PROJECT MUSE®

The Pseudo-Democrat's Dilemma

Hyde, Susan D.

Published by Cornell University Press
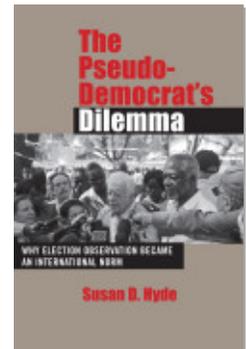
➡ For additional information about this book

https://muse.jhu.edu/book/26089

4

DOES ELECTION
MONITORING MATTER?

Are international election monitors more costly to pseudo-democrats than to true democrats? In this chapter, I continue investigating the consequences of internationally monitored elections as they relate to norm formation and show that the presence of international observers correlates with several types of costs to incumbent leaders. I also use experimental evidence involving the randomization of short-term election observers to demonstrate that international observers can have a direct deterrent effect on election day fraud. By causing a reduction in vote share through fraud deterrence, at the very minimum election monitors make it more difficult for pseudo-democrats to steal elections, a cost that must exist for election monitoring to be a credible signal of a government's commitment to democratic elections. I also show that for true democrats, the same cost does not exist and present evidence that observers can have unintended consequences that help incumbent true democrats. Chapter 5 continues this investigation by exploring the evolving game of strategy between international observers and pseudo-democrats. Here I provide a detailed examination of whether international observers have both direct and micro-level effects on election day behavior in two elections, either by deterring fraud or making it more likely that polling officials follow the rules.

An important empirical implication of my theory is that election observation must be more costly to pseudo-democrats than to true democrats. Supporters of election observation argue that observers improve the quality of elections and make it more likely that election fraud will be uncovered. Cross-nationally, the presence of international observers is correlated with a variety of outcomes that are costly to the incumbent governments that invite them. These correlations suggest that, on average, observed elections are more competitive, and the incumbent government

is more likely to lose such elections. I present these cross-national descriptive statistics to support my argument and to contrast them with the experimental evidence provided later in this chapter. Although it is true that incumbents are more likely to lose elections when monitors are present, such evidence cannot demonstrate a causal relationship between observers and increased costs to pseudo-democrats, unlike the natural and field experimental evidence shown in this chapter.

Incumbent Turnover and the Correlates of Observed Elections

In chapter 1, I argued that pseudo-democrats, all else equal, should perform worse when international monitors are present. The descriptive statistics below provide some support for this claim, but establishing causality is difficult. Critics of international election monitoring have been especially vocal in arguing that the central raison d'être of observers is to legitimize flawed elections.[1] It is true that observers sometimes legitimize flawed elections. My argument suggests, however, that the possibility that observers will not condemn a manipulated election is a central reason why pseudo-democrats are willing to invite them in the first place. Nevertheless, the fact that pseudo-democrats take this risk should increase the chances that they will lose, either directly, through fraud deterrence, or indirectly, by making election fraud more expensive and diverting resources that could be used elsewhere. Consider the following descriptive statistics for elections held between 1990 and 2006, when international observers were widely available (I exclude the consolidated democracies for ease of interpretation):

- Out of all elections in which the incumbent leader or governing party was replaced, 75% were internationally monitored (161 of 215).[2]
- Out of all elections in which the office of the incumbent executive was contested, and in which the incumbent both ran and was replaced, 68% were internationally observed (45 of 66).[3]

1. Abbink and Hesseling, *Election Observation and Democratization in Africa*; Geisler, "Fair?"

2. Data are from the Hyde and Marinov NELDA dataset described in appendix B.

3. Coding of leaders from the Archigos dataset, Goemans, Gleditsch, and Chiozza, "Introducing Archigos." Other data from NELDA data described in appendix B (Hyde and Marinov, "National Elections Across Democracy and Autocracy").

- From 1960 to 1989, observers were present at only 10% of elections in which there were significant concerns before the election that it would not be free and fair. After 1989, this number increases to 69% and to 75% when single-party elections are excluded.
- Out of all elections in which the vote count was considered a "gain for the opposition," observers were present at 70% (258 of 368).

These data suggest that observers are invited by incumbent governments to many elections that are more competitive and that are more likely to result in negative outcomes for the incumbent regime. Yet it is possible that governments were more likely to invite observers in these circumstances. Although they are suggestive, these cross-national data do not demonstrate that observers caused more competitive election outcomes.

An Alternative Micro-level Approach to Testing the Effects of Observers

One persistent alternative argument is that it is costless for leaders to invite international election monitors, and the phenomenon is therefore not interesting as a case of international norm formation. This chapter disputes this alternative explanation by showing that inviting monitors can have several types of direct effects on election day behavior. Using experimental methods involving the randomization of international election observers, I show that in Armenia (2003) the (pseudo-democratic) incumbent president performed much worse in polling stations that were monitored compared with those that were not. In Indonesia (2004), in what was widely considered a democratic election, observer presence actually increased votes cast for the incumbent true democrat, who went on to lose the election and peacefully transfer power.

Natural and field experimental tests are increasingly popular within the social sciences because of their potential to demonstrate cause and effect. The cases included in this chapter are described in greater detail elsewhere and are condensed here in order to link the results from these projects to the international norm of election observation.[4] 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 Alan

---

4. See Hyde, "The Observer Effect in International Politics"; Hyde, "Experimenting with Democracy Promotion"; Dunning and Hyde, "The Analysis of Experimental Data."

Gerber and Donald Green,[5] Edward Miguel and Michael Kremer,[6] Benjamin Olken,[7] and Leonard Wantchekon,[8] 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 was random.[9] 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.[10]

There are few published natural or field experiments within the discipline of comparative politics and fewer within international relations. Exceptions for comparative politics include the work of Mei Guan and Donald Green;[11] Macartan Humphreys, William Masters and Martin Sandbu;[12] and Wantchekon.[13] 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, and as has been noted in the leading political science journal, experiments have an "unrivaled capacity to demonstrate cause and effect."[14]

I evaluate the micro-level effects of observers using evidence from two different natural and field experiments surrounding elections in Armenia in 2003 and Indonesia in 2004. In both cases, I exploit the randomization of international observers during their election day observation in order to test for differences between the areas that were visited and the areas that were not. Due to the randomization (or in the case of Armenia, the "as if" randomization), any differences between the groups can be causally attributed to international election observers. By combining experimental tests of the effects of international observers with cross-national and qualitative evidence, I provide a more complete picture of the effects of election monitoring, provide some hard evidence that election monitors

5. Gerber and Green, "The Effects of Canvassing."
6. Miguel and Kremer, "Worms."
7. Olken, "Monitoring Corruption."
8. Wantchekon, "Clientelism and Voting Behavior."
9. For examples of natural experiments, see Snow, *On the Mode of Communication of Cholera;* Schargrodsky and Galiani, "Property Rights for the Poor."
10. Dunning, "Improving Causal Inference."
11. Guan and Green, "Noncoercive Mobilization in State-Controlled Elections."
12. Humphreys, Masters, and Sandbu, "The Role of Leaders in Democratic Deliberations."
13. Wantchekon, "Clientelism and Voting Behavior."
14. Druckman et al., "The Growth and Development of Experimental Research in Political Science," 627.

are costly to pseudo-democrats, and link my theory of international norm formation with its micro-level implications.

