DETECTING AND CORRECTING ELECTION FRAUD

Daniel Sutter
University of Oklahoma

INTRODUCTION

Free and fair elections are the cornerstone of democracy. Government reflects the will of the people only by accident unless citizens can express their preferences through free elections. Assuming that elections are fair, however, is a dubious proposition, even in the United States. Fraud helped sustain political machines in Chicago, New York, and Philadelphia, rendered elections irrelevant in the post-Reconstruction South, and may have swung the 1960 presidential election to Kennedy. The Watergate break-in during the 1972 campaign led to Nixon's resignation, under threat of imminent impeachment. Charges of vote buying in New Orleans surfaced in the 1996 Louisiana Senate race.[1]

How can voters ensure the integrity of elections? A whole range of potential problems exist. Election officials must detect fraud, which can be problematic with thousands of polling places. The election result must be overturned when fraud is detected. Despite monitors' best efforts, elections will be less than perfectly reliable. Will the resulting election contests serve voters' interests?

Political scientists commonly cite attitudes toward corruption, party competition, introduction of the secret ballot, and the rise of a mass electorate as determinants of the incidence of electoral fraud [Key, 1949; O'Leary, 1962; Benson, 1978; Goldberg, 1987]. But the question is more complicated than these explanations suggest. Fair elections are a public good, so attitudes toward corruption consider only the potential demand for, and not the cost of, fairness. Key expects party competition to check fraud because, "if two would-be thieves watch each other, an honest count will result" [1949, 201]. But this ignores possible collusion between parties to share power and save the costs of competition.[2]

The numerous public choice models of elections typically ignore the possibility of fraud. Yet relaxation of electoral constraints could raise politicians' utility. The effect of the introduction of the secret ballot on the market for votes, a type of election fraud, has received some attention [Anderson and Tollison, 1990; Heckelman, 1995], but these studies do not consider efforts to ensure fair elections. Election chicanery can be considered to be a form of corruption, which has received continuing attention from political economists [Rose-Ackerman, 1978; Shleifer and Vishny, 1993; Mauro, 1995; Bliss and Di Tella, 1997; Triesman, 2000]. But the literature on corruption has not specifically examined vote fraud.

I employ a statistical decision-making model to examine the monitoring of elections for fraud. Costly monitoring provides imperfect information regarding fair-

Daniel Sutter: Department of Economics, University of Oklahoma, Norman, OK 73019. E-mail: dsutter@ou.edu.

ness. Election fraud presents a complication when compared to the economics of crime [Becker, 1968; Stigler, 1970]: the perpetrators can subvert the judicial process and avoid punishment.[3] To account for this possibility I incorporate a protest success function which determines whether a fraudulent election can be overturned.

I establish several results. (1) Monitoring an election is not always worth the cost since the value of the observers' signal may be zero. Democracy consequently may not be optimal in all nations or at all times. (2) Standards of evidence in considering charges of fraud change with the expected probability of fraud and level of income. Optimal monitoring allocates effort toward detecting fraud (eliminating false charges) when the prior expectation of fraud is high (low). (3) Optimal monitoring shifts toward avoiding false fraud reports when the current report influences expectations of fraud in future elections. Ignorance of fraud can be bliss.

I proceed as follows. I begin by describing a model of elections with fraud followed by a consideration of citizens' decision to monitor elections and use the resulting evidence to protest the official election result. I then establish basic results concerning the value of monitoring and the optimal allocation of monitoring effort and conclude with suggestions for international organizations interested in supporting new democracies.

A MODEL OF ELECTIONS WITH FRAUD

I present a model of elections with fraud. I consider a single contest for one office with a representative citizen. One interpretation of the model is a transition election held in a democratizing nation. The election itself produces a nominal result, which may be in error due to fraud. Observing the election provides information concerning fraud in the official results. Citizens decide in advance whether and how closely to monitor the election.[4] Monitoring produces a noisy information signal regarding the election result [Laffont, 1989, 55-69].

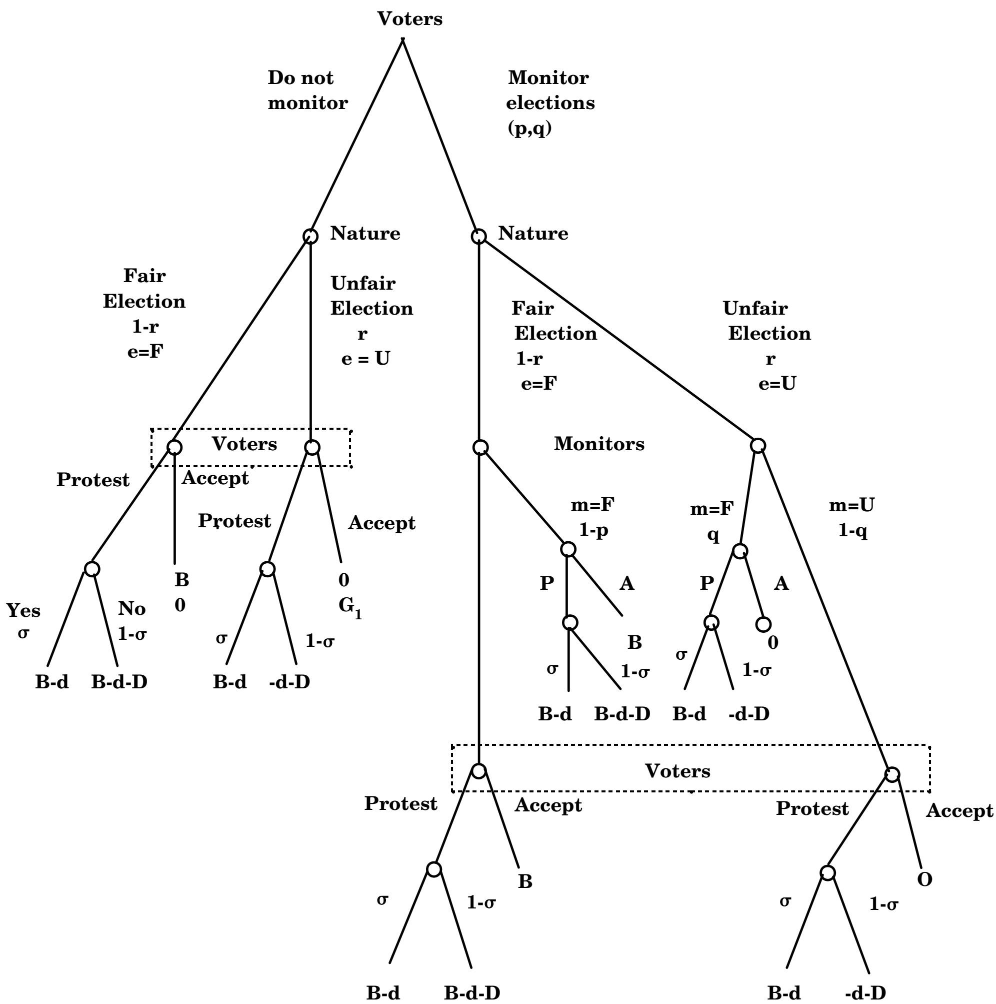
The representative citizen framework avoids differences in attitudes toward fairness among citizens. In particular, the representative citizen does not have preferences over the outcome of this election *per se*, but rather cares only about the fairness of the contest. Suppose we consider the constitutional level choice by citizens of the rules of the game. Citizens behind a veil of uncertainty abstract from whether they win or lose on individual plays of the game and consider the effect of rules on outcomes long term [Brennan and Buchanan, 1985]. From a constitutional perspective, the effects of institutionalized fair elections in providing an alternative to revolution in removing bad rulers and forcing politicians to consider the well-being of voters benefit all citizens. Citizens in a democratizing nation might truly consider establishing the precedent of fair elections to be more important to citizens than the winner of the first election. At the post-constitutional level, convergence to the median under a two party electoral system also leaves citizens concerned about only whether the election was fair, not the identity of the winner. Citizens benefit from institutionalized fair elections even if candidates' platforms end up being quite similar.

