Center for Mathematical Studies in Economics and Management Science
Northwestern University, Evanston, Illinois 60208

Discussion Paper No. 956

EFFECTIVENESS OF ELECTORAL SYSTEMS FOR REDUCING GOVERNMENT CORRUPTION: A GAME-THEORETIC ANALYSIS

by

Roger B. Myerson*

September 1991

Abstract. A theoretical model is developed for predicting the relative effectiveness of different electoral systems for reducing government corruption. We consider voting games in which parties with known corruption levels and known positions on a major policy question are competing for legislative seats. We find that approval voting and proportional representation are fully effective, in the sense that all equilibria exclude corrupt parties from legislative seats. Plurality voting is partly effective, in the sense that there always exist some equilibria that exclude corrupt parties. Borda voting is ineffective because, for some political situations, no equilibria can guarantee the exclusion of corrupt parties.

*Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60208.

EFFECTIVENESS OF ELECTORAL SYSTEMS FOR REDUCING GOVERNMENT CORRUPTION: A GAME-THEORETIC ANALYSIS

by Roger B. Myerson

1. Introduction

Government officials enforce laws to restrain citizens from destructive conflict, and citizens vote in elections to restrain government officials from abuse of their power. Riker (1982) has urged political theorists to recognize that removing government officials who are corrupt or abuse their power is a fundamental objective of democratic institutions. In this paper, we develop a game-theoretic model to show how electoral systems may differ in their effectiveness for reducing government corruption.

Theoretical models of politics have commonly assumed that government policy options correspond to points in some finite-dimensional Euclidean space, and that each voter prefers policies that are closer to his or her own ideal point, which is somewhere in this space. In this view, the political problem is caused by the fact that individual voters have different ideal points. Corruption levels could be considered as policy dimensions in such a model, where they would represent levels of spending for the personal benefit of government officials. However, these corruption levels would be dimensions in which a preference for less would be unanimously shared by all voters (other than the politicians and their families, who are assumed to be a negligibly small portion of the electorate).

If there are only two established parties, and they compete for votes in a single election by independent noncooperative selection of their policy positions, then in equilibrium they should select the minimal corruption levels. There are, however, some difficulties with this simple (Bertrand-like) conclusion. If it is really fixed that only two given parties can exist (and decision-making in each party is controlled by a rational optimizing leadership), then we might expect them instead to reach a collusive agreement that improves the welfare of the leaders in both parties by maintaining higher corruption levels. That is, the one-stage noncooperative analysis neglects many other equilibria that exist when two parties play in a repeated game; and we might expect that, in the long run, the leaders of the two established parties would collusively focus on high-corruption equilibria that are better for them both. So eliminating corruption may require the possibility of multiparty elections, in which parties that are known to behave corruptly can be challenged by new parties. There are many different multiparty electoral systems, however, and they differ greatly in their incentive properties.

In this paper we develop a theoretical model to predict the relative effectiveness of multiparty legislative electoral systems for eliminating elected officials who are known to be corrupt. We find that different electoral systems may indeed differ in their effectiveness for reducing corruption, even when some challengers are known to be incorruptible. In particular, our model will suggest that proportional representation and approval voting may be more effective at reducing corruption than plurality voting, and plurality voting may in turn be more effective than Borda voting.

Difficulty in removing corrupt officials can arise because, under some electoral systems, voters who desire some specific government policy may find that, to maximize the probability of achieving their desired policy, they must give support to corrupt candidates who are committed to this policy. That is, some electoral systems may enable a corrupt candidate to get support from voters essentially by holding their preferred policy position as a hostage. Even if there exist noncorrupt candidates who are committed to the same policy, some electoral systems can make it disadvantageous for individual voters to transfer support away from corrupt candidates, when others' expected votes are taken into account.

2. The basic model

We consider here a model of unicameral legislative elections in which candidates are known to differ in their corruption levels, and in which there is at least one additional policy question (other than that of reducing corruption) about which voters may have some disagreement. Subject to the inclusion of these parameters, our model is intended to be as simple as possible.