Direct Effects on Election Day Behavior

Although academics remain skeptical that observers can have direct effects on election day behavior, practitioners and proponents of international election observation frequently assert that international observers reduce election fraud and otherwise improve election quality. As referenced above, a cross-national study of this issue cannot demonstrate causality: elections that are clean because they are internationally observed are indistinguishable from elections that are observed but would have been clean without international observers. Although the cross-national data are consistent with the argument that observers reduce election fraud and observed elections are systematically more competitive than unobserved elections, there is no set of "control" variables that could disprove the alternative argument: leaders are more likely to invite observers when they know elections are going to be more competitive.

This chapter illustrates that the effects of international election observers can be measured by exploiting subnational variation in election results. In the most clear-cut case, if observers reduce election fraud directly owing to their presence in polling stations, this effect should be revealed by differences in voting patterns between monitored and unmonitored polling stations.

The typical election observation mission includes teams of observers who, throughout the course of election day, roam within predefined geographic areas. Observers take note of activities in and around polling stations and frequently talk to voters, polling station officials, political party witnesses, domestic nonpartisan observers, and to other international observers.

Observers are also able to observe fraud directly, and observers sometimes express surprise that election fraud is carried out blatantly in front of them. Teams of foreign observers have directly witnessed many forms of outright election day manipulation, including premarked and bundled stacks of ballots clearly visible within transparent ballot boxes; the allocation of multiple ballots to individual voters; theft of ballot boxes and burning of ballots before the count; overt vote-buying schemes; prohibitions on secret voting, including the presence of individuals who "assist" voters by filling out their ballots for them and/or accompanying each voter into the voting booth; illegal disenfranchisement of eligible voters; or widespread voting by children.

Less overt "irregularities" are also commonly witnessed by international observers, and these may or may not be a sign of intentional election manipulation. These more ambiguous irregularities include polling stations that open late or without sufficient materials, a lack of provisions for secret voting, the failure to follow procedures that prevent multiple voting, election day violence and disturbances that reduce voter turnout, extremely long lines, and the presence of unauthorized individuals, police, or military in polling stations.

Although it is clear from the records of election-monitoring missions that election manipulation tactics are frequently carried out in front of observers, observers may reduce election fraud without eliminating it entirely. Similarly, even though international monitors witness irregularities, polling station officials may be more likely to follow rules and regulations when they are being watched by international monitors. The remainder of this chapter tests for an "observer effect" across varying conditions, including in the presence of overt election day fraud (in the case of Armenia) and in relatively democratic elections (in the case of Indonesia). The inclusion of both studies in this chapter demonstrates that, at least in this comparison, observers were more costly to the pseudo-democratic government (in Armenia) than the incumbent government in Indonesia that was revealed to be a true democrat.

There are a few other points about election-observation methodology that are not unique to these countries but that are important to note. 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 potential attackers to anticipate where observers will be. It also makes it difficult for the competing parties to anticipate the arrival of observers and thereby restrict their cheating to polling stations at which international observers are not expected.

Second, international observers are usually mobile, moving from polling station to polling station throughout election day. During the course of one day, an observer team could visit between ten and twenty polling stations or neighborhoods 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 such short periods 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. 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 there throughout the day and are often able to report irregularities to the international observers.

Elections in Armenia and Indonesia

The 2003 presidential elections in Armenia provide the opportunity for a direct test of whether international observers can reduce election fraud. During these elections, international observers from the Organization for Security and Cooperation in Europe's Office for Democratic Institutions and Human Rights (OSCE/ODIHR) were assigned to polling stations in a way that closely approximates randomization, making this case a natural field experiment. In the 2004 presidential elections in Indonesia, I worked directly with the Carter Center in designing the deployment plans. The Carter Center, founded by former U.S. President Jimmy Carter and Rosalynn Carter and based in Atlanta, GA, is a well-respected international election-monitoring group. The field experiment was generated by randomizing the assignment of international observers within predefined geographic areas.

The case of Armenia demonstrates that when election day fraud occurs, election monitoring can cause a reduction in the vote share of the cheating candidate. The Indonesian case provides an interesting contrast because election day fraud was not widely expected, and the incumbent government was not accused of cheating. Even so, international observers increased the vote share of the incumbent candidate, I argue, by making it more likely that polling station officials followed election day regulations.

By presenting these cases together, I show that international election observers can have important effects on election quality. Particularly when election day fraud is widespread, observers can reduce election day fraud directly, making it more difficult for pseudo-democrats to steal elections. True democrats experience no such costs, and even when elections are widely perceived to be democratic, international observers can have important and sometimes unanticipated effects on election day behavior. In Indonesia, these unanticipated effects actually *increased* the vote share for the incumbent, who went on to lose the election and peacefully transfer

power. Overall, the results strongly refute the alternative hypothesis that inviting international observers is costless to pseudo-democrats or is equally costly to both types of leaders.

For both cases, I provide a brief background to the election, an evaluation of the randomization, and a summary of the empirical results and conclusions. I then conclude the chapter by discussing the results of these experiments in relation to my theory.

The 2003 Presidential Elections in Armenia

Of all countries that invite international election monitors, Armenia represents what might be termed a repeat offender. Between independence in 1991 and the 2003 presidential election, Armenia held six elections, none of which were viewed as democratic. Political participation by voters and by most candidates generally complied with democratic standards, but the executive office participated in elections in a manner that has been labeled "flagrantly undemocratic."[15] 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.[16] 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.[17] 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.[18] 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.

The two most prominent political figures in the postindependence period are Ter-Pertrossian and Robert Kocharian. The former was president until 1998, when he resigned amid wide public dissatisfaction as a result of his failure to increase the standard of living and his willingness to negotiate with Azerbaijan over the territorial conflict in Nagorno-Karabakh.

---

15. Welt and Bremmer, "Armenia's New Autocrats," 78.
16. Welt and Bremmer, "Armenia's New Autocrats."
17. Diamond, *Developing Democracy*, 55.
18. Welt and Bremmer, "Armenia's New Autocrats."

Kocharian, who was elected to replace Ter-Pertrossian 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.[19]

The only item on the ballot for the 2003 elections was the presidential race, in which nine candidates ran. 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.[20] The other notable challenger was Artashes Geghamian, the last Soviet-era mayor of the capital city of Yerevan.

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, was seen as a resolute supporter of its independence. Ter-Petrossian's willingness to negotiate with Azerbaijan over the territory 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, in 2003 he also campaigned on the promise of economic stability, as did all of the candidates. Thus, the relative homogeneity

---

19. OSCE/ODIHR, *Republic of Armenia Presidential Election 19 February and 5 March 2003*, 3.

20. 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.

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, making it a relatively clean setup to evaluate the effects of international observers on election fraud committed by the incumbent government.