A first concern is defining fraud. Bryce provides a classic categorization of the varieties of electoral misconduct: “The rational will which the citizens are expected to possess and express by their votes may be perverted in three ways: by Fear, when the voter is intimidated; by Corrupt inducements, when he is bribed; by Fraud, when the votes are not honestly taken or honestly counted" [1929, 49]. I consequently distinguish misconduct at the level of casting ballots (which includes intimidation, bribery, treating or repeating) and misconduct at the level of counting ballots (deliberately spoiling ballots or improper counting). Standards of electoral conduct have varied over the years; I do not discuss which specific acts are classified as corrupt.[5] Two alternatives exist in defining fraud within the model. One approach would focus on the acts of misconduct—bribing voters, stuffing or destroying ballot boxes—which constitute fraud. The regime or competing parties might choose a level of misconduct, and fraud exists if either side chooses a positive level. A second approach would define fraud to exist when misconduct reverses the election result. Misconduct could occur in an election judged fair by the second definition. The appropriate definition of fraud would depend on whether we consider election outcomes to be indivisible or if the margin of victory is also important. I opt for the second definition and define an election to be fraudulent when the "official" winner is not the candidate who would have won if the level of misconduct in the campaign were zero. Installation of the true (zero misconduct) winner is the unambiguous remedy with this definition of fraud. This approach is also consistent with most corrupt practices legislation, which requires widespread improprieties to invalidate a result [Sikes, 1928]. Recent international observation efforts focus on irregularities which are "extensive, systematic, or decisive in a close race" [Elklit and Svensson, 1997, 38].

Events in the election occur as follows. Citizens decide whether to monitor an already scheduled election, and if so, how carefully to monitor the count. Monitoring resources translate into reliability of a report concerning fraud. The election occurs and if monitored, citizens receive a report concerning observed irregularities. Citizens then decide whether to protest or accept the official election result. A successful protest results in the installation of the true winners of the election. Accepting the returns and unsuccessful protest both leave the "official" election winner in place. Figure 1 provides a diagrammatic overview of events.

The fairness of the election is in question. The election has two outcomes, fair (F) or unfair (U), with $e \in \{F, U\}$ denoting the true election outcome. The election outcome depends on the actions of the participants in the contest yet I do not model the actions which possibly rig the election. Rather, the probability the election is rigged is exogenously set at r. I do this to focus on citizens' problem in trying to detect and control electoral misconduct. Clearly the probability of detecting and correcting fraud affects the behavior of election contestants. Equilibrium in a well-functioning democracy may involve minimal fraud, but the final outcome does not reveal the efforts required to eliminate fraud. A second justification for a probability of fraud as exogenous is the numerous individuals with divergent interests who comprise a regime or a party. Parties are not unitary actors in a position to choose r strategically; leaders of a regime may desire to run a fair election but fail to control subordinates. Changing the value of r allows for the consideration of election environments of varying cleanliness on citizens' monitoring problem. Note that given my definition of fraud as reversing the no misconduct election result, the probability of fraud will also be a

FIGURE 1
Citizens' Payoff and Election Structure



function of the expected closeness of the election. Misconduct that is irrelevant in a landslide could decide a very close contest, as the 2000 U.S. Presidential contest demonstrates. Table 1 contains a summary list of notation.

If citizens choose to monitor the election, monitors provide a signal m that the election is either fair or biased, $m \in \{F, U\}$. The probability of a rigged election r is the citizens' common prior belief of fraud. The observers could conceivably make their report for strategic purposes; that is, the observers could be included as players in the game and choose m based on their payoffs. A false report would involve manufacturing or ignoring evidence of fraud. I do not allow this but instead treat the monitors' signal as a move by nature, which corresponds to a sincere report of their evidence. The availability of impartial monitors is critical in credibly detecting fraud, a point I return to in Section 5. Monitoring imperfectly detects fraud. Define $u(m|e)$, the probability signal m is sent given that the true election outcome is e, as follows:

TABLE 1
Summary Description of Notation

Variable	Description
e	True election outcome (F-fair, U-unfair)
w	Nominal (official) election outcome
m	Report from election monitors
B	Value of a fair election to citizen
d	Cost of protesting an election
D	Cost of a failed protest
$C(p,q)$	Cost of monitoring election
p	Probability of a false fraud signal from observers
q	Probability of a false fair signal from observers
v	Updated belief election is fair, based on monitors
r	Probability of fraud in a monitored election
α_0	Probability voter protests an unmonitored election
α_f	Probability voter protests on report of fair
α_u	Probability voter protests on report of unfair
σ	Probability a protest overturns official result

$$u(m = F\,|\,e = F) = 1 - p, \qquad u(m = U\,|\,e = U) = 1 - q,$$

$$u(m = U\,|\,e = F) = p, \qquad u(m = F\,|\,e = U) = q.$$

Observers make a false report of fraud with probability p and a false fair report with probability q. Assume $p, q \leq .5$.[6] Citizens update their beliefs using Bayes' Law based on the observers' signal. Let $v(e/m)$ be the posterior belief the true election outcome is e conditional on receiving m where:

$$v(e{=}F\,|\,m = F) = [(1{-}p)\cdot(1{-}r)]/[(1{-}p)\cdot(1{-}r){+}q\cdot r],\ v(e{=}U\,|\,m = F) = q\cdot r/[(1{-}p)\cdot(1{-}r) + q\cdot r];$$

$$(e{=}U\,|\,m = U) = [(1{-}q)\cdot r]/[(1{-}q)\cdot r + p\cdot(1{-}r)],\ v(e{=}F\,|\,m{=}U) = [p\cdot(1{-}r)]/[(1{-}q)\cdot r + p\cdot(1{-}r)].$$

