

Electoral Observers: The Implications of Multiple Monitors for Electoral Integrity

Ursula Daxecker and Gerald Schneider

Forthcoming in Pippa Norris, Richard W. Frank, and Ferran Martinez I Coma (Eds.): *Advancing Electoral Integrity*. Oxford University Press.

The 2004 legislative elections in Kazakhstan were monitored by three international organizations engaged in election observation. One might expect that the supervision of electoral processes by multiple organizations would increase the amount of information on election quality available to voters, but the three participating organizations came to quite different conclusions regarding the credibility of the elections. The Council of Europe (COE) and the Organization for Security and Cooperation in Europe (OSCE) criticized the election, alleging that serious irregularities occurred and condemning the electoral process. Yet in contrast, the Commonwealth of Independent States (CIS) delegation concluded that minor shortcomings in the electoral process did not affect either voters' free choice or the election results and came to an overall positive assessment.² Clearly, additional reports did not necessarily improve the quality of available information about this election's regularity. Kazakhstan is not an isolated case. On the contrary, incumbents concerned about negative observer reports can often expect a beneficial cacophony of friendly and critical observer opinions if they invite more than one observer group.³ Inviting organizations of varied quality would in this perspective be strategically advantageous by allowing governments under domestic pressure to highlight more positive assessments and selectively cite more critical ones – even if some observer organizations find serious irregularities. For instance, the Kazakh Electoral Commission's website for the 2004 elections includes a long quote by the head of the CIS delegation endorsing the election and cites more favorable portions from the OSCE and COE statements and reports, but it omits the more critical elements from these reports.⁴ In these elections, inviting a mix of observer organizations appears to have worked in the favor of the government. Despite allegations of major fraud and irregularities by opposition parties and some monitoring organizations, election results giving President Nazarbayev's party more than 60 % of the votes were left unchanged, and popular protests failed to materialize.

The 2000 presidential elections in Peru, conversely, show that the incumbent's gamble with the invitation of additional observers can also backfire. This case demonstrates how agreements on fraud between international election monitors bolster the legitimacy of their reports, thereby increasing the pressure on incumbents to rerun elections or even step down. All three observation missions invited to monitor the Peruvian elections were unanimous in condemning the elections as

flawed and falling short of international standards.⁵ Widespread protests over blatant election fraud followed, ultimately resulting in President Fujimori's resignation and his escape to Japan.⁶ While admittedly anecdotal, these cases illustrate how incumbents can benefit from inviting the right mix of friendly and critical organizations, but they also illustrate how multilateral monitoring by reputable organizations can reinforce the credibility of their assessments, help bring about democratic reforms, or support the consolidation of democracy.

Election observation of the kind conducted in Kazakhstan and Peru is a key new norm in international politics that has altered the strategic calculus of governments and opposition groups alike. Two encompassing monographs have recently focused on the diffusion of this practice, arguing that norms of democracy promotion have rendered an increasing number of leaders interested in inviting observers regardless of whether they are truly interested in establishing democracy.⁷ A growing number of formal models look through the lenses of decision or game theory how potentially fraudulent incumbents can profit from election monitoring even when they risk being caught cheating.⁸ These studies focus on how election monitoring provides information to domestic actors. However, they do not treat the observers as strategic actors whose reputation depends on the accuracy of the reports and the outcome of the election monitoring process. Combining the supply and the demand-side analysis of election monitoring, we examine the conditions under which asking for additional monitoring reports possibly offsets the informational gains that election observers provide to the domestic actors.

The formal literature on election monitoring suggests that voters and the opposition may perceive the invitation or acceptance of a less reputable monitoring mission as a sign of weakness.⁹ Our argument draws on this reasoning and implies that desperate leaders who face a high risk to be accused of election fraud might nevertheless accept the stigma of electoral manipulation. The tendency of trying to offset a possibly negative report through a more supportive one should, in our view, be particularly pronounced in institutionally weak, corrupt, and economically dependent countries. We expect that incumbents who invite the right combination of monitors are less likely to face post-election unrest.

We test these complementary claims about the causes of complex monitoring and its consequences using a new dataset for international election monitoring missions from 1980-2004 described by Judith Kelley and Kirik Kolev.¹⁰ We first model the causes of multilateral monitoring with multinomial logit models that distinguish the number and quality of invited monitoring organizations. The statistical evidence shows that leaders with a history of flawed elections and countries that are highly dependent on foreign aid are more likely to invite or accept a combination of low and high-quality monitors. Using evidence from these models in a matching procedure that

helps account for endogeneity concerns, we proceed to examining the consequences of multilateral monitoring invitations. We find that incumbents who invite a mix of lenient and critical organizations can successfully avoid the costly consequences of cheating documented in existing research.

Regime Complexity in International Election Monitoring

Election monitoring is a growing international practice that has received ample theoretical and empirical scrutiny in recent years. International organizations have become increasingly involved in the supervision of electoral processes, and this growing interest in elections documents itself both in the number of monitored elections and the number of organizations present at elections. Susan Hyde¹¹ and Judith Kelley¹² show that rate of internationally observed elections has increased markedly, from less than 10 % of elections being monitored before the late 1980s to approximately 80 % of elections today. However, not all elections are monitored. As Pippa Norris, Jørgen Elklit, and Andrew Reynolds point out in Chapter Three of this volume, the selection mechanisms at work result in ‘floor’ and ‘ceiling’ effects. Hence, monitoring organizations do not tend to send observers to countries in which they expect to encounter few problems, and they also avoid elections in which security threats are high. Finally, it is unlikely that the most authoritarian states invite them to monitor the vote.