The first round of Armenia's 2003 presidential elections 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% of the vote in the first round of the single-district national election. The official first-round vote share for Kocharian was 49.48%, 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 thirty-five participating countries. The second round was observed by 193 short-term observers from twenty-one countries.

The Natural Experiment Research Design in Armenia

In the Armenian election, international observers were assigned to polling stations on election day using a method that I did not supervise but that approximates random assignment. 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 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."[21] Fraud (and therefore fraud deterrence) can occur at many points in the electoral process.

---

21. *Economist*, "Democracy, It's Wonderful," February 22, 2003.

However, it would be difficult to test for election day costs to a pseudo-democratic government if no fraud occurred.

Eyewitness reports from international observers, domestic observers, and journalists documented many varieties of fraud. The OSCE/ODIHR observed "significant irregularities" in more than 10% 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 inside polling stations of government officials 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.[22] 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 "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 website.[23]

The third and most important favorable feature of the 2003 Armenian elections for analysis is that the international observers were assigned in a way that approximates random assignment. Although the OSCE/ODIHR mission did not assign observers using a random numbers table or its equivalent, their method would have been highly unlikely to produce a list of assigned polling stations that were systematically different from the polling stations not visited. 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 prediction of voting patterns and the choice of polling stations based on those predictions.

In this particular election, the delegation leaders gave each team of the short-term observers a preassigned list of polling stations to visit on

---

22. OSCE/ODIHR, *Republic of Armenia Presidential Election 19 February and 5 March 2003*, 19.
23. Government-reported election results were made available online at http://www.elections.am by the Central Election Commission of Armenia.

election day. These lists were made with two objectives in mind: (1) to distribute the observers throughout the whole 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 lists 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 not be predictable by external observers, including government officials.

The individuals who made these lists had little knowledge of polling-station characteristics other than their general geographic location. This is critical for the validity of the natural experiment. If the assignment of observers had considered other variables that might be correlated with the performance of the incumbent candidate, then the assignment could not be considered "as-if" random. In addition, the discussion of Armenian politics indicated few observable characteristics of the population (such as socio-economic 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 has 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 to choose 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.[24] The fact that Armenian politics are not predictable along partisan or demographic lines underscores that this type of 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 eliminated much of the agency's influence on the individual observer teams, which, in the absence of a directive, could choose to visit polling sites based on their own selection criteria within their assigned geographic area. When observers are given leeway in choosing polling stations, the two most common alternative selection criteria (based on observation missions outside of the two discussed in this chapter) are to choose polling stations that are considered to be either convenient or interesting. Each of these decision

---

24. Even if this information were inaccurately communicated to me, if observers were more likely to visit stations they believe to be problematic, then this would dampen an observed effect observers have on fraud. For the reasons cited, however, this is an unlikely scenario.

criteria may create significant bias in the types of polling stations that are observed. This has been pointed out as a problem by several critics of election observation.[25] Observer teams that select "interesting" polling stations typically go to areas in which problems are expected; teams using this criteria 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 mission.

Observer teams that go to convenient areas are criticized for being electoral tourists. 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 in the natural experiment proposed here, particularly because a clever politician could recognize the tendency of observers to travel in 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.

In sum, the assignment of international observers to polling stations for both rounds of the 2003 presidential elections in Armenia can be characterized as approximating randomization. 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 because of 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 on election day.

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 or past election results. These data were not available for Armenia at the polling-station level for the first round, but as table 4.1 shows, observer distribution does not appear to follow a clear pattern that would predict Kocharian vote share. Coverage

25. Carothers, "The Observers Observed"; Geisler, "Fair?"

varies by region from a low of 28% of polling stations monitored in Aragatsotn to 70% of polling stations monitored in the capital city of Yerevan. The last column of table 4.1 suggests that much of this difference is due to voter density, because 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% of all polling stations are nonurban. In the first round of the election, international observers visited 38% nonurban polling stations, and 35% nonurban in the second round. Given that there are more voters in each urban polling station, observers covered nonurban areas extensively.

Because the same method of assigning observers was used in both rounds of the election, it is possible to check the second round's randomization against covariates from the first round of the election (table 4.2). Second-round observation, because it occurred three months later and was random, should not be systematically related to any first-round election outcomes. The round two presence of observers is compared with vote share and turnout in round one. These round-one outcomes should be equal between polling stations that were monitored in the second round and those that were not. As expected with near-random assignment, the presence of observers in the second round of the election is unrelated to voter turnout or to vote share for either of the leading candidates.

Table 4.1. Armenia round one observer coverage by region

Region	Total polling stations	Percentage monitored (%)	Average voters per polling station	Voters/total monitored polling stations
Aragatsotn	133	27.82	701	2,520
Ararat	137	53.28	1,355	2,543
Armavir	153	38.56	1,219	3,161
Gegharkunik	148	32.43	1,140	3,515
Kotayk	132	43.18	1,429	3,309
Lori	226	33.63	1,059	3,148
Shirak	273	25.64	907	3,537
Syunik	54	37.04	859	2,319
Tavush	80	28.75	1,152	4,007
Vayots Dzor	40	37.50	1,017	2,712
Yerevan	388	69.59	1,751	2,516

Table 4.2. Logistic regression of round two monitoring on background covariates

Variable	(1)	(2)	(3)
Round 1 Turnout	0.425		
	(0.608)		
Round 1 Kocharian Vote Share		–0.751	
		(0.832)	
Round 1 Demirchian Vote Share			0.114
			(1.083)
Constant	–0.878	–0.220	–0.636
	(0.690)	(0.653)	(0.418)
Observations	1,763	1,763	1,763

Notes: Robust standard errors in parenthesis, clustered by region. *Significant at 5%; **Significant at 1%

Data and Results

The central measurable effect of observers on election day fraud is to decrease the vote share for the incumbent. With 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 all else being 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. There was some between-round overlap in the polling stations visited by international observers. 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 had any lasting effect in the second round. Approximately 43% of polling stations were not observed in either round of the election, and about 21% were observed in both rounds.[26]

26. Outside of the Yerevan region (where polling stations were equally likely to be visited in both rounds) polling stations that were visited in round one were twice as likely to be visited again in round two.

The Effect of Monitors on Vote Share

The dependent variable is the vote share for the incumbent presidential candidate, Robert Kocharian, in the first and second rounds.[27] The results presented in table 4.3 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 6% in the first round (Model 1) and by about 2% in the second round (Model 2). 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 4.1 illustrates this difference with a kernel density plot of Kocharian's round-one vote share in monitored and unmonitored polling stations. Note the unusual distribution of vote share in unmonitored polling stations.