Citizens decide how closely to observe the election if they decide to monitor at all. Devoting more resources to monitoring produces a more accurate information signal concerning the incidence of fraud. Effective monitoring of an election requires execution of a number of tasks over a period of several months before and after polling day. Tasks include registering voters, guaranteeing opportunities for campaigning by candidates, preventing intimidation of voters, maintaining the secrecy of and an accurate count of ballots, and resolving complaints about the process [Padilla and Houppert, 1993; Elklit and Svensson, 1997]. A poorly funded observation effort might only observe a fraction of polling places. Citizens can also direct monitors to guard against one of the two types of error in issuing their final report on the integrity of the election. By declaring the election to be fraudulent based on any observed irregularities, monitors can reduce the probability of a false fair report, although the probability of a false allegation of fraud will rise. Error cannot be eliminated totally: diligent and numerous observers will not detect all fraud (the perpetrators hide their misdeeds), while honest behavior may appear hopelessly fraudulent.[7] Let $C(p,q)$ be the cost of monitoring as a function of the reliability of the signal. Smaller values of p and q imply more accurate monitoring. Monitoring costs increase with the reliability of the signal. I impose the following properties on the cost function: C_p, $C_q < 0$, $C_p(0,\cdot) = C_q(\cdot,0) = -\infty$, C_{pp}, $C_{qq} > 0$, and $C_{pq} \approx 0$ (subscripts denote partial derivatives).

Setting aside a fraudulent result may be problematic in transition elections; the rule of law may not be well established and the regime able to subvert the judicial process. Observers may scrupulously document existence of fraud, only to have the regime jail the poll watchers, destroy their evidence and declare martial law. To allow this possibility, citizens decide following the election or receipt of the observers' report whether to "protest" the election, which may overturn the official result. I leave aside exactly what "protest" involves (letters and phone calls, litigation, marches and rallies, a general strike). Let a be the action chosen by voters; $\alpha \in \{A, P\}$ where P = protest and A = accept, in which case the official election result stands. A successful protest ensures the election outcome is fair; the official election result stands if the protest is unsuccessful. Protest in the model refers to citizens taking to the streets to effect a change in government.

Several factors jointly determine the probability of a successful protest, σ. The first is the extent of participation by citizens in protest. A fair election is a public good. Although the representative citizen approach employed here abstracts from explicit consideration of collective action problems, citizens might avoid participating in protest, particularly if the regime might resist and the protest turn violent. The effects of free riding can be incorporated by adjusting σ. If the number of participants required for a successful protest is sufficiently small, free riding may not have much effect on σ. A second determinant of σ is the regime's willingness to use violence and risk civil war to hold on to power. As Przeworski [1991] emphasizes, rulers in a democracy must more or less voluntarily give up power. My approach is sufficiently general to incorporate a leadership bound by the rule of law; this would amount to $\sigma = 1$ and a very low cost of protest (defined below). Upon presentation of sufficiently credible evidence of fraud citizens would then redress electoral chicanery. Also the rule of law would minimize the extent of collective action needed for protest.

Figure 1 presents the citizens' payoffs along with the structure of the election. Citizens' payoffs have the following components. The value of a fair election is $B > 0$, which essentially is the instrumental value of fair elections in controlling government and improving policy choices. The value of fair elections is a function of the amount of damage politicians unchecked by elections can impose on citizens. If fraud is sufficiently prevalent, the regime becomes authoritarian despite the trappings of democracy. Citizens receive this benefit if the election is fair, $e = F$, or if they successfully protest an election. Citizens get a payoff of 0 when an unfair election is not overturned. Participation in protest, regardless of its outcome, involves a cost $d > 0$ for the representative citizen. (Strictly speaking this cost is excludable and free riders can avoid it.) An unsuccessful protest entails further costs; a lengthy general strike which results in conflict entails added costs. Let D be the cost of an unsuccessful protest; this cost is non-excludable. The cost to participants in an unsuccessful protest is $d + D$.

THE VALUE OF MONITORING ELECTIONS

I first consider citizens' protest decision when elections are not monitored. Citizens' payoff if they accept the election is $(1 - r)\cdot B$ while their expected payoff from protest is

$$\sigma\cdot(B-d) + (1-\sigma)\cdot[(1-r)\cdot(B-d-D) - r\cdot(d+D)]. \quad (1)$$

Citizens will protest, provided $\sigma > 0$, if the probability of a fraudulent election is high enough,

$$r \geqslant (1-\sigma)\cdot(d+D)/(\sigma\cdot B) + d/B. \quad (2)$$

Let $\alpha_0{}^*$ denote the citizens' optimal action without information. The inequality in (2) cannot be satisfied for any $r \leqslant 1$ if σ is sufficiently small; citizens will not protest if their efforts are futile. Higher values of r and σ make citizens more likely to protest—that is, citizens will protest without information if protest is sufficiently likely to be rigged.

Let $\alpha_f{}^*$ and $\alpha_u{}^*$ give the citizens' optimal protest decisions upon receiving fair and unfair signals from the monitors. The availability of a protest option distinguishes two types of fraudulent elections. Politicians need not hide their misconduct if citizens lack the ability to protest. The occurrence of fraud is then common knowledge. Politicians must hide their offenses (take advantage of $q > 0$) when citizens can overturn an election result. Blatant fraud may disappear once citizens have an effective protest option, but elections may still be rigged. Both free riding and the lack of an effective protest option pose difficulties preventing election fraud.[8] Hence political science explanations of election fraud based solely on citizen attitudes B fail to address these aspects of the problem.

The value of monitoring depends on the citizens' actions given each signal m and also citizens' payoff with an unmonitored election. I first consider adoption of a monitoring system of exogenous quality—fixed p and q—and then consider citizens' optimal choices of p and q. To simplify the expressions and focus on the role of monitoring, I assume now that protest is perfectly effective, $\sigma = 1$.