Yet in addition to a higher rate of monitored elections, a growing number of elections are also observed by more than one organization. As Judith Kelley notes,¹³ efforts by Western states to give primary responsibilities in election monitoring to the UN after the end of the Cold War failed, and more than 20 nongovernmental and intergovernmental organizations are or have been active in international election monitoring as a result.¹⁴ Since 1975, approximately half of all observed elections were monitored by more than one international organization.¹⁵ Similar to the rate of monitored elections, this pattern has become more pronounced over time. Before 1990, less than 20% of elections were observed by more than one monitoring mission; now over 60% of observed elections are monitored by two or more organizations.¹⁶

Importantly, these organizations differ significantly in their willingness and ability to assess election quality. The election monitoring arms of the African Union (AU), the South African Development Community (SADC), or the CIS, for example, have a reputation for (almost) never criticizing even highly problematic elections.¹⁷ Some argue that the CIS monitoring organization, for example, was created by regional governments specifically interested in counteracting the frequently critical reports by the OSCE for post-Soviet elections.¹⁸ Arguably, even reputable organizations may at times fail to adequately assess the quality of elections. For example, as Chapter

Two illustrates, the Carter Center mission observing the 2012 presidential elections in Egypt was hampered by substantial restrictions imposed by the electoral commission. Presumably, the emergence of this 'shadow market' could benefit incumbents who face pressure to invite monitors, but at the same time want to avoid the punishments that can follow critical reports.¹⁹ The foreign aid literature suggests along these lines that donor fragmentation decreases the effectiveness of development assistance and the administrative capacity of the receiving state.²⁰

The implications of multilateral election monitoring have received relatively little attention. While recent research has improved our understanding of the intended and unintended consequences of election observation, researchers have primarily compared situations in which observers were present to those in which they were not.²¹ With regard to the benefits of election monitoring, studies have shown that international election observation, in particular observation by reputable organizations, can deter fraud and raise the costs of cheating for incumbents because of post-election punishment. For example, Susan Hyde²² and Judith Kelley²³ find that the presence of reputable international observers can help reduce fraud and thus improve election quality. Regarding the increased costs of cheating, Donno²⁴ argues that the presence of high-quality observers in fraudulent elections mobilizes opposition and citizens and therefore increases the chance of punishment by regional organizations, which she supports with evidence from Latin American elections. Similarly, Susan Hyde and Nikolay Marinov²⁵ expect that information on fraud provided by credible observers increases the probability of post-election protests, which could lead to costly consequences such as the ousting of incumbents or require the use of violence to contain anti-regime protests.²⁶ Also drawing on an informational logic, Ursula Daxecker²⁷ finds that fraudulent elections monitored by reputable international observers have a greater likelihood of post-election violence.

With regard to the unintended consequences of monitoring, Judith Kelley²⁸ documents that international monitors at times legitimize flawed elections. Other research has shown that the growing international attention to electoral processes can induce shifts in the manipulative strategies used by incumbents, leading to increasing pre-electoral fiscal manipulation, lower quality of governance, or a higher rate of boycotts in elections where monitors were present.²⁹

Yet while these studies have produced valuable information on the domestic and international implications of the presence of (reputable) monitors compared to their absence, they do not examine whether or how the growing number and differing quality of organizations engaged in monitoring could affect these findings, nor do they evaluate whether some incumbent governments might strategically invite more than one observer organization or a mix of observers. One notable exception is Judith Kelley³⁰ who argues that the density of international election

monitors has potential costs, such as coordination problems or competition between monitoring organizations, but also benefits, such as the mutual reinforcement of democratic norms or the availability of multiple institutional channels to ‘help avoid deadlock and paralysis.’³¹ Although Judith Kelley³² does not examine these scenarios with statistical analysis or case studies, she presents several illustrative examples that support the presence of costs and benefits in the international election monitoring regime. In our argument, developed in more detail below, we will outline the conditions under which the invitations of multiple organizations, particularly a mix of low and high-quality organizations, could be exploited by incumbent governments and thus undermine the benefits provided by the international regime on election monitoring. Conversely, our argument implies that multilateral monitoring by credible organizations can help improve electoral integrity by either deterring cheating in the first place, or by punishing incumbents for cheating through the incidence of post-election unrest.

The Selection and Impact of Second Election Monitors

Formal theoretical work has analyzed how incumbent governments might profit from the reliance on election monitors or similar mechanisms like the appointment of an electoral commission. Beatriz Magaloni analyzed the credibility dilemma an autocrat faces vis-à-vis an opposition that might challenge any election regardless of whether it was clean or fraudulent.³³ However, delegating the organization of an election solves the dilemma that announcing orderly elections is not credible. Andrew Little shows along these lines that inviting election monitors and cheating in front of them is not paradoxical because it lowers citizens’ expectation about the level of fraud that will accompany an election.³⁴ Modeling the interactions between the incumbent and the citizens as a ‘global game’ and thus a specific type of limited information model, Little argues that incentives to cheat linger on as long as the monitoring is imperfect, but beliefs that massive fraud has occurred increase the risk of mass protests. Milan Svolik and Svitlana Chernykh similarly examine election observation as an information-gathering device, focusing on the observer’s concern, labeled the ‘monitor’s restraint,’ of falsely rejecting a claim by the incumbent that she has won the election.³⁵ More particularly, their model demonstrates that appointing biased observers might be helpful as a rejection of the incumbent’s claim of an election victory by highly restrained monitors is much more informative for the opposition than the equivalent opinion of less constrained observers.³⁶

Susan Hyde and Nikolay Marinov support the contention advanced by these recent game-theoretic models that election monitoring is largely an information device and helps dissatisfied citizens to coordinate their protests.³⁷ However, governments have largely had a free hand in the

selection and appointment of monitors, and they might have incentives to accept one or more additional observer missions in case that they fear a particularly negative assessment by an international monitor. As Judith Kelley shows, appointing multiple election observers has become more frequent. In her view, ‘the ability of governments to manipulate the election monitoring experience increases as the number of organizations available for monitoring grows.’³⁸ Nevertheless, appointing a second or third monitor – or portraying these appointments as an act by the international community – creates the risk that these reports reinforce each other and that an accumulation of negative election assessments will force the incumbent out of office. However, contradicting evaluations by the international community will be beneficial and help the incumbent to avoid post-election protest and unrest.