The results also suggest that the effect of monitoring in the first round of the election carried over into the second round. This type of effect could occur if those committing fraud anticipated that polling stations visited in the first round were more likely to be visited again. If the effect of international observers had lasting effects on fraudulent behavior, the polling stations that were monitored in the first round should also experience less fraud in the second round. If first-round observation had no effect on second-round

Table 4.3. Effects of observations on vote share for President Robert Kocharian

Variables	(1)	(2)	(3)	(4)	(5)
	Round 1 Vote share for Kocharian			Round 2 Vote share for Kocharian	
Observed R1	–0.059**	–0.059**	–0.055**		–0.040**
	(0.010)	(0.010)	(0.009)		(0.008)
Observed R2				–0.020*	
				(0.008)	
Urban			–0.012		
			(0.010)		
Near Nagorno-Karabakh			0.213**		
			(0.019)		
Constant	0.542**	0.542**	0.536**	0.693**	0.702**
	(0.007)	(0.007)	(0.009)	(0.005)	(0.005)
Observations	1,764	1,764	1,764	1,763	1,763
R^2	0.02	0.02	0.07	0.00	0.01

Notes: Robust standard errors in parentheses, clustered by region. *Significant at 5%; **Significant at 1%

27. The same tests for the other candidates are consistent with the conclusions drawn from these results.
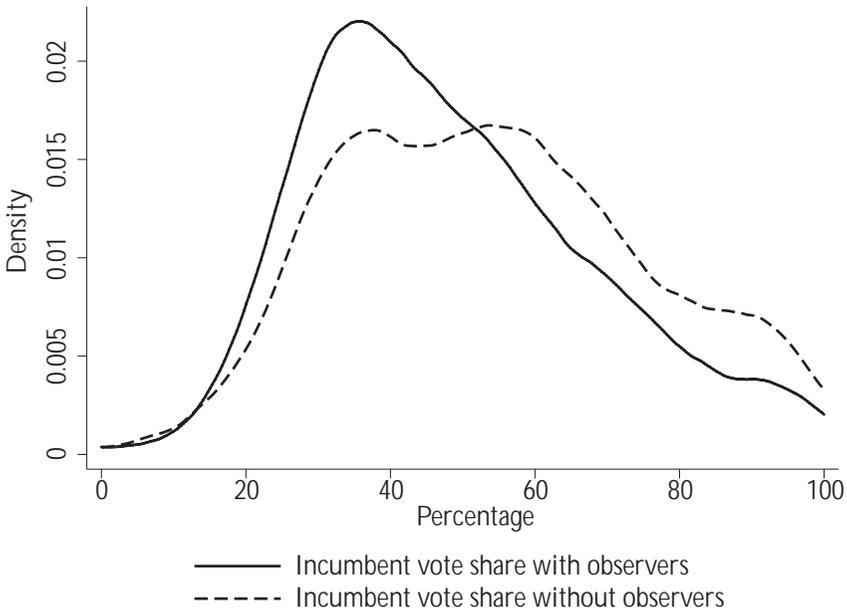
Figure 4.1. Round one vote share for incumbent in monitored vs. unmonitored polling stations

fraud, then the difference in the second-round incumbent vote share be-tween the groups in the first round should be close to zero. Model 5 in table 4.3 illustrates that a first-round visit by observers continued to have a clear effect in the second round. A first-round visit caused the incumbent to perform about 4% worse in the second round among polling stations that were not visited again in the second round. 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. Overall, the Armenia election shows that under conditions of widespread election day fraud, observers reduced the vote share for the incumbent, therefore making it more difficult or more costly for his government to steal votes. The magnitude of the effect of observers in the first round may have caused the second round runoff, because the national average deterrent effect was greater than the margin by which the incumbent failed to win the first round outright.

The 2004 Presidential Elections in Indonesia

Armenia and Indonesia differ in many ways, most obviously in size. Indo-nesia is one of the largest and most geographically diverse election-holding

countries in the world. There were fewer than 1,800 polling stations for the 2003 Armenian elections. In contrast, elections in Indonesia took place across more than 17,000 islands, approximately 155 million eligible voters, and more than 500,000 polling stations.

The 2004 presidential elections in Indonesia were the first direct presidential elections in the country's history. Legislative elections held in 1999 and in April of 2004 were widely considered successful given the size of the country and the country's newly democratizing status. Prior to these elections, the president was selected indirectly. The incumbent in the 2004 elections, Megawati Sukarnoputri, had been in office since her 2001 appointment by the People's Consultative Assembly. There were two rounds of the 2004 presidential election; due to the fact that first-round election results are unavailable, I focus here on only the second-round runoff between the incumbent candidate Megawati Sukarnoputri (commonly referred to as Megawati or Mega) and the leading challenger, Susilo Bambang Yudhoyono (commonly referred to as SBY). Expectations were high leading up to the 2004 elections, which were viewed as a crucial step in Indonesia's democratization. Many believed that the elections were likely to go well, and the most common concerns in advance of election day pertained to logistical factors and the administration of an election in such a large and diverse country. However, because of the scope of the election reforms leading to the 2004 elections and the recent transition to democratic institutions, some observers worried that the election could deteriorate into violence or fraud.[28]

For the 2004 presidential elections in Indonesia, I had the opportunity to attempt random assignment of international observers for the Carter Center's election day deployment rather than rely on "as if" random assignment as in the Armenian election. To my knowledge, this was the first attempt to randomly assign observers within the field of international election observation.[29] The case of Indonesia was selected because the opportunity to attempt random assignment of international observers was made available by the Carter Center. The introduction of randomly assigned international observers had been met with some skepticism by other practitioners. Although international election observation missions regularly

28. European Union, *European Union Election Observation Mission to Indonesia*; Carter Center, *The Carter Center 2004 Indonesia Election Report*.

29. Since that time, randomized assignment of international observers has been conducted by the Carter Center in Nicaragua (2006), by a Yale University student delegation participating in a U.S. Embassy mission in Mauritania (2007), and by the National Democratic Institute (NDI) in the 2006 Palestinian elections.

use randomization to assign international observers to vote-counting centers at the end of election day as part of a parallel vote tabulation,[30] random assignment of international observers during polling was thought unnecessary, logistically too difficult, or contrary to some of the other goals of election observation.[31]

Prior to the election, there were reports of "money politics" and other forms of intimidation, complaints related to restrictions placed on domestic election observers, as well as violations of laws restricting campaign activity. However, overall, the environment leading up to the presidential elections was guardedly optimistic, and observers hoped that the election would be carried out peacefully. Thus, the anticipated effect that international observers could have on election day behavior was moderated by the expectation that the election would be relatively clean. Clear-cut cases of blatant election day fraud would have made this a more straightforward baseline study of whether election observers change election quality, as in the Armenian case. Theoretically speaking, Indonesia was more complicated. Although many experts in Indonesia politics had concerns in advance of the election, blatant election day fraud—such as ballot box stuffing—was not expected. In designing this study, it was not clear in advance of the election which candidate would be more likely to benefit from the presence of observers. In countries that experience widespread election day manipulation, the incumbent party is frequently the primary sponsor, and as shown in the Armenia case, evidence suggests that observers can deter blatant election day manipulation. However, in Indonesia's 2004 election, the incumbent president had never stood for direct election to the presidency and did not have a reputation for carrying out widespread election day

30. The parallel vote tabulation, or quick count, provides an independent measure of the election results, within a margin of error, and is traditionally more reliable than exit polling. Observers (domestic or international) are assigned to a random sample of polling stations to directly observe the counting process, and they report the tallies from the vote count. Because the sample is random, quick counts typically provide very accurate estimations of the election results and thus guard against manipulation during the counting process; Estok, Nevitte, and Cowan, *The Quick Count and Election Observation.*

31. For example, one strategy for election monitoring is to send observers to the areas that are expected to have problems or to send observers to areas that would "benefit from seeing an international presence." These strategies create clear bias in the content of election day observations, but are perceived as politically important. (Personal conversations between the author and international election observation professionals from NDI, the EU, the OSCE/ODIHR, and The Carter Center.) Of course, it would be possible to randomize within regions that are expected to have problems in order to alleviate this concern.

manipulation. Additionally, ahead of the second-round runoff, Mega-wati had already lost the first round of the election to SBY and was not expected to win. In hindsight, the fact that the incumbent ran, allowed democratic elections, lost, and peacefully transferred power makes clear that she met my definition of a true democrat. Observers should not, therefore, have been costly to Megawati in terms of her election day performance.