We let T denote the set of *parties* which are competing for seats in the legislature, where T is a nonempty finite set. Each party offers candidates for every seat in the legislature. For each party r in T, we assume that party r has a *corruption level* c(r) which is fixed and known to all of the voters. That is, we are simplifying our model by assuming that each politician's propensity to take money from the taxpayers is a known function of his party affiliation; and we are ignoring the difficult problems of monitoring government officials and discovering their abuses of power. (To justify such simplifications, we may ask, if an electoral system cannot effectively eliminate corruption in such simple situations, then how can we hope for it to be effective in more complicated situations?)

For each party r, we assume that the corruption level is nonnegative,

$$c(r) \geq 0, \quad \forall r \in T. \tag{1}$$

Thus, a party has a zero corruption level iff its candidates all attain the ideal of perfect virtue and incorruptibility.

For simplicity, we assume that total cost of corruption by members of party r will be proportional to its corruption level c(r) and to the number of seats that party r wins in the election. The numerical value of each corruption level c(r) is interpreted as the dollar cost of corruption that each voter would have to bear if party r won all of the seats in the legislature. Thus, if $\sigma(r)$ denotes the fraction of the legislative seats that are won by candidates from party r, for each party r in T, then the total cost of corruption that each tax-paying voter must bear is

$$\Sigma_{r \in T}\ \sigma(r)c(r).$$

Equivalently, if there are S seats in the legislature and there are M voters, then we may say that c(r)M/S is the total amount of money that each elected legislator from party r would take from the treasury.

If there were no other policy questions at stake in the election, then almost any electoral system would be effective for minimizing the cost of corruption, because of the voters' unanimous preference for reducing this cost. However, the addition of any other policy question, about which the voters do not all agree, can lead to a situation in which some voters may have an incentive to support parties with positive corruption levels.

To keep things simple, we assume that there is just one major Yes-or-No policy question (say: "Shall our nation join a regional military alliance?") about which voters disagree and parties have taken different positions. Let T_Y denote the set of parties that are committed to voting "Yes" on this question in the legislature; we may call these the _affirmative parties_. Let T_N denote the set of parties that are committed to voting "No" on this question; we may call these the <u>negative parties</u>. We assume that all parties have publicly known positions on this question, so

$$T = T_Y \cup T_N, \quad \text{and} \quad T_Y \cap T_N = \emptyset. \tag{2}$$

If all the affirmative parties in T_Y had positive corruption levels, then it would not be surprising for these corrupt parties to get support from voters who prefer an affirmative policy outcome. To avoid such a trivial reason for supporting corrupt parties, we assume here that there is at least one affirmative party that has a zero corruption level, and there is at least one negative party that has a zero corruption level. That is

$$T_Y \cap \{r \mid c(r) = 0\} \neq \emptyset \quad \text{and} \quad T_N \cap \{r \mid c(r) = 0\} \neq \emptyset. \tag{3}$$

Thus, on each side of the policy question, there is at least one noncorrupt candidate available for every seat in the legislature.

Let J denote the set of voters, where J is a nonempty finite set. For each voter i in J, let v(i) denote the value to voter i of an affirmative policy decision by the government. This value v(i) may be positive or negative, and we normalize payoffs by supposing that the value of a negative policy decision would be zero to every voter. So $v(i) > 0$ means that voter i prefers an affirmative policy outcome, and $v(i) < 0$ means that voter i prefers a negative policy outcome. We may say that

$$\{i \in J \mid v(i) > 0\}$$

is the set of <u>affirmative voters</u>, and

$$\{i \in J \mid v(i) < 0\}$$

is the set of <u>negative voters</u>.