Election observers should theoretically only be concerned with the regularity of the election process. However, they might also have previously formed contacts with either the government or the opposition, which may bias them favorably towards one of the candidates or parties. Election observers are therefore strategic actors who have to carefully evaluate whether it is worth taking the risk of siding with the wrong side. Whatever the bias and personal interest of an election observer in the outcome of an election might be, she faces the double risk of committing type I (accusing an honest incumbent of cheating) or type II errors (supporting a cheating incumbent). Whitewashing a stained incumbent through a rosy election observer report diminishes the monitor’s reputation. However, some election observers might not care about such a loss very much, as they might want to continue to have close contact with the incumbent after the elections. The Kazakh example introduced above where the CIS provided a supportive statement of Nazarbayev’s party is a telling example that not only the demand for electoral monitoring but also that its supply need to be considered. In other words, electoral observation should be analyzed like general international conflict management by jointly examining those who ask for the service and those who provide it.³⁹

To this, ‘complex’ monitoring missions add the chance that a contradicting view will inevitably face accusations from either the government or the opposition. While an assignment improves an organization’s reputation, delivering a dissenting vote carries the risk of destroying – or at least undermining – it. This danger is especially large for Type II errors, so non-reputational benefits must exist for producing an assessment that contradicts a negative report. Unsurprisingly, such positive assessment often come from organizations in which the incumbent is itself a major player and which are therefore most likely to provide biased assessments. We believe that incumbent governments will rely most often on observers who have not build up a reputation of delivering high quality reports in the past and who are politically close to team up with another monitor. Specifically, we maintain that observer missions that depend on the incumbent

government are much more likely to provide a positive assessment of an election. Hence, interest collusion between the observer and the observed might hinder the effective monitoring of suspicious elections. Liisa Laakso provides evidence for these shady tendencies for the 2000 election in Zimbabwe, noting that 'some observers from the neighboring countries were astonishingly pro-government in their statements.'⁴⁰

In sum, there are strategic reasons for why we observe multiple observers in some elections, but not in others. This means, by extension, that multilateral monitoring efforts are not a random sample of all monitoring cases and that the average quality of unilateral and multilateral election observer mission reports should differ. The theoretical argument that we will develop below takes this stepwise decision-making process into account.

Selection of multilateral monitoring: Like other conflict resolution mechanisms, election monitoring is, by and large, a voluntary process on which governments and one or several outside parties have to agree. This implies that incumbent leaders have an almost free hand in asking additional observers to provide their monitoring services. The formal models summarized above suggest that asking for a second opinion might decrease the benefits that an incumbent government receives from its willingness to expose its election organization to independent scrutiny, especially if the additional observer is not highly reputable and likely to produce a rosy election assessment. Hence, asking for a second opinion reduces the likelihood that an independent monitor's report will be taken seriously, whatever the evaluation's outcome will be. Some governments, however, might accept the risk of losing face if they believe that the danger of being stigmatized as an election fixer is too large. As Bruce Bueno de Mesquita Alastair Smith, Randolph M. Siverson and James D. Morrow and others have shown, the risk of losing office affects the present-day behavior of political leaders and forces them to use carrot and stick policies towards their 'selectorates' to survive an imminent election or revolutionary challenge.⁴¹ Asking for a second report therefore comes close to what George Downs and David Rocke have called 'gambling for resurrection.'⁴²

The danger of a negative report looms particularly large in two contexts. First, the incumbent government might fear the shadow of past election fraud. If it or a predecessor had given the orders in the last election to manipulate the results, it might therefore believe that election observers will be especially careful in their evaluations. Second, incumbent governments in anocracies might fear that the international community does not trust that the elections for which they are responsible will be sufficiently free and fair compared to the leaders of established democracies. Anocracies might also bear a disadvantage compared to autocracies because they do

not necessarily have sufficient means to quell public protests after the publication of a negative assessment of the elections.

H1: Incumbent governments that face a higher chance of negative reports because their countries lack a tradition of free and fair elections will have a more pronounced tendency to invite a mix of observers. This tendency manifests itself either a) through a recent history of rigged elections or b) through the anocratic state of the political institutions.

A second reason for inviting an additional election observer might be international pressure. This is particularly the case for smaller developing countries in whose economies foreign aid plays an important role. Leaders in such countries might try to counteract a possibly negative report by inviting or accepting a second election observer team. Note, however, that foreign leader whose countries depend on foreign aid do not necessarily have to fear single observer missions. Judith Kelley reports for instance that observers are more likely to endorse elections in such countries.⁴³ Case studies furthermore suggest that the strategic interest of donors has occasionally prolonged the tenure of leaders such as Kenyan President Daniel arap Moi.⁴⁴ However, aid donors frequently consider election observation as a pre-condition of further assistance. The call for monitors will, according to this reasoning, be most pronounced in countries in which the international community can make such demands quite credibly.

H2: The more that an incumbent government profits from the international community's support the higher the chance that a government invites a mix of observers.

How multilateral monitoring affects post-election unrest: Incumbent governments invite a mix of critical and friendly observers to avoid outcomes undesirable to them, such as being faced with post-election protests and rioting or having to use repression to contain these anti-regime protests. This gamble, however, only pays off if observers disagree or if, against all odds, all reports agree that the elections were free and fair. The principal-agent literature strongly suggests that multiple principals (and the disagreements associated with them) increase the power of the agent.⁴⁵ Observer disagreement mainly benefits those leaders who believe that fixing election results is the only thing that might help them stay in power. Conversely, more settled or self-confident leaders who do not see a need to cheat might prefer to stick to a unilateral mission because they do not want to signal

their pre-election desperation to the voters and the opposition through the invitation of a team of monitors.

Technically, increasing potential conflict among observers increases the size of the core and thus the stability of the decision making process. A growing literature in international political economy supports the conjecture that growing disagreement among principals increases the slack of the agent.⁴⁶ Although in their relationship with the outside world incumbent governments are not agents in the strict sense, they are thus nevertheless considerably weakened through negative reports. Hence, their ‘gamble for resurrection’ will only be successful if they can count on at least some disagreement between the international monitors or an overall positive assessment by this team.

H3: Inviting a mix of low- and high-quality observers reduces the risk of a) post-election unrest and b) post-election repression.

Research Design

Data and Methodology

We start by creating a dataset of all election rounds held from 1980 to 2004.⁴⁷ Elections data come from the National Elections Across Democracy and Autocracy (NELDA) dataset.⁴⁸ NELDA includes all countries that held competitive elections since 1945. The unit of analysis in our data is the election-round, meaning that runoff elections and multiple-round legislative elections are coded as separate cases.⁴⁹ The dataset includes 1,638 elections rounds for the time period under analysis.