Field Experimental Design in Indonesia

The Carter Center's mission for the second round of the election consisted of fifty-seven observers and twenty-eight observer teams, twenty-three of which participated in the randomization. The long-term election observers and the Jakarta-based staff of the Carter Center selected areas of Indonesia (primarily *kabupaten* and *kota*, or districts and cities) to which the Carter Center would send election observers. The selection of districts to be visited by the Carter Center was not random. In order for an area to be selected, it had to be accessible by car or aircraft within one day's travel time and had to have basic accommodations for the observer team that were judged as sufficiently safe.[32] There was also some effort made to avoid extensive overlap with the European Union election observation mission, as well as consideration for whether access was granted to areas in which foreigners are frequently prohibited from traveling, such as Banda Aceh, Ambon, and parts of Papua. For the participating teams, random assignment was applied within each district or pair of districts where Carter Center observers were sent.

The lists of villages and neighborhoods assigned to each participating Carter Center team were generated within each preselected geographic area using systematic random sampling (also called patterned sampling).[33]

32. Security concerns are relatively standard on election observation missions but were heightened in Indonesia because of recent Western-targeted bombings of hotels and the Australian embassy.

33. For a given block (city or district) to which a Carter Center team was assigned, a complete list of villages and neighborhoods was compiled. The total number of units within each block, or N_i, was sorted by an identification number that roughly identified the units geographically but was not otherwise organized in any systematic pattern. For each block, a target number of randomly selected units, n_i, was produced in negotiation with regional experts and the Carter Center staff, and for logistical reasons a greater proportion of selected units were allowed within some blocks. Given n_i, every kth unit was selected, with $k = N_i/n_i$ for all i blocks. Randomization requires that every unit within a given block has an equal probability of being selected. The first village chosen in the skipping pattern was selected arbitrarily from all villages within the block.

These lists were not released to anyone other than the Carter Center staff and the observer teams assigned to each area.

The unit of analysis in this study is the village/neighborhood. However, each village or neighborhood contained one or more polling stations, which observers were instructed to select using a method that approximates randomization.

To check the randomization of villages within each experimental block, I used logistic regression with assignment to the treatment group as the dependent variable. Because this was the first direct presidential election in Indonesian history, little historical precedent existed on which to base predictions of vote patterns for the 2004 elections. Only one background covariate is available at the village level: the total number of registered voters. Table 4.4 presents the results of the randomization check. Across all blocks, the null hypothesis—that assignment to the treatment group is not related to the number of registered voters—cannot be rejected. Also in table 4.4, when all blocks are pooled, assignment to the treatment group is unrelated to the number of registered voters, similarly indicating that there is no significant

Table 4.4. Logistic regression of assigned-to-treatment group on registered voters

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Registered Voters (1,000s)	–0.103 (0.215)	0.028 (0.019)	–0.001 (0.110)	–0.030 (0.102)	0.104 (0.136)	0.124 (0.074)	–0.078 (0.144)
Constant	–1.125* (0.471)	–1.704** (0.333)	–0.072 (1.231)	–1.173** (0.355)	–1.500** (0.451)	–1.645** (0.353)	0.345 (1.086)
Observations	90	163	23	186	117	136	45

Block	(8)	(9)	(10)	(11)	(13)	(14)	(15)
Registered Voters (1,000s)	0.175 (0.128)	0.006 (0.032)	–0.004 (0.043)	–0.021 (0.054)	–0.029 (0.075)	–0.058 (0.048)	0.160* (0.078)
Constant	–1.515* (0.674)	–1.495** (0.379)	–1.265** (0.202)	–1.231 (0.674)	–1.326** (0.513)	0.339 (0.801)	–2.304** (0.524)
Observations	46	156	243	42	31	24	103

Block	(16)	(17)	(18)	(19)	(20)	Pooled estimate	
Registered Voters (1,000s)	–0.007 (0.060)	–0.097 (0.223)	0.170 (0.235)	–0.230 (0.136)	0.016 (0.048)	0.011 (0.112)	
Constant	–1.363* (0.576)	–1.197* (0.530)	–1.645** (0.586)	–0.015 (0.496)	–1.527** (0.346)	–1.34** 0.259	
Observations	103	60	52	56	146	1822	

Notes: The "pooled estimate" includes dummy variables for each block (not reported). Standard errors in parentheses. *Significant at 5%; **Significant at 1%

difference in the number of registered voters between treatment and control groups.[34]

Table 4.5 summarizes the areas observed by Carter Center observers at the village level. Out of all villages in the visited regions, Carter Center observers were assigned to visit 482 villages, 95 of which were actually visited. Within these 95 assigned and visited villages, 147 individual polling stations were visited. Note that a small proportion of villages in the control group were visited.[35]

Data and Results for 2004 Indonesian Election

In the second round of the 2004 presidential elections, Susilo Bambang Yudhoyono and his running mate Jusuf Kalla were the leading candidates, having won 34% of the votes cast in the first round in a five-candidate field. The incumbent president, Megawati Sukarnoputri, won 27% in the first round. The runoff was held on September 20, 2004. SBY won the presidency with 61% of the vote. Due to the size of the country, Carter Center observers could not spread out across the entire country, as was attempted by the OSCE/ODIHR mission in Armenia. The randomization within blocks prevents generalization across the country, but within-block comparisons remain valid. Estimating the average effect across included blocks is also possible, although it is more complicated and should not be confused with randomization across the whole country.

Government-reported unofficial election results for the total number of votes cast for each candidate and the number registered voters for all villages were recorded in the second round of the 2004 presidential election. Polling station-level data for the same regions were also collected but are not analyzed here.[36] The unofficial results were made public by the Indonesian *KPU* (the general elections commission) for most of the country. These aggregate results were uploaded by regional election officials to a central government-run website and are subject to the usual disclaimers about unofficial election results. Unfortunately, data were incomplete for three of the districts where teams from the Carter Center were deployed

---

34. Pooling all blocks is more complicated because of variation in the size of the blocks.

35. For some teams, visiting control group villages or neighborhoods was accidental and resulted from visiting polling stations near the border between urban neighborhoods. Other teams encountered logistical (usually transportation related) problems that caused them to choose to visit villages outside of their assigned list. This information is only available anecdotally and was not coded in the dataset.