The actual government policy decision will be "No" unless a majority of legislators are elected from affirmative parties, in which case the government policy decision will be "Yes." Thus, if $\sigma(r)$ denotes the fraction of the legislative seats that are won by party r, for each r in T, then the <u>utility payoff</u> for each voter i will be

$$v(i) - \Sigma_{r\in T}\ \sigma(r)c(r) \quad \text{if} \quad \Sigma_{r\in T_Y}\ \sigma(r) > 1/2,$$

$$0 - \Sigma_{r\in T}\ \sigma(r)c(r) \qquad \text{if} \quad \Sigma_{r\in T_Y}\ \sigma(r) \le 1/2.$$

To avoid situations in which everyone votes frivolously because he anticipates that no race will be decidable by one vote, we assume that each individual in J has an independent probability ε of forgetting to vote or being <u>inactive</u> on election day, where

$$0 < \varepsilon < 1. \tag{4}$$

The results of the election can depend only on the votes that are cast by the active voters (that is, by those who remember to vote on the election day). Let us assume that, if everyone forgot to vote on election day (an event which has probability $\varepsilon^{|J|}$), then the election would be rescheduled for the next day, and each voter's behavior on the next day would be determined by another independent draw from the same distribution. The results of the scheduled election will be final, however, if at least one individual in J remembers to vote. So when an individual remembers to vote, he knows that there is a small positive probability ($\varepsilon^{|J|-1}$) that he may be the only active voter, in which case his vote alone would determine the entire allocation of legislative seats.

Finally, to allow the possibility that the voters may be divided into districts, we let D denote a partition of the set J into disjoint subsets, which we interpret as electoral districts. Thus, in our model, the elements of a <u>political situation</u> are summarized by the following parameters:

$$(T,\ (c(r))_{r\in T},\ T_Y,\ T_N,\ J,\ (v(i))_{i\in J},\ \varepsilon,\ D),$$

which must satisfy formulas (1)-(4) above.

To complete the definition of a voting game, it only remains to describe the rules of the electoral system. The electoral system must specify a set of feasible ballots for each voter, and a rule for determining the distribution of legislative seats $\sigma = (\sigma(r))_{r \in T}$ as a function of the ballots cast by the active voters. We assume that the set of possible ballots for each voter is a nonempty finite set, so that the existence of at least one Nash equilibrium (in randomized strategies) can be guaranteed for the resulting voting game. However, we must anticipate that the voting game may have multiple equilibria.

In the context of this model, the best property that an electoral system could have is to guarantee that, in all possible equilibria of all political situations, no corrupt party ever wins any legislative seats. Failing this property, we might ask that there should always exist at least some equilibrium such that no corrupt party wins any legislative seats (and we might then hope that voters would focus on playing according to such an equilibrium). If neither of these properties is satisfied, then there must exist political situations in which rational and intelligent behavior by voters is not compatible with the elimination of corrupt government officials.

Thus, our model enables us to divide electoral systems into three catagories. We say that an electoral system is <u>fully effective</u> iff, for every political situation as defined above, in every Nash equilibrium of the voting game that results from applying this electoral system, it can be guaranteed (with probability one) that all legislative seats will be won by parties with zero corruption levels. We say that an electoral system is <u>partly effective</u> iff, for every political situation as defined above, the voting game has at least one Nash equilibrium such that, with probability one, all legislative seats will be won by parties with zero corruption levels. We say that an electoral system is ineffective iff there exist some political situations in which, for every Nash equilibrium of the resulting voting game, there is a positive probability that a positive fraction of the legislative seats will be won by parties that have positive corruption levels.

3. Comparison of electoral systems

These catagories can be used to compare various electoral systems. We consider here plurality voting, approval voting, Borda voting, and proportional representation.

Plurality voting, approval voting, and Borda voting are examples of scoring rules that can be used with single-seat districts. Under each of these electoral systems, each voter must choose a vote vector that lists a number of points for each party. The vote total for a party in a district is the sum of points given to this party by all active voters in the district. Under plurality voting, each voter must choose a vote vector that gives 1 point to one party and gives 0 points to all other parties. Under approval voting, each voter must choose a vote vector that gives each party either 0 points or 1 point, but he can give 1 point to as many or as few parties as he wants. Under Borda voting, each voter must choose a vote vector that assigns 0 points to one party, 1 point to another party, 2 points to a third party, and so on, up to a maximum of $|T|-1$ points for some party.