Our empirical analysis consists of two parts. First, we model incumbents’ decision to invite international election observer organizations to examine whether our hypotheses regarding the invitation of a mix of observers is supported. Since incumbents can decide whether to invite one or more organizations or low or high-quality organizations, we use a multinomial logit model that can distinguish among different types of invitations.⁵⁰ Second, we examine the consequences of these different types of observer invitations for a variety of post-election outcomes. We use logit models to assess the effect of different types of observer invitations on post-election leadership protests and repression. Since our models of observer invitations show that certain elections are more likely to be observed, we use matching to preprocess the data and reduce concerns on endogeneity bias.

Dependent Variables

Observer Invitation: For the first two hypotheses, our dependent variable is a categorical measure of the number and quality of international election observer organizations present at each election. Recall that Hypotheses 1 and 2 focus on observer invitations and expect that incumbent governments with a history of cheating, those with questionable commitments to democracy, and those facing international pressure are more likely to invite a mix of low and high-quality observer organizations. We use the Data on International Election Monitoring (DIEM) to indicate the number and quality of observer organizations.⁵¹ The DIEM data are based on information from 592 reports by 19 organizations that monitored elections from 1980 to 2004 and are the most detailed available account of the activities of international monitors. The data include information on the number, names, and assessments of monitoring organizations present for each election, which we use to create a categorical variable indicating whether an election was observed by one or more organizations and the quality of the organization involved.

We create five categories for observed elections that seem most theoretically relevant: elections observed by one low-quality organization, elections observed by two or more low-quality organizations, elections observed by a combination of low and high-quality organizations, elections observed by one high-quality organization, and elections observed by two or more high-quality observers. Establishing the number of organizations is uncontroversial, but the categorization of organizations as low or high-quality merits explanation. Following Alberto Simpser and Daniela Donno⁵² and Susan Hyde and Nikolay Marinov,⁵³ we define high-quality organizations as those that have been willing to criticize fraudulent elections in the past. The following organizations were coded as high-quality: Asian Network for Free Elections (ANFREL), the Carter Center, Commonwealth, Council of Europe, Electoral Institute of Southern Africa (EISA), European Parliament (EP), European Union (EU), International Republican Institute (IRI), National Democratic Institute (NDI), Organization of American States (OAS), Organization for Security and Cooperation in Europe, and the United Nations (UN). Four organizations in DIEM remain and were then defined as low-quality: The African Union (AU), the Commonwealth of Independent States (CIS), the International Human Rights Law Group (IHRIG), and the South African Development Community (SADC).⁵⁴ Of the four low-quality organizations, the CIS, SADC, and IHRIG have never criticized a problematic election, and the AU has criticized highly fraudulent elections less than 15% of the time.⁵⁵ One might object that our definition of high-quality organizations is rather lenient, but we are most interested in cases where incumbents try to counteract the risk of a negative report by inviting organizations that are almost certain to provide an endorsement.⁵⁶ As we argued in our theoretical section, combining an invitation to friendly observers with those the international

community deems reputable maximizes incumbents' chance to avoid the costly consequences of inviting observers. Table 5.1 presents the regional distribution of observer invitations with respect to the number and quality of organizations invited.

TABLE 5.1 ABOUT HERE

Table 5.1 shows that of 388 of 1,638 elections in the data were monitored by international observers, leaving 1,250 election rounds without monitors. Since no election-rounds were monitored by two low-quality organizations, we do not show them in the table and exclude this category from our analyses.⁵⁷ Globally, we observe 36 elections that were monitored by a single low-quality organization. Elections observed by a mix are more common (16%), although the most frequent missions consist of one single high-quality organization (36%) and two or more high-quality organizations (39%). Table 5.1 also presents the regional distribution of observer organizations. We observe the highest percentage of mixed monitoring missions in Africa (30%), followed by the Middle East and North Africa (25%), and Asia (18%).

The dependent variable in the statistical analysis contains five categories and is coded zero for elections without observers, one for those with one low-quality organization, two for mixed observer groups, three for one high-quality observer, and four for two or more high-quality organizations.

Post-Election Unrest: We examine two dependent variables for our analysis of post-election unrest. The third hypothesis anticipates that inviting a mix of observers, particularly in elections where incumbents cheated, will lower the risk of post-election protests and post-election repression, outcomes incumbents would prefer to avoid. Data for post-election protests and riots, and post-election repression by the government, respectively, come from NELDA. The first measure is a dichotomous variable indicating whether riots and protests related to the handling or outcome of elections occurred in elections' aftermath. 150 election rounds (14%) in the data experienced post-election protests and riots. The second variable measures whether the government responded to such riots and protests with the use of violence. The repression variable is dichotomous and governmental violence occurred in 70 election rounds (or 6%).

Independent Variables

Observer Invitation: In hypotheses 1 and 2, we argued that inviting a mix of friendly and critical observers would be particularly attractive to incumbents who worry about a negative report but are under international pressure to invite monitors. Incumbents fearing a critical report should be those

who are inclined to manipulating elections. While we cannot empirically measure incumbents' intent to cheat, past fraudulent behavior should be a reasonably proxy for capturing this concept. We use the Quality of Elections Data (QED) to create this variable. QED uses information from U.S. State Department Human Rights Reports and codes whether an election was acceptable, ambiguous, or unacceptable. We transform this measure into a dichotomous variable coded one for cases in which previous elections were considered unacceptable, zero otherwise.⁵⁸ We also argued that uncertainty over whether a government is institutionally able to commit itself to democracy should lead to more observer invitations.⁵⁹ This risk should be particularly high in semi-democracies. To test for the possible curvilinear impact of democracy, we include the Polity IV democracy measure in its simple and its squared version.⁶⁰ The coefficient for the squared polity measure is expected to be significant and negative, indicating a curvilinear and n-shaped relationship between democracy and observer invitations.

Our second hypothesis regarding the invitation of multiple monitors focused on international pressure, and we expected that incumbents receiving large amounts of official development assistance should be particularly inclined to invite a mix of lenient and reputable observer organizations. We use data on official development assistance (ODA) from the AidData project to create this variable, which measures each country's net ODA flows in constant U.S. dollars.⁶¹ We take the natural log of ODA flows because the values are right-skewed, and we also lag values by one year to account for possible endogeneity.