36. Data were downloaded from the KPU website, http://tnp.kpu.go.id/ (accessed March 2007).

Table 4.5. Carter Center observation coverage of villages in Indonesia

Study or Region	Province	District or city	Assigned treatment group	Assigned control group	Treated	Treated in control group	Total voters
1	Nanggroe Aceh Darussalam	Kota Banda Aceh	19	71	3	1	173,265
2	East Java	Kota Surabaya	34	129	3	0	2,078,486
3	Nusa Tenggara Barat	Kota Mataram	11	12	4	0	241,483
4	East Java	Sampang	41	145	5	0	569,216
5	Bali	Tabanan	27	90	6	0	325,701
6	East Java	Situbondo	32	104	10	0	488,633
7	DI Yogyakarta	Kota Yogyakarta	20	25	4	0	327,873
8	East Java	Kota Kediri	15	31	11	0	200,137
9	North Sumatera	Kota Medan	30	126	5	2	1,525,526
10	Riau	Kampar, Kota Pekan Baru	53	190	5	0	740,924
11	East Kalimantan	Kota Samarinda	8	34	4	2	453,693
13	Central Kalimantan	Kota Palangka Raya	6	25	1	4	123,596
14	West Kalimantan	Kota Pontianak	9	15	4	3	371,780
15	West Sumatra	Kota Padang	20	83	4	1	525,422
16	South Sumatra	Palembang	20	83	11	1	906,169
17	North Sulawesi	Kota Bitung	12	48	3	4	120,637
18	North Maluku	Kota Ternate	11	41	5	1	95,771
19	Maluku	Kota Ambon	18	38	4	1	192,097
20	South Sulawesi	Kota Makassar	28	118	2	2	812,977
	Total		414	1408	94	22	10,273,386

in the second round: Mimika, Kupang, and Manokwari. These regions (and the three corresponding Carter Center teams) were dropped from the analysis, leaving twenty experimental blocks.

Another important issue in analyzing the results pertains to treatment rates. Table 4.5 shows that the treatment rates, or the percentage of villages randomly assigned to be monitored that were, in fact, monitored, vary considerably across blocks. Several issues must therefore be considered in the analysis of the experiment. There is one block in particular (block 12, in Cianjur) in which there was no experiment to speak of, with monitors going only to less than 2% of both the treatment and control groups. It is also a block with an unusually large number of villages, representing a significant portion of all villages in the "failure to treat" category. The reason for the implementation failure in this block was that the team of monitors assigned there did not attempt to comply with the assigned list of villages, a decision that was not influenced by the characteristics of the block. Although I present the summary data for this block in table 4.5, I exclude it from the remainder of the analysis.

Table 4.6 presents aggregate summary statistics for the 1,822 remaining village-level observations. I downloaded, compiled, and merged the international observer data with the unofficial election results. All comparisons include only districts in which Carter Center observers were deployed, where they participated in the randomization, and where village-level elections results were reported for the entire district.

Another issue stems from treatment rates. Given some failure to treat, it is tempting to move the untreated villages into the "control" group and

Table 4.6. Summary statistics for all available village-level variables

Variables	Observations	Mean	Standard deviation	Minimum	Maximum
Observed	1,822	0.064	0.224	0	1
Sample	1,822	0.227	0.419	0	1
Megawati (total votes)	1,822	1,394	1,662	1	22,494
SBY (total votes)	1,822	2,486	2,538	5	19,618
Overall Turnout	1,822	0.715	0.117	0.1	1
Ballots Received	1,822	5,645	5,826	35	59,567
Valid Ballots	1,822	3,880	3,848	6	42,112
Invalid Ballots	1,822	83	131	0	1,582
Extra Ballots	1,822	6	26	0	345
Damaged Ballots	1,822	43	276	0	5,923
Ballots Not Used	1,822	1,645	2,071	0	16,612
Total Registered Voters	1,822	5,639	5,821	35	59,567

simply compare the subset of villages visited by international observers with those that were not. This comparison yields biased estimates when nontreatment is correlated with the dependent variable and does not take advantage of the randomization. To illustrate with this experiment, because it is plausible that some villages were more difficult for observers to locate than others and that this "findability" could be related to voting behavior, it cannot be assumed that the factors determining which villages were actually monitored were completely random. All comparisons must therefore exploit the randomization by using assignment to the treatment group rather than actual treatment. As the central dependent variable of interest, I use the total number of votes cast for Megawati (logged).

The most straightforward method requiring the fewest assumptions is to estimate the intent-to-treat (ITT) effect within each block for one of the candidates. Here, the estimated ITT effect within each block i is the average difference between treatment and control groups in incumbent performance in villages. It would be possible to estimate the ITT effect without accounting for any other observed differences between villages, but regression analysis allows the inclusion of covariates that reduce the unexplained variance in vote share between villages.

I calculate the ITT effect using ordinary-least squares (OLS) regression. The central dependent variable is the performance of the incumbent candidate, measured as the total number of votes cast for Megawati in each village (logged). To account for varying village sizes, an additional independent variable measuring the total number of registered voters in the village (logged) is included in the regression. This basic model can therefore be expressed as:

$$\log (Y_j) = \beta_0 + \beta_1 T_j + \beta_2 \log (X_j) + \mu_j,$$

where Y is the total votes vast for Megawati in village j, $T_j = 1$ if the village was assigned to the treatment group, X is a variable representing the total number of registered voters in the village, and μ represents unobserved causes of votes for Megawati.

Table 4.7 presents estimates of the effect of being assigned to the treatment group within each regional block. Even given the relatively low rate of assigned villages that were actually visited (as shown in table 4.5), assignment to the treatment group is associated with *improved* performance for Megawati in 15 out of the 20 blocks, a result that is unlikely to be due to chance. I also provide a pooled estimate with fixed effects for each experimental block. Note that in these models, *Treatment Group* is a measure of the intent to treat the village, not the actual presence of observers

Table 4.7. Estimated effects of intent to treat on total votes for Megawati (ordinary least squares, ln)

Block	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment Group	0.14	0.02	0.07	0.24*	0.02	0.15*	0.11
	(0.11)	(0.05)	(0.15)	(0.10)	(0.06)	(0.07)	(0.06)
Registered Voters	0.97**	1.04**	0.88**	1.36**	0.85**	0.94**	0.90**
	(0.06)	(0.03)	(0.19)	(0.08)	(0.05)	(0.05)	(0.10)
Constant	–2.16**	–1.65**	–0.69	–4.34**	0.68	–0.81*	–0.50
	(0.42)	(0.26)	(1.78)	(0.60)	(0.40)	(0.41)	(0.89)
Observers	90	163	23	186	117	136	45
R^2	0.77	0.90	0.51	0.64	0.71	0.74	0.66