Under each of these scoring rules, we assume that each district in D elects one seat in the legislature (that is, a $1/|D|$ portion of the legislature), which will be awarded to the party that gets the highest vote total in the district, provided that there is at least one active voter in the district. If there are no active voters in a district (that is, if everyone in the district has forgotten to vote on election day) then that district loses its legislative seat and the other districts increase their shares of the legislature, unless everyone has forgotten to vote in all districts, in which case the whole election will be rescheduled.

To complete the definition of these scoring rules, we must describe how the seat will be allocated in case of a tie. Here we will assume that, if there is a tie in a district, then a randomly selected active voter in the district will make the selection among the parties that are tied for having the highest vote total. We also assume that this voter's tie-breaking selection will be made before he gets any other information about the voting results. One way to implement this tie-breaking scheme is to ask each voter to submit a two-page ballot in which he specifies his vote vector on the first page and he specifies a rank-ordering of the parties on the second page. In the event of a tie, one of these second pages would be drawn at random, and the district's seat would be awarded to the party that is, according to the ranking on this page, highest among the parties that have the maximal vote total in the district. (We do not need to require any consistency between an individual's vote vector and his secondary ranking. That is, the party that is highest in the second-page ranking does not need to get the most points in the vote vector on the first page of the ballot.)

A proportional representation system is defined here in the pure sense, with fractionally divisible seats. Under proportional representation, each voter must choose a vote vector that gives 1 point to one party and gives 0 points to all other parties, and each party r in T will get a fraction of the legislative seats $\sigma(r)$ that is equal to the fraction of all active voters who give their votes to party r.

We can now state the main result of this paper.

Theorem. Proportional representation and approval voting are fully effective electoral systems. Plurality voting is partly effective but is not fully effective. Borda voting is ineffective.

Proof. We verify first that plurality voting is partly effective. Let r and s be parties such that $r \in T_Y$, $s \in T_N$, and $c(r) = c(s) = 0$. There is a Nash equilibrium in which all affirmative voters give their votes to party r, and all negative voters give their votes to party s, so that only the noncorrupt parties r and s win seats in the legislature.

We now show that plurality voting is not fully effective. To prove this claim, it suffices to consider one example that has an equilibrium in which corrupt parties win seats, even though noncorrupt parties are available on both sides of the policy question. So let us consider a simple example in which there is one district (so $D = \{J\}$), and there are four parties, numbered 1 to 4. Suppose that $T_Y = \{1,3\}$, $T_N = \{2,4\}$, $c(1) = c(2) = 100$, and $c(3) = c(4) = 0$. That is, parties 1 and 3 are affirmative parties, parties 2 and 4 are negative parties, parties 1 and 2 have the same positive corruption level, but parties 3 and 4 are noncorrupt. Suppose that there are 5 affirmative voters, each with $v(i) = 1$, and 5 negative voters, each with $v(i) = -1$, and let $\varepsilon = 1/1000$. Now consider a scenario in which all affirmative voters plan to vote for party 1 and all negative voters plan to vote for party 2. A voter who unilaterally deviated from this scenario would change the policy outcome to his own disadvantage with probability almost 1/2 because, when everyone votes (which has probability greater than .99), his deviation would break the tie by reducing the vote total of the party that he favors among parties 1 and 2. On the other hand, his unilateral deviation could give the seat to a less corrupt party (party 3 or party 4) only if at least 7 other voters forget to vote, which has probability less than 10^{-19}. So this scenario, in which a corrupt party is sure to win the seat, is an equilibrium (because 100×10^{-19} is less than $1 \times 1/2$). In fact, this equilibrium is perfect, and it does not involve the use of any weakly dominated strategies. Furthermore, if parties 1 and 2 are old established parties, but parties 3 and 4 are new entrants to the political arena, then the weight of tradition could make this equilibrium focal, even though it is Pareto-inferior to the equilibrium in which all voters switch to the new parties.

