud will be caught. International observers are not a magic bullet, but this article has shown that they can influence the quality of elections.