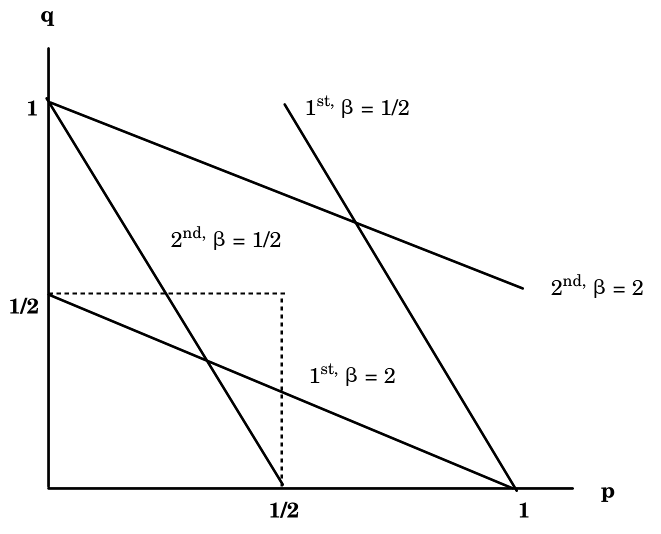
The net payoffs from protest for citizens without monitoring and then contingent on each possible report are

$$NPP(m{=}\varnothing) = r\cdot B - d;$$

$$NPP(m{=}U) = [B\cdot(1-q)\cdot r]/[p\cdot(1-r) + (1-q)\cdot r - d];$$

$$NPP(m{=}F) = [B\cdot q\cdot r]/[(1-p)\cdot(1-r) + q\cdot r - d].$$

FIGURE 2



Voters respond optimally in each case. Monitoring information must allow citizens to discriminate in their response, $\alpha_f^* = A$ and $\alpha_u^* = P$, or its value is zero, as the following proposition establishes. (All proofs appear in in the Appendix.)

Proposition 1

The value of the monitoring report is zero unless

(3) $$(1-p)/q \geqslant [(B-d)\cdot r]/[d\cdot(1-r)] > p/(1-q).$$

Figure 2 illustrates the two constraints in (3). Let $\beta = (B-d)\cdot r/\ [d\cdot(1-r)]$. The first constraint in (3) can be written as $q \leqslant 1/\beta - p/\beta$ and the second as $q < 1 - p/\beta$. Figure 2 graphs each constraint for sample values of $\beta = 1/2$ and $\beta = 2$. Note that only one of the two constraints affects the region $p, q \leqslant 1/2$, so only one of the constraints binds.

The first constraint in (3) ensures that voters accept the election result on a fair report from the observers. The cost of accepting a fair report is the possibility of allowing a fraudulently elected government to hold power. Citizens must trust the monitors' report sufficiently to allow a fair report to stand. The second constraint ensures that voters protest on a report of fraud. The cost of error to voters here is unnecessarily incurring the cost of protest. Citizens must have enough confidence in the monitors' report to act on a signal of fraud. The dual constraints in (3) combine to generate counterintuitive results. For instance, intuitively a high value of fair elections B should make voters more likely to employ a given monitoring system. Yet this intuition is only half right. A large value of B helps ensure satisfaction of the second inequality but also makes violation of the first inequality more likely. The value of fair elections may render the cost of possibly letting a fraudulent result stand excessively high.[9] A low prior estimate of fraud r can lead to violation of the second inequality. If citizens believe fraud is unlikely, they dismiss a report of cheating as erroneous instead of incurring the cost of overturning the official result. Note that Proposition 1 establishes only when the value of monitoring information is positive; this value may still be less than the cost of monitoring.

Optimal choices of p and q maximize the expected net value of monitoring,

$$NV = B\cdot(1-q\cdot r) - d\cdot[(1-q)\cdot r + p\cdot(1-r)] - C(p,q). \tag{4}$$

Let p^*, q^* maximize (4) and let NV* be the maximized net value of monitoring. (3) must be satisfied and NV* ≥ 0 for the election to be worth monitoring. Assuming these conditions are met, the following first order conditions describe optimal monitoring:

$$-C_p = d\cdot(1-r); \tag{5}$$

$$-C_q = (B-d)\cdot r. \tag{6}$$

Optimal monitoring equates the marginal cost of improving each element of signal quality with its marginal benefit. A reduction in p, the probability of a false fraud signal, avoids the cost of protest against a fairly-elected government. A reduction in q, the probability of a false fair signal, increases the likelihood of a fairly elected government, so the benefit is B less the cost of protest. The comparative statics of optimal monitoring are:

Proposition 2

The comparative statics of optimal monitoring are as follows:

Parameter	**Effect on false rigged report *p****	**Effect on false fair report *q****
Probability of fraud, r	increase	decrease
Value of fair elections, B	no effect	decrease
Cost of protest, d	decrease	increase
Marginal cost of verifying charges, C_p	increase	no effect
Marginal cost of detecting fraud, C_q	no effect	increase

Comparative statics indicate that changes in the exogenous variables do not have parallel effects on each of the error probabilities. This is a consequence of these variables' differential impact on the cost of each type of error for the voters. For instance, the value of a fair election B affects the cost of allowing a rigged election to stand. Consequently an increase in B reduces the optimal probability of a false fair signal, q^*, and has little effect on p^*. An increase in the cost of protest d reduces the net value of correcting a fraudulent election and has the opposite effect of an increase in B on q^*. But an increase in d also raises the cost of protest against a fair result, which decreases p^*. An increase in the marginal cost of either element of signal quality increases the optimal probability of that type of error.

An increase in the prior probability of fraud r reallocates monitoring effort from verifying fraud charges (the probability of a false bias signal p increases) to detecting fraud (the probability of a false fair signal q falls). Citizens guard against fraud going unnoticed if they expect a rigged election, but guard against false fraud signals if they expect a fair election. Note that the expected closeness of an election also affects the probability of fraud. Consequently, optimal monitoring involves greater effort to detect fraud in a close election and more effort to verify charges in a lopsided contest.

Proposition 2 suggests that cycles in fraud might be possible. Although my model is static, consider the following dynamic. When politicians behave and elections are typically fair, prior probability of fraud r falls and voters allocate less effort toward detecting fraud that they do not expect to find (or dismiss evidence of fraud). Politicians note this and are emboldened to push the envelope. Voters revise upward their expectation of fraud and generate momentum for a new wave of reforms, leading to closer monitoring and inducing politicians to behave once more. Sabato and Simpson [1996] present evidence of fraud in several recent elections and argue that fraud indeed is on the rise in the United States. The dynamic suggested here is similar to Clark and Lee's [2001] argument that cycles in trust in government could occur.

Optimal monitoring depends on many exogenous factors. Consequently the same values for prior probability of fraud and the probability of a false fair signal will not be optimal for all parameter values. Optimal standards of proof for election fraud will vary between democracies and over time. Western democracies must be careful not to export their own evidentiary standards for validating election fraud charges to a new democracy. Expectations of fraud are typically high in new democracies, so voters will allocate effort toward detecting fraud. Citizens in this case will want to expose cheaters, even at the expense of falsely accusing some politicians of misconduct. Carothers [1997] notes, and according to Proposition 2 rightly criticizes, the tendency of many international observers to tolerate fraud in nations with little history of democracy. As democracy matures, expectations of fraud fall and voters allocate more effort toward investigating the validity of fraud charges with something approximating an "innocent until proven guilty" standard being applied.