Next we show that Borda voting is ineffective. It suffices to show one example of a political situation in which all equilibria assign seats to corrupt parties with positive probability. Consider a situation in which there is one district, and $J = \{1,2\}$. Voter 1 is a negative voter with $v(1) = -1$, and voter 2 is an affirmative voter with $v(2) = 1$. Now suppose that there are three parties, numbered 1, 2, 3. Party 1 is a negative party, whereas parties 2 and 3 are affirmative parties. The party corruption levels are

$$c(1) = c(2) = 0, \quad \text{and} \quad c(3) \geq 0.$$

Suppose first that $c(3)$ is equal to zero. Then the voting game is a two-person constant-sum game. In any equilibrium of this game, voter 1 randomizes between submitting the vote vectors (2,1,0) and (2,0,1), each with probability 1/2, and voter 2 randomizes between submitting the vote vectors (0,1,2) and (0,2,1), each with probability 1/2. (There are multiple equilibria of this game, but only because there is indeterminacy about how each voter will rank parties 2 and 3 on his secondary tie-breaking page.) Thus, in all equilibria with $c(3) = 0$, there is a probability 1/4 that the vote totals will be (2,1,3), in which event party 3 will win the seat (regardless of the tie-breaking pages).

The set of Nash equilibria depends upper-hemicontinuously on the parameters of the game. So for all sufficiently small positive values of $c(3)$, we still get a game in which, under every Nash equilibrium, the probability of party 3 being the unique top-scoring party (and thus winning the seat) is close to 1/4. (In fact, it can be shown that this probability is actually greater than 1/4 and is increasing in $c(3)$, when $c(3)$ is between 0 and 1/2.)

Next, we show that <u>proportional representation is fully effective</u>. When there are m active voters, each voter allocates $1/m$ of the legislature with his vote. By changing his vote from a party with positive corruption level to a noncorrupt party that agrees with him on the policy question, a voter will decrease his expected corruption cost and will not decrease the probability of getting the policy decision that he favors. Thus, for each voter, the only strategies that are not strongly dominated are the strategies in which he votes for a party that has zero corruption level and agrees with him on the policy question. No Nash equilibrium can involve strategies that are strongly dominated, so the corrupt parties get no votes in equilibrium.

Finally, we show that <u>approval voting is fully effective</u>. Consider any equilibrium of the voting game under approval voting. (The equilibrium may be in either pure or randomized strategies. For a randomized equilibrium, phrases below like "voters who give approval votes to party r" should be interpreted as applying after the voters have carried out their randomizations and determined their actual ballots.) Fix any given district in D. Let p denote the probability that this district's representative can determine whether the government's policy is affirmative or negative. (That is, p is the probability that, from all other districts, the number of seats won by affirmative parties is either equal to or one more than the number of seats won by negative parties.) For any voter i and any party r, let $u_i(r) = pv(i) - \delta c(r)$ if $r \in T_Y$, and let $u_i(r) = -\delta c(r)$ if $r \in T_N$, where δ denotes the district's expected share of the legislature (so δ is close to $1/|D|$ when ε is small). Then, $u_i(r)$ denotes the net change in voter i's utility from giving this district's seat to party r, relative to the alternative of giving this district's seat to a noncorrupt negative party. Voter i's true preference ranking of the possible winners of his district's seat will therefore be in order of these $u_i(r)$ numbers. In any weakly undominated voting strategy, a voter i in this district should list the parties in his secondary ranking in order of these $u_i(r)$ numbers, because his secondary ranking only matters when it is alone responsible for choosing among tied parties.

Adding approval votes for noncorrupt affirmative parties can never hurt an affirmative voter, because adding such votes can only change the outcome by transferring the seat to a noncorrupt affirmative party, which is the best possible outcome in this district for an affirmative voter. Thus, any ballot that an affirmative voter may submit is weakly dominated for him by another ballot that differs only in adding approval votes for all noncorrupt affirmative parties.

So let us say that to <u>rectify</u> a ballot for an affirmative voter i means to transform the ballot by adding approval votes for all noncorrupt affirmative parties (if any were not approved in the given ballot), and by ranking the parties on the secondary tie-breaking page according to the $u_i(r)$ numbers (with r above s if $u_i(r) > u_i(s)$). Similarly, let us say that to rectify a ballot for any negative voter j means to transform the ballot by adding approval votes for all noncorrupt negative parties, and by ranking the parties on the secondary tie-breaking page according to the $u_j(r)$ numbers. Then rectifying a ballot for any voter i transforms it into a ballot that weakly dominates the original ballot as a strategy for a voter i. Thus, in any Nash equilibrium, there cannot be any positive-probability event in which rectifying a voter's ballot for him would strictly improve the outcome for him (because otherwise the voter's expected utility could be strictly increased by rectifying his ballot).