Block	(8)	(9)	(10)	(11)	(13)	(14)	(15)
Treatment Group	0.11	–0.07	0.03	0.06	0.19	–0.11	0.03
	(0.06)	(0.09)	(0.09)	(0.10)	(0.21)	(0.15)	(0.10)
Registered Voters	0.91**	1.04**	0.86**	0.87**	0.87**	1.14**	0.80**
	(0.05)	(0.03)	(0.04)	(0.05)	(0.05)	(0.10)	(0.06)
Constant	–0.29	–2.00**	–0.88**	–0.52	–0.42	–2.98**	–0.86
	(0.39)	(0.29)	(0.28)	(0.45)	(0.36)	(0.94)	(0.52)
Observers	46	156	243	42	31	24	103
R^2	0.90	0.87	0.70	0.89	0.91	0.88	0.63

Block	(16)	(17)	(18)	(19)	(20)	Pooled estimate	
Treatment Group	–0.04	–0.14	0.19	0.04	0.16	0.065*	
	(0.04)	(0.07)	(0.17)	(0.18)	(0.15)	(0.026)	
Registered Voters	0.89**	0.83**	0.51**	0.99**	0.92**	0.94**	
	(0.03)	(0.03)	(0.09)	(0.08)	(0.07)	(0.014)	
Constant	–0.08	0.28	1.69*	–0.95	–1.80**	–1.95**	
	(0.25)	(0.25)	(0.64)	(0.67)	(0.55)	(0.11)	
Observers	103	60	52	56	146	1822	
R^2	0.91	0.92	0.42	0.73	0.58	0.87	

Notes: Pooled estimate includes block fixed effects. Standard errors in parentheses.
*Significant at 5%; **Significant at 1%

on election day. Because *Treatment Group* is dichotomous, the coefficients represent the percentage of change in total votes cast for Megawati given that *Treatment Group* changes from zero to one and all else is constant. In the pooled estimate in table 4.7, assignment to the treatment group causes a 6.5% positive change in the number of votes cast for Megawati. To put this number in context, the average number of votes cast per village for Megawati is 1,394, and assignment to the treatment group is associated with an average increase of about 91 votes.

This estimation of the intent-to-treat effect is the least biased estimate, but it does not count for actual treatment rates. Because relatively few assigned villages were visited by observers, the estimated effect of observers on villages that were visited is much larger. The average treatment rate across all blocks in the experiment was about 23%: a little more than one out of every five villages assigned to observers was actually visited on election day. Given low treatment rates, as in this case, the figures in table 4.7 likely underestimate the magnitude of the observers' effect.

Following previous applications in field experiments, I use instrumental variable techniques to estimate the magnitude of the treatment effect.[37] Very generally, this estimate can be understood as the ITT effect divided by the actual treatment rates.[38] Table 4.8 presents estimates from two-stage least-squares regression (2SLS). Using 2SLS, for an instrument to be valid it must be correlated with the actual treatment but not correlated with the error term in the model. Assignment to the treatment group of villages within a region is random, and there is therefore no reason that it should be correlated with the error term. Actual treatment, or being visited by international observers, is a function of a village being assigned to the treatment group. When the actual visit by observers to a given village is used as an explanatory variable, assignment to the treatment group satisfies the conditions for a valid instrument.

Like the ITT estimates presented in table 4.7, total registered voters (logged) are included as an independent variable. The difference between the results presented in tables 4.7 and 4.8 is that the table 4.8 results account for the fact that observers did not visit all assigned villages. Consistent with expectations, the estimate of the effect of observers on treated villages in the treatment group (also called the treatment-on-treated effect) is substantively larger and is associated with a +32% change in votes cast for Megawati, an average increase of 446 votes per treated village.[39] Note that this is an estimate of the size of the effect if observers had visited all villages in the assigned-to-treatment group.

The same estimates using vote share rather than votes cast were also conducted, but results are not presented here because the varying size of

37. Angrist, Imbens, and Rubin, "Identification of Causal Effects Using Instrumental Variables"; Gerber and Green, "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment."

38. This use of experimental treatments as instrumental variables is described in greater detail by Gerber and Green 2000, "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment," 657–658.

39. Note that both tables 4.7 and 4.8 present pooled estimates in which all blocks are combined.

Table 4.8. Estimated effect of observers on total votes for Megawati in observed villages (two-stage least-squares regression)

Block	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated (Observed)	1.04	0.22	0.18	1.96*	0.09	0.50*	0.57
	(0.79)	(0.59)	(0.42)	(0.98)	(0.25)	(0.22)	(0.38)
Registered Voters (ln)	1.00**	1.04**	0.80**	1.37**	0.85**	0.93**	0.90**
	(0.06)	(0.03)	(0.27)	(0.09)	(0.05)	(0.05)	(0.12)
Constant	−2.39**	−1.66**	0.03	−4.40**	0.69	−0.69	−0.48
	(0.49)	(0.26)	(2.40)	(0.69)	(0.40)	(0.42)	(1.07)
Observers	90	163	23	186	117	136	45
R^2	0.76	0.90	0.50	0.53	0.71	0.73	0.52

Block	(8)	(9)	(10)	(11)	(13)	(14)	(15)
Treated (Observed)	0.16	−0.49	0.32	0.13	2.41	−0.32	0.17
	(0.09)	(0.61)	(0.93)	(0.23)	(4.08)	(0.43)	(0.55)
Registered Voters (ln)	0.88**	1.05**	0.86**	0.86**	0.45	1.21**	0.80**
	(0.05)	(0.04)	(0.04)	(0.05)	(0.71)	(0.12)	(0.07)
Constant	−0.06	−2.08**	−0.86**	−0.47	2.16	−3.63**	−0.84
	(0.43)	(0.32)	(0.28)	(0.47)	(4.27)	(1.04)	(0.54)
Observers	46	156	243	42	31	24	103
R^2	0.89	0.86	0.69	0.88	0.80	0.88	0.62

Block	(16)	(17)	(18)	(19)	(20)	Pooled estimate
Treated (Observed)	−0.08	−0.72	0.46	0.19	2.89	0.32*
	(0.08)	(0.46)	(0.43)	(0.89)	(3.24)	(0.13)
Registered Voters (ln)	0.89**	0.94**	0.48**	0.98**	0.91**	0.93**
	(0.03)	(0.08)	(0.10)	(0.09)	(0.08)	(0.015)
Constant	−0.04	−0.46	1.91**	−0.88	−1.76**	−1.87**
	(0.25)	(0.52)	(0.71)	(0.69)	(0.67)	(0.12)
Observers	103	60	52	56	146	1822
R^2	0.91	0.87	0.37	0.71	0.40	0.87

Notes: Instrumented variable: Village visited by international observers. Exogenous variable: Village in treatment group. Pooled estimate includes block fixed effects. Standard errors in parentheses. *Significant at 5%; **Significant at 1%

villages and blocks complicate the analysis. These estimates using alternative specifications of the dependent variable are similar.[40] As an additional check, the estimates presented in Tables 4.7 and 4.8 were also conducted

40. Additionally, I estimated all models with a variable indicating the presence of EU observers. EU observers were not randomly assigned. Out of the 1,822 villages included in this study, EU observers visited 61. Of these 61 villages, 4 were in the treatment group and also visited by Carter Center observers, and 6 of which were in the assigned treatment group but not visited by Carter Center observers. The inclusion of this variable has minimal influence on the sign and significance of the (randomized) Carter Center observation variable.

on total votes cast for SBY and are available on request. There is no significant relationship between observer presence and the performance of the winning candidate, SBY.