A country with unmonitored elections probably could not be classified as a genuine democracy. Protest involves taking to the streets to enforce a change in government. If citizens always protest without monitoring, α_0^* = protest, the politics of the street demonstration prevail and elections play a minimal role. But citizens always accepting the result, α_0^* = accept, is unsustainable as an equilibrium. Citizens think fraud is sufficiently unlikely that they accept election results; but if politicians know citizens will accept the election result, they will steal the election, disconfirming citizens' prior estimate of fraud.[10] Democracy would have to involve only occasional protest, regulated by reports of fraud. Nonetheless, citizens may not choose to monitor elections, $NV^* < 0$, *even though* they value fair elections, $B > 0$, because monitoring does not provide an improvement over street demonstrations.

Application of the envelope theorem to (4) indicates which factors increase the net value of monitoring, and therefore when elections are likely worth monitoring. An increase the value of fair elections raises NV^* (provided monitoring ensures $q \cdot r < 1$). Numerous studies document the empirical correlation between income and democracy; most nations above a threshold level of per capita income (about US$2,000 in 1990) are democratic, while most nations below the threshold are not [Huntington, 1991]. Many scholars have noted the connection and have accepted a preference-driven view—that democracy is a normal good [Minier, 2001]. Monitoring provides a mechanism by which income leads to democracy. If B increases with national income, a threshold level of national income exists (given the other parameters) for elections to be worth their cost. Low income countries do not value fair elections sufficiently (given other budgetary priorities) to fund the investment necessary to monitor the contest worthwhile. On the other hand, given the level of monitoring poor countries can afford, the resulting elections might well be rigged. Hence citizens have little reason to agitate for democracy before they can afford to conduct fair elections.[11]

A reduction in the cost of monitoring elections nations increases the net value of democracy. Differences in monitoring costs across nations might exist due to geography, communications, or transportation. Innovations such as registration, the secret ballot, and the voting machine lowered the cost of detecting fraud at the level of casting ballots [O'Leary, 1962; Fredman, 1968; McCormick, 1953]. The development of parallel vote tabulations, based on results from a subset of polling places and released shortly after the polls close, is a recent innovation in this regard [Garber and Cowan, 1993]. When sufficient data on the electorate is available, only a small but statistically representative number of polling places must be observed, reducing cost further. Recent relaxation of registration requirements, the Motor Voter law, and California's "honor system" in voting, however, increase the cost of detecting fraud. Fraud at the level of counting votes, due to possible collusion between poll workers, is not affected as much by such reforms. Hence the mere adoption of the Australian ballot was insufficient to eliminate fraud, as experience in the U.S. South demonstrates [Kousser, 1974].

EXTENSIONS: FUTURE ELECTIONS AND LOCAL ELECTIONS

Repeated elections magnify the effect of an increase in the prior probability of fraud. Suppose voters update their expectation of fraud in future elections using the report from today's election. Assume r is higher for the next k elections if m_1 = U than if m_1 = F. A current report of fraud reduces the value of the next election by

$NV_2{}^*[r_2(m_1\text{=F})] - NV_2{}^*[r_2(m_1\text{=U})]$ where $r_2(\cdot)$ is the prior for election #2 contingent on m_1. Let δ be the discount factor. The total cost of fraud report in today's election is

$$E = \sum_{j=1}^{k} \delta^j * \{NV_j * [r_{j+1}(m_1 = F)] - NV_j * [r_{j+1}(m_1 = U)]\} \tag{7}$$

The cost is incurred if and only if m_1 = U, so the net expected value of monitoring the current election is now:

$$NV = \text{B}\cdot(1-q\cdot r) - (d + E/N)\cdot[(1-q)\cdot r + \text{p}\cdot(1-r)] - \text{C}(p,q)/N. \tag{8}$$

The effect on optimal monitoring is as follows.

Proposition 3

A dependence of the prior estimate of fraud in future elections the report on the current election report reduces the optimal value of p and increases the optimal value of q.

The influence on future expectations raises the cost of m = U and increases the optimal probability of a false fair report. But actual fraud then has a higher probability of not being reported. Ignorance of election fraud can be bliss. Loss of confidence in the system offsets the value of reversing today's electoral crime. While Nixon's loss to Kennedy in the 1960 election may have been due to fraud, the cost of righting the wrong may well have exceeded the benefit.

Repeated elections also provide the opportunity for voters to punish a party guilty of fraud in an earlier election. The Republicans suffered heavy losses in the 1974 elections three months after Nixon's resignation from office. The prospect of future punishment even if the current election stands provides an incentive for contestants to hold a fair election. Future electoral punishment for current fraud, though, requires prior establishment of some degree of fairness. Citizens cannot punish the perpetrators of today's fraud at the polls tomorrow if all elections are totally rigged.

Election fraud was an integral component of the political machines that at times dominated New York, Chicago, Philadelphia, Kansas City, Jersey City, and other cities. Yet no similar machines have existed at the national level. British electoral history also suggests election fraud has occurred more often at the local than national level.[12] What might explain this difference?

Fair elections are a public good; fair municipal elections are a local public good. Local communities can cater to the different preferences of citizens for fair elections in the manner of the Tiebout model of local public goods provision. The model in Section 2 assumes all voters place the same value on fair elections. But suppose citizens are heterogeneous, specifically that there are two types of preferences toward fair elections, B_h and B_ℓ with $B_h > B_\ell$. Assume that citizens who place a high value on fair elections would choose to have elections monitored while citizens with a low value of fairness do not find elections worth monitoring given the available technology; that is, the maximized net value of monitoring from (4) is positive for B_h and negative for B_ℓ. Citizens with B_ℓ would prefer to live in a community with unmonitored elections than to pay higher taxes for fair elections. A stable equilibrium could involve citizens with B_h living in community A where elections are monitored and citizens with B_ℓ living in community B. Although constitutional choice—determination of which public goods will be provided in which communities— is exogenous in the Tiebout model (a range of communities is assumed to exist), the model could apply to monitoring elections. A national election could well be monitored in this case because the average value of B might be sufficiently high that the maximized net value of a monitored election is positive. Within a federal system equilibrium could involve some communities with fair elections and some rotten boroughs, while national elections would be monitored.[13]

Movement of individuals across communities could help entrench corruption if citizens are not initially perfectly sorted across communities. A citizen with B_h living in city B where elections are not monitored will move to A if the cost of moving is less than the net value of fair elections. The cost and possible ineffectiveness of a protest option strengthens the incentive for citizens with B_h to leave community B, especially since the probability of successful protest might be zero once there are sufficiently few B_h's in community B. The citizens who place the greatest value on fair elections, the potential constituency for local election reform, exit first, an example of exit deflating the effectiveness of voice, as emphasized by Hirschman [1970].

DISCUSSION

I have considered a single election in this paper. Several simultaneous elections allow political parties to share the gains from fraud by dividing up the available positions, leading to fraud at the level of counting ballots.[14] Party 2's election officials may overlook miscounting of votes for one office if they are being compensated elsewhere. Allegations of fraud in the 1993 Russian elections involved a trade of votes for Yeltsin's constitution for seats in the Duma [Myagkov and Sobyanin, 1995].