We also include several control variables. First, Judith Kelley and Susan Hyde⁶² expect that first multiparty elections and first elections after suspension are more likely to attract monitors. NELDA includes a variable that indicates whether elections were the first with multiple parties or after suspension and we create a dummy variable coded one for such elections, zero otherwise. Second, we control for the effect of economic development by including logged GDP per capita values and also lag values by one year. Third, we include dummy variables for the Americas, Europe, Africa, the Middle East and North Africa, and Asia to account for regional differences. Coefficients for dummy variables (with the Middle East and North Africa as the excluded category) are not reported to preserve space. Finally, following recommendations by Nathaniel Beck, Jonathan Katz, and Richard Tucker⁶³ we create a measure that counts the number of years since a government last invited multiple election observer organizations and three cubic splines to account for temporal dependence. We present results for the observer year variable but exclude splines to preserve space.

Post-Election Unrest: The third hypothesis examines the effect of different types of observer invitations on the probability of post-election protests and governmental violence against

protesters. Our main expectation contends that a mix of friendly and critical monitoring organizations makes it easier for incumbents –particularly those who cheated - to avoid anti-regime rioting or the use of repression. As highlighted in the example of the 2004 Kazakh elections, incumbents can exploit contradictions between organizations of differing quality by highlighting the endorsements of friendly monitors and omitting more critical ones. We use the dependent variable in the invitations model to operationalize different types of observer organizations and disaggregate it into a four dummies: a dichotomous variable for elections monitored by a single low-quality organization, a second variable for elections monitored by one low-quality and one high-quality organization, a third variable for elections observed by one high-quality organization, and a fourth variable for elections monitored by two or more high-quality organizations. Each of these variables is coded one if the respective configuration was present, zero otherwise.

We argue that the benefits of inviting a mix of observers materialize especially when incumbents resort to cheating. While an invitation to a friendly monitoring organization likely undermines the credibility of an incumbent, he or she is unlikely to be punished for inviting such organizations if the elections are assessed as acceptable by all monitoring organizations present. For that reason, we expect that the advantages of strategic monitoring invitations are especially pronounced in fraudulent elections. To examine the combined effect of mixed monitoring missions and fraud in elections, we create an interaction between the dummy variable for elections observed by low and high-quality monitors and a variable indicating the presence of serious fraud in elections. We use the QED data to measure fraud. The variable is coded one if elections were considered unacceptable, zero otherwise. We also create interactions between fraud and the remaining types of monitoring missions (one low-quality observer, one high-quality observer, and two or more high-quality observers).

We include several control variables in the post-election models. First, we create a dummy variable that indicates whether opposition parties experienced gains in their vote shares. The NELDA data include this variable and we expect that an increase in the number of votes for opposition parties would reduce the probability anti-regime unrest. Second, we retain the development assistance variable from the selection equation to examine whether incumbents receiving aid are more likely to be removed or experience protests as a result of greater international pressure. Third, we again include a regional dummy variables. Finally, we control for temporal dependence by creating variables that indicate the number of years since the protest and repression, together with three cubic splines. We use standard errors clustered by country in all models.

In the post-election models, we use matching methods to address threats to causal inference. In particular, international election monitoring could be the result of anticipating protests

or violence after elections rather than having an effect on them. Observers may be more likely to monitor elections that have a high risk of leading to unrest and an empirical relationship could thus be driven by observers' anticipation of these outcomes rather than the presence of observers. We use coarsened exact matching to reduce the risk of such bias and match on variables that are likely to influence observer anticipation of election fraud, expecting that fraud in the preceding election-round, level of democracy, GDP per capita, first elections, and development assistance affect the likelihood of observation.⁶⁴ Matching preprocesses the data on these variables and excludes observations from the treatment group (i.e. observed elections) and control group (i.e. unobserved elections) that differ fundamentally on covariate values and thus might be driven by the absence or presence of observers.⁶⁵

Results

Observer Invitations

The models in Table 5.2 present findings for our hypotheses on the likelihood of inviting observer organizations.

TABLE 5.2 ABOUT HERE

Our main theoretical focus is on incumbents who invite a combination of low and high-quality observers (model 2). We find that incumbents who cheated in the past or whose commitment to democracy is uncertain (Hypothesis 1) and those who are dependent on international development assistance (Hypothesis 2) are more likely to invite a mix of observer organizations. The coefficients for all three variables are in the expected direction and significant at the 95% confidence level and therefore support the first two hypotheses. While uncertain commitment to democracy and international assistance are significant for all observer invitations (model 1) and remain significant for high-quality observer invitations (models 3 and 4), the coefficient for fraud in previous elections is significant only in model 3. The findings therefore confirm that incumbents considering manipulation but simultaneously facing international pressure to invite observers are most inclined to ensure at least some positive reports by inviting a combination of lenient and reputable observers.

While less important for our argument, we also observe that the determinants of observer invitations are quite different for different types and numbers of observers. The coefficient for first elections is significant only for elections monitored by two or more high-quality organizations, suggesting that they are particularly likely to receive multiple high-quality observers, which could be

due to greater international pressure or the need of incumbents to strongly signal their commitment to democracy. Table 5.3 presents the marginal effects for significant variables in Table 5.2.

TABLE 5.3 ABOUT HERE

Post-Election Unrest

The third hypothesis argues that inviting a mix of monitoring organizations – particularly if incumbents manipulated election outcomes – reduces the likelihood of post-election unrest and repression after elections.

TABLE 5.4 ABOUT HERE

Findings for post-election protests and repression are presented in Table 5.4. The first two models show coefficients without interactions between fraud and observer types, but since we would anticipate that only cheating incumbents have to worry about punishment, we are most interested in models including interactions (models 2 and 4). Findings confirm that the combination of inviting a mix of observers together with fraud in elections somewhat reduces the risk of protests and the need to use repression. The coefficient for the interaction between mixed observation missions and fraud is insignificant in the both the protest and the repression model, but coefficients for interactions between fraud and one high-quality observer group (model 4), and two or more high-quality observer groups, respectively (models 2 and 4) are positive and significant. Hence, the punishment for cheating resulting from the judgments of high-quality observers does not occur in mixed observation missions. While our findings for the punishment imposed for cheating documented by high-quality monitors are thus somewhat inconsistent, we should consider collapsing these two categories since the determinants of high-quality observer invitations were similar regardless of their number.