Overall, the results of the Indonesia field experiment show that the incumbent candidate performed better and the challenger performed about the same in villages and neighborhoods assigned to be monitored by Carter Center observers. This result was not anticipated and highlights a central advantage of using field experimental methods: the possibility that they can reveal effects that are not anticipated by scholars or practitioners.[41] Such a surprising result nevertheless requires some speculative explanation and analysis of the unique circumstances surrounding this election. Why might the presence of observers increase votes cast for Megawati but not decrease votes cast for SBY? Why did observers influence what was widely viewed as a democratic election?

The reports of international observers, journalists, and analysts suggest several possible explanations. Although all major international observer organizations judged the observed problems with the election to be insignificant, a number of irregularities were documented and described in the postelection reports of international observers. The most plausible explanation for this finding stems from the early closing of polling stations. The official election day was from 7:00 a.m. to 1:00 p.m., but after the first round of the presidential election, the *KPU* ruled that polling stations could close after 11:30 a.m., provided that all eligible voters had voted. If this rule was followed correctly, it should not have produced significant problems, and only those polling stations that reached 100% turnout should have closed early. Reports suggest, however, that a number of polling stations closed before all eligible voters had cast a ballot and well before the earliest legal closing time of 11:30.[42]

During the course of their observation, many Carter Center observers announced or implied that they could return later in the day to observe the closing. The presence of observers could have influenced the decision by election officials to close early by making it more likely that polling stations in visited areas would stay open until the mandated time. If Megawati supporters were less likely to turn out to vote without being mobilized to do so by party representatives or election officials, correctly following the regulations surrounding the length of election day would have disproportionately benefited Megawati voters. Local party officials

41. Banerjee and Duflo, "The Experimental Approach to Development Economics."

42. European Union, *European Union Election Observation Mission to Indonesia*, 58; Carter Center, *The Carter Center 2004 Indonesia Election Report*, 63.

would have more time to mobilize voters, and poll workers would have had greater incentive to prove that all voters had cast a ballot so that they could close early without violating electoral regulations. One potential explanation is therefore that nonobserved villages were more likely to close before less-motivated or reluctant voters had shown up and were less likely to follow the electoral regulations about staying open until 1:00 p.m. or until all registered voters had cast a ballot.

Several additional pieces of evidence support the possibility that Megawati supporters were more reluctant to turn out and also suggest that she was not in control of the party or state machinery that would have been required to engage in widespread election day fraud. First, her party performed poorly in the April legislative elections and in the first-round presidential elections. Second, in the weeks leading up to the runoff election, it was widely speculated in the media that she would lose, with public opinion polls from several organizations predicting support for SBY at about 60% and support for Megawati at around 29%.[43] Third, although Megawati had some incumbency advantages, including the ability to make public appearances throughout the country outside of the legal three-day campaign period, her support from several prominent parties was unstable. For example, Megawati was endorsed by the powerful Golkar party, which won the April 5 legislative elections and which possessed well-developed local party machinery that could have been used to mobilize the vote for Megawati. But several weeks before the election, national and local party leaders split publicly over the decision to endorse Megawati, and before the election, analysts predicted that "Golkar will not be able to fully bring its formidable party machinery behind Megawati."[44] Postelection polling revealed that the vast majority of Golkar voters who cast a ballot voted against their party's endorsement and for SBY.[45] Relative to incumbent presidential candidates in other countries, Megawati's election day advantage was minimal.

If Megawati supporters were reluctant to turn out, she should have performed better in those areas in which turnout was higher. Although it is not conclusive evidence, scatter plots of votes cast for Megawati versus turnout across all 1,822 villages included in the experiment illustrate

43. "Indonesia's Megawati Heading for Defeat, Two Polls Show," Associated Press Worldstream, September 15, 2004.

44. "Golkar Party Leaders Split as Internal Rift Deepens," *The Jakarta Post*, September 1, 2004; "What Lies Ahead After Indonesia's Election," United Press International, September 14, 2004.

45. Liddle and Mujani, "Indonesia in 2004."

that Megawati does somewhat better in villages with higher turnout and SBY does worse, on average, in villages with higher turnout. These comparisons do not prove that increasing turnout would have necessarily increased votes for Megawati, but they are consistent with the idea that Megawati's supporters were more reluctant to turn out and that her performance increased when voter mobilization increased.

The results presented here show a clear difference between observed and unobserved villages, but they are subject to interpretation. The most likely explanation for this finding, in my view, is that observers made polling station officials more likely to follow electoral regulations and therefore caused visited polling stations to stay open later than they would have if observers had not visited. Given that the election was expected to be relatively free of election day irregularities, the fact that any significant effect of observers was found is noteworthy. This result does not imply election fraud. If widespread election fraud by one candidate had taken place, and this fraud was deterred by observers, the cheating candidate should have performed worse in areas that were observed. Even though Megawati benefited from observers, the results do not show that SBY performed significantly worse when observers were present, as would be expected if observers reduced ballot box stuffing or other forms of direct election fraud. Rather, I argue that election officials were more likely to follow the letter of the election law pertaining to closing time after having been visited by international observers.

The Carter Center mission concluded that "voters were able to exercise their democratic rights in a peaceful atmosphere and without significant hindrance."[46] The results presented here do not contradict this conclusion.

Inviting Observers Changes Behavior

Taken together, the micro-level results from elections in Armenia and Indonesia show that observers can be more costly to pseudo-democrats than to true-democrats, at least in terms of their election day performance. Particularly for pseudo-democrats who commit election day fraud, this chapter demonstrates that observers can make stealing elections more expensive or more difficult by directly reducing fraud through their presence

---

46. Carter Center, *The Carter Center 2004 Indonesia Election Report*, 13. See also the EU final report "*European Union Election Observation Mission to Indonesia.*"

in polling stations. Even in the case of Indonesia, in which the election was viewed as a successful exercise in democracy, the presence of observers had a direct and measurable effect on election day behavior, perhaps inducing increased compliance with government regulations and having an unanticipated effect on turnout among the more reluctant supporters of the incumbent. The Indonesia study demonstrates that the incumbent government was not significantly harmed by the presence of observers, and even benefited from their presence for idiosyncratic reasons.

More generally, these findings indicate that inviting observers can be consequential for leaders who invite monitors and attempt to steal the election and that these costs do not necessarily exist for incumbent true democrats. I also showed that competitive elections outcomes are positively correlated with the presence of observers. In chapter 3, I showed that pseudo-democrats face consequences if they refuse to invite observers or if they are caught cheating by international observers. As I will show in the next chapter, in addition to potentially reducing blatant election day fraud, many leaders now invite observers and work to avoid international criticism, making the changing forms of election manipulation another consequence of internationally monitored elections.