My effort in this paper has been theoretical—to understand what voters must do to ensure elections are fair. The comparative statics results established here provide potentially testable hypotheses about monitoring effort and the incidence of fraud if a good measure of fraud were available, a substantial qualification. Allen and Allen emphasize the lack of systematic evidence concerning fraud: "The evidence to demonstrate the existence of election fraud in the literature is not only anecdotal, it is unsystematic, impressionistic, and by and large inconclusive" [1981, 179]. Political historians have produced a number of case studies documenting fraud in a specific locale, but these would not suffice for an empirical test. Contested elections in Congress or state legislatures comprise a possible measure of fraud, but as Argersinger [1985] notes, these complaints were typically decided on partisan grounds and may not be a very accurate measure of fraud. Convictions of individuals for violations of election laws would be another measure, but would require a tremendous research project to search court records and create a tally suitable for empirical work. As a final complication, recall that in the absence of an effective protest option politicians do not have to hide their misdeeds from public view. The publicly perceived level of fraud is not necessarily an accurate measure of misbehavior.[15]

My model emphasizes the relevance of both demand and cost factors in determining whether to monitor elections. These factors provide insight on the elimination of corrupt election practices in the United States and Britain. On the cost side, innovations such as registration lists, the Australian ballot, and the voting machine lowered the cost of producing a signal of a given quality. The availability of a protest option, the ability to overturn the results of a rigged election, was another significant factor. Parliament in Britain played a role in investigating fraud and disfranchising rotten boroughs. The federal and state governments in the United States provide a check on local misconduct; federal intervention can overturn a fraudulent election even if the local administration is thoroughly corrupt. The failure of the Lodge elections bill in 1890 weakened the protest option in the post-Reconstruction South. On the demand side real per capita income grew in the United States and Britain in the 1800s, increasing people's willingness to pay for fair elections. An increase in scope of government also increased the importance of fair elections in ensuring the quality of policymaking. As long as government remained small and distributed patronage, the quality of decision-making was not very important. When government exercises more control over the economy and peoples' lives, fair elections become more important.[16]

Recent treatments of the monitoring process rely on the integrity of observers to honestly report their findings: "Essential to the success of any election observation, whether international or domestic, is credibility. Credibility flows from neutrality, the nonpartisan and evenhanded character of the observation effort" [Nevitte and Canton, 1997, 50]. Similarly I do not include monitors as strategic players in this model. Parties may provide observers, but if false reports of fraud can be cheaply generated, citizens may discount these monitors. International observers with no ties to any faction in the contest provide the most common solution to a lack of neutral domestic observers. Thus organizations like the Carter Center, the National Democratic Institute for International Affairs, the International Republican Institute, the Organization of American States, and the United Nations have provided observers for recent transition elections. Nonetheless, even international observers may have an agenda to pursue and be less than neutral, as Carothers [1997] notes. If impartial observers were unavailable either domestically or internationally, citizens might need to pay observers. Election monitors could, in the spirit of efficiency wage theory, be paid a premium for their services and monitored themselves for evidence of inaccurate reporting concerning misconduct. This would increase the cost of monitoring, and would be another example of Clark and Lee's [2001] argument that a lack of trust increases the cost of government.

How effective have international observation efforts been in recent decades? Some notable examples of blatant fraud being uncovered and punished exist, the most notable probably being the Philippines in 1986. As discussed above, a reliable general measure of the integrity of elections is elusive, because the presence of monitors and the availability of a protest option may force fraud to take on less perceptible forms. The comparative country ratings provided by Freedom House provide some evidence on this point. Between 1991-92 and 2000-01, the Freedom House political rights score improved for 57 countries and got worse for 41 countries. During this decade 147 countries in the world had less-than-perfect political rights scores for at least one year, so almost forty percent of countries that could improve did so.[17] Some prominent improving countries like El Salvador, India, Indonesia, the Philippines and South Africa had closely observed elections during the decade.

My investigation reveals several useful roles for organizations trying to further the cause of democracy worldwide. One is the provision of neutral election observers. Perhaps only outsiders without ties to any faction may be trusted by all sides to provide a sincere fraud report. The second is subsidization of the cost of monitoring an election. A poor nation may be unable to afford very reliable election results and may not value the elections it can afford. Provision of in-kind aid to stage an election, such as election observers, minimizes the probability of misappropriation of assistance for unintended purposes. The third is strengthening the protest option. Locals may lack the ability to punish perpetrators of election fraud, due to either collective action problems or the regime's willingness and ability to use force to retain power. The international community can tie aid, loans, and development projects to the correction of a fraudulent election result and employ economic sanctions and military intervention if necessary. The 1994 U. S. intervention in Haiti represents a foreign supplied punishment option.

APPENDIX

PROOF OF PROPOSITION 1

Citizens' payoff for $\alpha = P$ is $B - d$, B for $\alpha = A$ if $e = F$, and 0 for $\alpha = A$ if $e = U$. Let $u(\alpha_m)$ be citizens' expected utility from action α when the monitors' report is m. These payoffs are

(A.1a) $$u(\alpha_u{=}P) = (B-d)\cdot v(e{=}F \mid m{=}U) + (B-d)\cdot v(e{=}U \mid m{=}U) = B - d;$$

(A.1b) $$u(\alpha_u{=}A) = B\cdot v(e{=}F \mid m{=}\mathrm{U}) + 0\cdot v(e{=}U \mid m{=}U) = [Bp(1-r)]/[(1-q)r+p(1-r)];$$

(A.1c) $$u(\alpha_f{=}P) = (B-d)\cdot v(e{=}F \mid m{=}F) + (B-d)\cdot v(e{=}U \mid m{=}F) = B - d;$$

(A.1d) $$u(\alpha_f{=}A) = B\cdot v(e{=}F \mid m{=}F) + 0\cdot v(e{=}U \mid m{=}F) = [\,\mathrm{B}(1-p)(1-r)]/[(1-p)(1-r)+qr].$$

Citizens respond optimally to each signal. $NPP(m{=}U) = u(\alpha_u{=}P) - u(\alpha_u{=}A)$ while $NPP(m{=}F) = u(\alpha_f{=}P) - u(\alpha_f{=}A)$. The optimal action on a given signal is protest if $NPP(\varnothing) > 0$ and accept if $NPP(\varnothing) \leqslant 0$. Since $NPP(m{=}U) > NPP(m{=}F)$, citizens have three possible optimal decision rules contingent on the monitors= report: [1] $\alpha_f^* = \alpha_u^* = A$; [2] $\alpha_f^* = A$, $\alpha_u^* = P$; [3] $\alpha_f^* = \alpha_u^* = P$. Let $\pi(m{=}F)$ and $\pi(m{=}U)$ be the probability each signal is received from the monitors. These probabilities are