Taken together, while findings for post-election unrest do not support our expectations as strongly as expected, incumbents extending these invitations do not seem to incur punishments for manipulating elections despite the presence of some high-quality observers. Our findings thus imply that invitations of a combination of friendly and critical observers can help counteract the costly consequences of negative observer assessments for incumbents.⁶⁶

Conclusion

Election monitoring has turned into a global norm almost no government can escape. However, incumbent governments have considerable possibilities to undermine the credibility of the international observers. One particularly frequent means is the appointment of a mix of low and

high-quality observers. As governments and the international community agree on 'complex' monitoring missions, they are not representative of all elections to which monitors are assigned. In this chapter we contend that we can only understand the effect of these election monitors on political outcomes if we jointly examine these results with the origins of mixed monitoring missions. In our view, governments will invite a mix of friendly and critical observers if they face a relatively high risk of being confronted with a negative observer report but simultaneously face international pressure to invite reputable observers. This argument suggests that incumbents with a history of cheating as well as those under the close scrutiny of the international community will call for a second report by friendly monitors in the hope that this observer will offset a possibly negative report through a more positive assessment.

The empirical analysis lends qualified support to this double conjecture. We show that incumbents who cheated in the past and face international pressure to extend invitations are more likely to invite a combination of low and high-quality observer organizations. As we have argued, these governments are more likely to face a critical report and therefore try to counteract this risk by inviting at least one friendly organization. While we do not find that a mix of observers in fraudulent elections significantly lowers the risk of post-election unrest, we argue that the absence of a positive relationship still marks an improvement from the perspective of incumbents. Since our findings show that elections certified as fraudulent by high-quality international monitors face a greater risk of anti-regime protests and repression, a mix of observers successfully undermines the costly consequences of electoral manipulation.

We should note that the supply of low-quality monitors is concentrated in Africa and the post-Soviet region and thus not equally available to all incumbents, implying that our findings are particularly important for reputable organizations active in those regions. Yet a potentially concerning development is the recent involvement in election monitoring by the Union of South American States (UNASUR), an organization that has not endorsed the Declaration of Principles for International Election Monitoring and sent its first mission to 'accompany' the 2012 presidential elections in Venezuela.

More generally, our argument demonstrates how the growing norm of international election monitoring has led to strategic adaptation by both governments and international organizations. Absent a central organization in charge of coordinating election monitoring, a 'shadow market' of lenient monitoring organizations has emerged, and this chapter helps clarify the conditions under which governments are particularly inclined to use this increasing supply of monitoring organizations to their advantage.

TABLES

Table 5.1: Regional Distribution of Number and Quality of Observer Organizations across Elections, 1980-2004

Variable	Americas	Europe	Africa	Middle East & North Africa	Asia	Total (%)
One Low Quality	5 (6.4)	0 (0.0)	27 (29.7)	3 (37.5)	1 (1.35)	36 (9.3)
Mix of Low and High Quality	2 (2.6)	18 (13.1)	27 (29.7)	2 (25.0)	13 (17.6)	62 (16.0)
One High Quality	33 (42.3)	43 (31.4)	26 (28.6)	3 (37.5)	34 (46.0)	139 (35.8)
Two or More High Quality	38 (48.7)	76 (55.5)	11 (12.1)	0 (0.0)	26 (35.1)	151 (38.9)
Total	78	137	91	8	74	386

Note: No elections were monitored by two or more low-quality organizations.

Table 5.2: Multinomial Logit Models of Observer Invitations, 1980-2004

Variables	One Low Quality	Mixed Quality	One High Quality	Two High Quality
Fraud _{t-1}	-0.459 (0.499)	1.306** (0.433)	0.244 (0.346)	0.063 (0.411)
First Election	-0.100 (0.415)	0.013 (0.428)	0.249 (0.353)	0.715* (0.348)
ODA logged _{t-1}	-0.211 (0.151)	0.406** (0.098)	0.302** (0.077)	0.310** (0.075)
Polity _{t-1}	-0.069 (0.058)	0.092* (0.044)	0.023* (0.030)	0.038** (0.034)
Polity Squared _{t-1}	-0.031** (0.008)	-0.027** (0.007)	-0.012* (0.005)	-0.036** (0.005)
GDP logged _{t-1}	-0.173 (0.291)	0.046 (0.302)	0.132 (0.214)	0.295 (0.209)
Observer Years	0.014 (0.030)	-0.115** (0.044)	-0.057** (0.022)	-0.070* (0.030)
Constant	0.095 (3.051)	-5.525+ (2.840)	-5.605** (2.556)	-15.909** (2.039)
Observations	1,332	1,332	1,332	1,332
Number of Countries	145	145	145	145

Standard errors in parentheses and clustered by country. Models present coefficients for different numbers and quality of observer organizations using no observers as the baseline category.

** p<0.01, * p<0.05, + p<0.1 (two-tailed tests)

Table 5.3: Marginal Effects for Significant Variables in Table 13.2

Variables	One Low Quality	Mixed Quality	One High Quality	Two High Quality
Fraud _{t-1}	-	+0.063	-	-
First Election	-	-	-	+0.051
ODA logged _{t-1}	-	+0.012	+0.019	+0.014
Polity ^a	-0.039	-0.054	-0.018	-0.190
Observer Years	-	-0.004	-0.003	-0.003

Note: Marginal effects calculated by varying continuous variables $\pm 1SD$ from the mean and 0 to 1 for dichotomous variables.

^a Polity and Polity squared are varied jointly from the mean to +1SD above the mean.

Table 5.4: Logit Models of Protests and Repression After Elections, 1980-2004

Variables	Protest 1	Protest 2	Repression 1	Repression 2
One Low Quality	0.234 (0.430)	-0.240 (0.563)	0.659 (0.655)	0.370 (0.829)
Mixed Quality	0.906* (0.454)	0.429 (0.528)	0.884+ (0.481)	0.324 (0.668)
Two High Quality	0.951** (0.361)	0.474 (0.439)	0.785 (0.512)	0.258 (0.563)
One High Quality	0.101 (0.392)	-0.331 (0.502)	-0.042 (0.535)	-1.360 (1.109)
One Low*Fraud	-	1.240 (0.887)	-	0.702 (1.248)
Mixed*Fraud	-	1.281 (1.000)	-	1.444 (1.130)
One High*Fraud	-	1.298 (0.791)	-	2.568+ (1.314)
Two High*Fraud	-	2.171* (0.866)	-	1.746* (0.872)
Fraud	-0.183 (0.300)	-0.729 (0.380)	-0.172 (0.403)	-0.707 (0.486)
Opposition Gain	-0.689* (0.320)	-0.686* (0.320)	-1.013** (0.374)	-1.029** (0.372)
ODA logged _{t-1}	-0.074 (0.087)	-0.071 (0.088)	-0.142 (0.139)	-0.156 (0.146)
Protest Years	-0.267** (0.085)	-0.275** (0.085)	-	-
Repression Years	-	-	-0.130 (0.184)	-0.124 (0.190)
Constant	-0.031 (0.811)	-0.393 (0.826)	-0.228 (1.139)	-0.269 (1.199)
Observations	1,083	1,083	1,087	1,087
Number of Countries	145	145	145	145

Standard errors in parentheses and clustered by country.

** p<0.01, * p<0.05, + p<0.1 (two-tailed tests)

Acknowledgments: Previous versions of this chapter have been presented at the annual convention of the International Studies Association, San Francisco, April 2-5, 2013, the ENCoRe meeting in Amsterdam, April 24-26, 2013, and the Harvard Workshop of the Electoral Integrity Project in Cambridge, MA, June 3-4, 2013. The authors would like to thank the Leibniz Foundation and European Commission for travel support and workshop participants as well as Nikolay Marinov, Andrea Ruggeri, and Brian Burgoon for helpful comments.

² Information on observer assessments comes from the Data on International Election Monitoring (DIEM) collected by Judith Kelley, available at <http://sites.duke.edu/kelley/>.

³ Kelley, Judith. 2009a. 'The More the Merrier? The Effects of Having Multiple International Election Organizations.' *Perspectives on Politics* 7(1): 59-64.

⁴ <http://www.kazelection2004.org/observers.htm>.

⁵ The organizations present were the Organization of American States, the OSCE/Helsinki Center, and a joint mission by the Carter Center and the National Democratic Institute. For an example of their assessments, see the Carter Center/NDI report at <http://www.cartercenter.org/documents/292.pdf>.

⁶ The negative repercussions for Fujimori raise the question of why he invited high-quality monitors and then cheated in front of them. However, the Fujimori regime had strong control over the media and he may have hoped to suppress negative reports (Conaghan 2002). In addition, Cooper and Legler (2005) provide evidence of the strong control of the Fujimori regime over the OAS monitoring mission, while the opposition hoped to profit from the international media campaign, mistrusting the OAS because of its pro-authoritarian bias of the organization during the Cold War.

⁷ Hyde, Susan D. 2011. 'Catch Us if You Can: Election Monitoring and International Norm Diffusion.' *American Journal of Political Science* 55(2): 356-369; Kelley, Judith. 2012. *Monitoring Democracy: When International Election Observation Works, and Why It Often Fails*. Princeton, NJ: Princeton University Press.

⁸ Magaloni, Beatriz. 2010. 'The Game of Electoral Fraud and the Ousting of Authoritarian Rule.' *American Journal of Political Science* 54(3): 751-65. Little, Andrew T. 2012. 'Elections, Fraud, And Election Monitoring in the Shadow of Revolution.' *Quarterly Journal of Political Science* 7(3): 249-283. Svobik, Milan W. and Svitlana Chernykh. 2012. 'Third-Party Actors and the Success of Democracy: How Electoral Commissions, Courts and Observers Shape Incentives for Election Manipulation and Post-Election Protest.' *Unpublished Working Paper*, University of Illinois/Oxford University.

⁹ Magaloni, 2010.

¹⁰ Kelley, Judith and Kiril Kolev. 2010. 'Election Quality and International Observation: Two New Datasets.' Working paper, *Duke University*.

¹¹ Hyde 2011: 356.

¹² Kelley 2012: 16-17.

¹³ Kelley 2009a.

¹⁴ Kelley and Kolev, 2010: 6.

¹⁵ Kelley 2009a: 59.

¹⁶ Kelley 2009a.

¹⁷ Kelley 2012: 53.

¹⁸ Tolstrup, Jakob. 2009. 'Studying a Negative External Actor: Russia's Management of Stability and Instability in the 'Near Abroad.' *Democratization* 16(5): 922-944.

¹⁹ Kelley 2012.

²⁰ e.g. Knack, Stephen and Aminur Rahman. 2007. 'Donor Fragmentation and Bureaucratic Quality in Aid Recipients.' *Journal of Development Economics* 83(1): 176-197.

²¹ Hyde, Susan D. 2007. 'The Observer Effect in International Politics: Evidence from a Natural Experiment.' *World Politics* 60(1): 37-63; Donno, Daniela. 2010. 'Who is Punished? Regional Intergovernmental Organizations and the Enforcement of Democratic Norms.' *International Organization* 64(4): 593-625; Daxecker, Ursula E. 2012. 'The Cost of Exposing Cheating: International Election Monitoring, Fraud, and Post-Election Violence in Africa.' *Journal of Peace Research* 49(4), 503-516; Hyde, Susan D. and Nikolay Marinov. 2013. 'Information and Self-Enforcing Democracy: The Role of International Election Observation.' *International Organization* (in press); Kelley 2012.

²² Hyde 2007.

²³ Kelley 2012.

²⁴ Donno 2010.

²⁵ Hyde and Marinov 2013.

²⁶ Whether protests or violence actually increase the probability of democratic elections or transitions to democracy in the long run should ultimately be an empirical question, and neither study examines these long-term consequences.

²⁷ Daxecker 2012.

²⁸ Kelley, Judith. 2009b. 'D-Minus Elections: The Politics and Norms of International Election Observation.' *International Organization* 63(4): 765-787.