(A.2a) $$\pi(m{=}F) = (1-p)\cdot(1-r) + q\cdot r;$$

(A.2b) $$\pi(m{=}U) = p\cdot(1-r) + (1-q)\cdot r.$$

Expected utility for each rule is $u(\alpha_f^*) - \pi(m{=}F) + u(\alpha_u^*{=}U) - \pi(m{=}U)$. Using the payoffs in (A.1) and the probabilities in (A.2) we have for each rule:

(A.3a) $$[1]:\ B-(1-p)\cdot(1-r) + B\cdot p\cdot(1-r) = B\cdot(1-r);$$

(A.3b) $$[2]:\ B\cdot(1-p)\cdot(1-r) + (B-d)\cdot[(1-\mathrm{q})\cdot r + p\cdot(1-r)] = B\cdot(1-\mathrm{q}\cdot r) - d\cdot[(1-\mathrm{q})\cdot r + p\cdot(1-r)];$$

(A.3c) $$[3]:\ (B-d)\cdot[(1-p)\cdot(1-r)+\mathrm{q}\cdot r] + (B-d)\cdot[(1-\mathrm{q})\cdot r+p\cdot(1-r)] = B - d.$$

The value of the information signal also depends on citizens' action and utility with and without information. Since $\mathrm{NPP}(m{=}U) > \mathrm{NPP}(\varnothing) > \mathrm{NPP}(m{=}F)$, four cases must be considered to establish when the value of information is positive.

[a]: [1] with $\alpha_0^* = A$; [b]: [3] with $\alpha_0^* = P$;
[c]: [2] with $\alpha_0^* = A$; [d]: [2] with $\alpha_0^* = P$.

Examination reveals that citizens receive the same expected utility with or without information in cases [a] and [b], so the value of information then is zero. A necessary condition for the value of the monitors' signal to be positive is therefore $\alpha_f^* = A$ and $\alpha_u^* = P$. Consequently we must have $\mathrm{NPP}(m{=}F) \leqslant 0$ and $\mathrm{NPP}(m{=}U) > 0$. From (A.1c) and (A.1d) we see that $\mathrm{NPP}(m{=}F) \leqslant$ if and only if

(A.4) $$B\cdot(1-p)\cdot(1-r) \geqslant (B-d)\cdot[(1-p)\cdot(1-r) + q\cdot r].$$

Simplification and rearrangement yield

(A.5) $$d\cdot(1-p)(1-r) \geqslant (B-d)\cdot q\cdot r,$$

from which the first constraint in (3) follows. From (A.1a) and (A.1b) we see that $\mathrm{NPP}(m{=}U) > 0$ if and only if

(A.6) $$(B-d)\cdot[(1-q)\cdot r + p\cdot(1-r)] > B\cdot p\cdot(1-r).$$

Again simplification and rearrangement yield

(A.7) $$(B-d)\cdot(1-\mathrm{q})\cdot r > d\cdot p\cdot(1-r),$$

from which the second constraint in (3) follows.

PROOF OF PROPOSITION 2

The Hessian matrix of (4) is

(A.8)
$$H = \begin{bmatrix} -C_{pp} & -C_{pq} \\ -C_{pq} & -C_{qq} \end{bmatrix}.$$

$|H| = (C_{pp} \cdot C_{qq} - C_{pq}{}^2)$, which is greater than zero since $C_{pq} \approx 0$, so the second-order condition for maximization is satisfied.

The comparative statics effect of any parameter a on the optimal choices of p and q is given by

(A.9)
$$\begin{bmatrix} \frac{\delta^2 NV}{\delta p^2} & \frac{\delta^2 NV}{\delta p \delta q} \\ \frac{\delta^2 NV}{\delta p \delta q} & \frac{\delta^2 NV}{\delta q^2} \end{bmatrix} \begin{bmatrix} \frac{\delta p*}{\delta q} \\ \frac{\delta q*}{\delta a} \end{bmatrix} = \begin{bmatrix} \frac{-\delta^2 NV}{\delta p \delta a} \\ \frac{-\delta^2 NV}{\delta q \delta a} \end{bmatrix}.$$

The effects of the parameters in Proposition 2 are as follows. All derivatives contain a common factor $1/|H| > 0$.

	$\delta p*$	$\delta q*$
δr	$-C_{qq} \cdot (B-d) < 0$	$C_{pp} \cdot d > 0$
δB	$-1 \cdot r \cdot C_{pq} \approx 0$	$-1 \cdot r \cdot C_{pp} < 0$
δd	$-C_{qq} - (1-r) < 0$	$C_{qq} \cdot r > 0$
δC_p	$-C_p - C_{qq} > 0$	$C_p \cdot C_{pq} \approx 0$
δC_q	$C_q - C_{pq} \approx 0$	$C_q \cdot C_{pp} > 0$

NOTES

I would like to thank Tyler Cowen, Amihai Glazer, Harvey Palmer, Marc Poitras, Alex Tabarrok, the editor of this *Journal*, three referees, and participants at the Southern Economics Association Public Choice Society meetings for helpful comments. I am solely responsible for any errors and omissions in the text.

1. Hersch [1997] presents evidence of fraud in the 1960 presidential contest. Sabato and Simpson [1996] detail the persistence of fraud in U.S. elections.
2. Incumbent politicians might also collude to rig elections against challengers.
3. The majority party in the legislature could also fail to act on evidence of fraud. Argersinger [1985, 682] offers the following description of the behavior of state legislatures and Congress in the late 1800s: "It made little sense to go to the expense and trouble of contesting elections, regardless of grounds, unless one 's own party already controlled the chamber. ... The Republican defense against such charges of was that such practices were followed by all parties, and therefore, not of particular relevance."
4. My model does not explicitly allow for ex post investigations of fraud like the special commissions established to examine rotten boroughs in Britain in the 1800s [O'Leary, 1962] or two-part screening of elections.
5. Misconduct at the level of casting ballots which involves vote selling could allow registering the strength of preferences and be Pareto improving (Anderson and Tollison 1990). My model takes designation of corrupt practice as exogenously determined and could allow vote buying to be legal.
6. Given two outcomes, .5 is a maximally uninformative signal. A larger probability of error implies all information provided by the signal is not being used.