²⁹ Hyde, Susan D. and Angela O'Mahony. 2010. 'International Scrutiny and Pre-Electoral Fiscal Manipulation in Developing Countries.' *Journal of Politics* 72(2):690-704; Simpser, Alberto and Daniela Donno. 2012. 'Can International Election Monitoring Harm Governance?' *Journal of Politics* 74(2): 501-513; Beaulieu, Emily and Susan D. Hyde. 2009. 'In the Shadow of Democracy Promotion: Strategic Manipulation, International Observers, and Election Boycotts.' *Comparative Political Studies* 42 (3): 392-415.

³⁰ Kelley 2009a.

³¹ Kelley 2009a: 61.

³² Kelley 2009a.

³³ Magaloni 2010.

³⁴ Little, Andrew T. 2012. 'Elections, Fraud, And Election Monitoring in the Shadow of Revolution.' *Quarterly Journal of Political Science* 7(3): 249-283.

³⁵ Svulik, Milan W. and Svitlana Chernykh. 2012. 'Third-Party Actors and the Success of Democracy: How Electoral Commissions, Courts and Observers Shape Incentives for Election Manipulation and Post-Election Protest.' *Unpublished Working Paper*, University of Illinois/Oxford University.

³⁶ The analogous relationship holds for a highly informative endorsement of the incumbent by unrestrained observers and for support with little informative value by highly restrained monitors.

³⁷ Hyde and Marinov 2013.

³⁸ Kelley 2009a: 63.

³⁹ Bercovitch, Jacob and Gerald Schneider. 2000. 'Who Mediates? The Political Economy of International Conflict Management.' *Journal of Peace Research* 37(2): 145-165.

⁴⁰ Laakso, Liisa. 2002. 'The Politics of International Election Observation: The Case of Zimbabwe in 2000.' *Journal of Modern African Studies* 40(3): 437-464.

⁴¹ Bueno De Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA : MIT Press. Conaghan, Catherine M. 2002. "Cashing in on Authoritarianism: Media Collusion in Fujimori's Peru." *The Harvard International Journal of Press/Politics* 7(1): 115-125.

⁴² Downs, George W. and David M. Rocke. 1994. 'Conflict, Agency, and Gambling for Resurrection: The Principal-Agent Problem Goes to War.' *American Journal of Political Science* 38 (2): 362–380.

⁴³ Kelley 2009b.

⁴⁴ Brown, Stephen. 2001. 'Authoritarian Leaders and Multiparty Elections in Africa: How Foreign Donors Help to Keep Kenya's Daniel Arap Moi in Power.' *Third World Quarterly* 22(5): 725-739.

⁴⁵ Mastenbroek, Ellen, Heike Klüver, Gerald Schneider and Dietrich Drüner. 2013. 'The Core or the Winset? Predicting Policy Change and Decision-Making Efficiency in the European Union.' *Unpublished Manuscript*, Radboud Universiteit/Universität Konstanz.

⁴⁶ e.g. Copelovitch, Mark. 2010. 'Master or Servant? Common Agency and the Political Economy of IMF Lending.' *International Studies Quarterly* 54(1): 49-77; Nielson, Daniel L. and Michael J. Tierney. 2003. 'Delegation to International Organizations: Agency Theory and World Bank Environmental Reform.' *International Organization* 57(2): 241-276.

⁴⁷ Election monitoring data described in Kelley and Kolev (2010) are available for this time frame.

⁴⁸ Hyde, Susan D. and Nikolay Marinov. 2012. 'Which Elections Can Be Lost?' *Political Analysis* 20(2): 191-210.

⁴⁹ Concurrent elections (legislative and executive elections held on the same day) are coded as a single event.

⁵⁰ Hausman and Small-Hsiao tests conducted for multinomial logit models showed that the assumption of IIA is not violated.

⁵¹ (Kelley and Kolev 2010). For most multiple-round elections, the organizations present in one round also observe subsequent rounds. Data are available at:

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/31461/version/1>.

⁵² Simpser and Donno 2012.

⁵³ Hyde and Marinov 2013.

⁵⁴ While DIEM includes data on 19 organizations, reports were not available for three organizations and we can therefore not assess whether they criticized elections.

⁵⁵ Kelley 2012: 53. A concern with this definition of low quality is that organizations becoming engaged in election monitoring only recently have fewer opportunities to criticize problematic elections. Yet several high-quality organizations have been active for similar or shorter time frames

than low-quality ones. For example, ANFREL and EISA have been active since 1998, whereas the AU has started to monitor elections in 1992 and the SADC in 1999.

⁵⁶ Kelley (2012) defines low-quality organizations as those criticizing problematic elections less than 50 percent of the time, yet we believe that a more restrictive definition of low-quality is appropriate for our argument since we focus on incumbents who attempt to ensure at least one supportive report. In addition, a lenient definition avoids accidentally labeling an organization as low quality for sometimes failing to criticize fraudulent elections. As discussed earlier, even credible organizations are sometimes limited in their ability to provide systematic assessments, or may refrain from providing them because of political concerns (see Davis-Roberts and Carroll chapter in this volume and Kelley 2012: 57).

⁵⁷ This absence is unsurprising given the small number of low-quality organizations and the lack of legitimacy to be gained from the invitation of two organizations of questionable quality as opposed to just one.

⁵⁸ While we would like to examine whether alternative measures of fraud would produce similar results, few other systematic sources are available. NELDA includes a variable that indicates the presence of pre-election concerns over fraud, but the variable does not measure the actual incidence of fraud, nor does it specify how serious these concerns were.

⁵⁹ Kelley 2012; Hyde 2011.

⁶⁰ Marshall and Jagers 2011. We lag the Polity variables by one year to reduce concerns regarding endogeneity bias.

⁶¹ Available at <http://www.aiddata.org/content/index>.

⁶² Kelley 2012; Hyde 2011.

⁶³ Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. 'Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable.' *American Journal of Political Science* 42(4): 1260-1288.

⁶⁴ Iacus, Stefano M., Gary King and Giuseppe Porro. 2012. 'Causal Inference Without Balance Checking: Coarsened Exact Matching.' Forthcoming in *Political Analysis*.

⁶⁵ Pre-processing of the data removes 335 observations from the data.

⁶⁶ Hyde and Marinov 2013; Daxecker 2012.