7. Bribes make both the candidate and the voter better off so the participants have an incentive to prevent detection of the payment. Detection of mutually beneficial corruption is more difficult than intimidation, which makes the affected voters worse off [Shleifer and Vishny, 1993]. The number of polling places is often quite substantial; the 1988 elections in Chile involved 22,000 polling sites, Pakistan in 1988 33,000 sites, and the Philippines in 1986 over 85,000 sites. For details on recent election observation efforts see Padilla and Houppert [1993], Garber and Cowan [1993] and Nevitte and Canton [1997].
8. A fair election is a step public good and is analogous to the Paradox of Revolution [Tullock, 1971].
9. If the value of fair elections varies across the population, those with the largest values of B may reject election procedures the rest of the population finds acceptable. The charges of the Progressives can be reconsidered in this regard.
10. Citizens might protest a certain percentage of elections to try to deter fraud. But without information about whether fraud has occurred, the official winner must be punished in all protested elections. Since both innocent and guilty politicians are punished, the deterrent effect is minimal.
11. Triesman [2000] finds very strong evidence that economic development reduces corruption, although he does not discuss whether increased income allows employment of better technologies to control corruption.
12. For evidence on this point see Bryce [1908, II], O'Leary [1962], Cook [1973], and Allen and Allen [1981].
13. Triesman [2000] finds that the incidence of corruption is higher in federal states than unitary systems, consistent with this conclusion. Communities could also cater to preferences for party government. Local Republican and Democratic communities within the same metropolitan area could be provided through Tiebout competition. Note that collective action problems should be more severe at the national level, hindering participation in monitoring efforts or protest, which might even eliminate the incidence of fraud.
14. An incentive for reform in Britain was the rising cost to the parties of fraudulent activities [O'Leary, 1962]. Fraud at the level of casting ballots, such as bribes, is expensive since the parties compete in these illicit activities. Fraud at the level of counting ballots involves collusion and can substantially reduce competitive expenditures (both legitimate and illegitimate) by the parties.
15. Carothers observes: "Although blatant fraud still occurs, efforts by entrenched leaders to manipulate electoral processes to their advantage have become more subtle as such leaders have been socialized into the new world of global democracy and internationally observed elections" [1997, 22].
16. Causation may run in the opposite direction here: only when fair elections ensured control of politicians were citizens willing to delegate additional tasks to government.
17. The political rights rating ranges from 1, most free, to 7, least free. The ratings were taken from Freedom House 's website at www.freedomhouse.org/ratings/index.htm.

REFERENCES

Anderson, G. M. and Tollison, R. D. Democracy in the Marketplace, in *Predicting Politics,* edited by W. M. Crain and R. D. Tollison. Ann Arbor: University of Michigan Press, 1990.

Allen, H. W. and Allen, K. W. Vote Fraud and Data Availability, In *Analyzing Electoral History,* edited by J. M. Clubb, W. H. Flanigan, and N. H. Zingale. Beverly Hills: Sage Publications, 1981.

Argersinger, P. H. New Perspectives on Election Fraud in the Gilded Age. *Political Science Quarterly*, Winter 1985, 100, 669-87.

Becker, G. S. Crime and Punishment: An Economic Approach. *Journal of Political Economy*, March/April 1968, 76, 169-217.

Benson, G. C. S. *Political Corruption in America*. Lexington MA: Lexington Books, 1978.

Bliss, C. and **Di Tella, R.** Does Competition Kill Corruption? *Journal of Political Economy*, October 1997, 105, 1001-1023.

Brennan, G. and Buchanan, J. M. *The Reason of Rules*. Cambridge: Cambridge University Press, 1985.

Bryce, J. *The American Commonwealth*. New York: Commonwealth Publishing, 1908.

____________. *Modern Democracies*. London: Macmillan and Company, 1929.

Carothers, T. The Observers Observed. *Journal of Democracy*, July 1997, 7, 17-31.

Clark, J. R. and **Lee, D. R.** The Optimal Trust in Government. *Eastern Economic Journal*, Winter 2001, 19-34.

Cook, F. J. *American Political Bosses and Machines*. New York: Franklin Watts, 1973.

Elklit, J. and Svensson, P. What Makes an Election Free and Fair? *Journal of Democracy*, July 1997, 32-46.

Fredman, L. E. *The Australian Ballot: The Story of an American Reform*. East Lansing MI: Michigan State University Press, 1968.

Garber, L. and Cowan, G. The Virtues of Parallel Vote Tabulations. *Journal of Democracy*, April 1993, 95-107.

Goldberg, R. Election Fraud: An American Vice, in *Elections American Style*, edited by A. J. Reichley. Washington, DC: Brookings Institution, 1987.

Heckelman, J. C. The Effect of the Secret Ballot on Voter Turnout Rates. *Public Choice*, January 1995, 107-124.

Hersh, S. M. *The Dark Side of Camelot*. Boston: Brown, Little and Company, 1997.

Hirschman, A. O. *Exit, Voice, and Loyalty*. Cambridge, MA: Harvard University Press, 1970.

Huntington, S. P. *The Third Wave*. Norman, OK: University of Oklahoma Press, 1991.

Key, V. O. *Southern Politics*. New York: Alfred A. Knopf, 1949.

Kousser, J. M. *The Shaping of Southern Politics*. New Haven, CT: Yale University Press, 1974.

Laffont, J. J. *The Economics of Uncertainty and Information*. Cambridge, MA: MIT Press, 1989.

Mauro, P. Corruption and Growth. *Quarterly Journal of Economics*, August 1995, 681-712.

McCormick, R. P. *The History of Voting in New Jersey*. New Brunswick, NJ: Rutgers University Press, 1953.

Minier, J. Is Democracy A Normal Good? Evidence from Democratic Movements. *Southern Economic Journal*, April 2001, 996-1009.

Myagkov, M. and Sobyanin, A. December 1993 Elections in Russia: Fraud or Irregularities? Paper presented at Public Choice Society Meetings, 1995.

Nevitte, N. and Canton, S. A. The Role of Domestic Observers. *Journal of Democracy*, July 1997, 47-61.

O'Leary, C. *The Elimination of Corrupt Practices in British Elections 1868-1911*. Oxford: Oxford University Press, 1962.

Padilla, D. and Houppert, E. International Election Observing: Enhancing the Principle of Free and Fair Elections. *Emory International Law Journal*, 1993, 73-132.

Przeworski, A. *Democracy and the Market*. Cambridge: Cambridge University Press, 1991.

Rose-Ackerman, S. *Corruption: A Study in Political Economy*. New York: Academic Press, 1978.

Sabato, L. J. and **Simpson, G. R.** *Dirty Little Secrets*. New York: Times Books, 1996.

Shleifer A. and Vishny, R. W. Corruption. *Quarterly Journal of Economics*, August 1993, 599-617.

Sikes, E. R. *State and Federal Corrupt-Practices Legislation*. Durham, NC: Duke University Press, 1928.

Stigler, G. The Optimum Enforcement of Laws. *Journal of Political Economy*, May/June 1970, 526-36.

Treisman, D. The Causes of Corruption: A Cross-National Study. *Journal of Public Economics*, June 2000, 399-457.

Tullock, G. The Paradox of Revolution. *Public Choice*, Fall 1971, 11, 89-99.