We claim now that, for any party r, there must exist a noncorrupt affirmative party s such that, in equilibrium, all affirmative voters who give approval votes to party r also give approval votes to party s. This claim is trivially true if party r is a noncorrupt affirmative party, so suppose that party r has a positive corruption level or is a negative party. If this claim fails to be true, then we can find a set G with the properties that G is a nonempty set of affirmative voters who all give approval votes to party r, and there is no noncorrupt affirmative party that also gets approval votes from all the voters in G, and G is a minimal set with these properties. In the event that G is the set of active voters in the district, the winning party in this district must get approval votes unanimously from all active voters (because party r gets such unanimous approval), and so the winning party cannot be a noncorrupt affirmative party. But then, in the positive-probability event that G is the set of active voters and voter i's secondary ranking is used to break any ties, voter i could cause the district's seat to be transferred to a noncorrupt affirmative party by rectifying his ballot. (We use the minimality of G here to guarantee that, for any voter i in G, there would be a unanimously approved noncorrupt affirmative party if i added approval votes for all such parties.) Thus, for any of the affirmative voters in G, there is a positive-probability event in which rectifying the ballot would strictly improve the outcome. This result contradicts the conclusion of the preceding paragraph, and this contradiction proves that no such set G can exist. That is, there must exist a noncorrupt affirmative party s that gets approval votes from all the affirmative voters who vote for party r.

Let d denote the highest corruption level of any party that ever wins the legislative seat in this district, in the given Nash equilibrium under approval voting. Let H be a minimal set of voters such that a party with corruption level d can win the seat with positive probability when H is the set of active voters in the district. Henceforth in this proof, let r be a party with corruption level d that can win when H is the set of active voters. Without loss of generality, let us suppose that r is an affirmative party. (The negative case can be handled symmetrically.) Contrary to our theorem, suppose that d, the corruption level of party r, is strictly positive.

If j is a negative voter in H who is giving an approval vote to party r in this equilibrium, then let j consider dropping his approval vote for r and for all parties that are at least as corrupt as r (if he is giving votes to any such corrupt parties). Such a change by j could never make any more corrupt party win the seat, because no party more corrupt than r ever wins the seat when j is inactive. Furthermore, when H is the set of active voters, such a change would transfer the seat to a party that is strictly less corrupt than r (because otherwise such a corrupt party could win when j is removed from the active set, which contradicts the definition of H). The negative voter j would strictly prefer such a transfer (because, if the less corrupt party differed in policy from r, then it would be a negative party which j prefers). So voter j would expect to gain by dropping his approval votes for r and all more corrupt parties. Thus, no negative voters in H give approval votes to party r in equilibrium. Party r can only get votes from affirmative voters in H.

We have shown that there exists some noncorrupt affirmative party s that also gets approval votes from all the affirmative voters who vote for party r. So this noncorrupt affirmative party s must be tied for the maximal number of approval votes when H is the set of active voters, because s gets get at least as many approval votes in H as the winning party r. Thus, when H is the set of active voters, any voter in H who would have selected party r in the secondary tie-breaker could improve the outcome by rectifying his ballot (and thus moving the noncorrupt affirmative party s above the corrupt affirmative party r). This result contradicts the fact that, in equilibrium, rectifying a voter's ballot cannot strictly improve the outcome for a voter in any positive-probability event. But this contradiction was derived from a supposition that r had a strictly positive corruption level. Thus, $c(r) = d = 0$, and so no positively corrupt party can ever win the district's seat in equilibrium under approval voting. Q.E.D.

This proof relies on the assumption that ties are broken essentially by letting one randomly selected voter cast a second vote. Other ways of breaking ties could be considered, and some of these may lead to similar results. One method which would not work, however, is to simply select one of the tied parties at random. Under approval voting, for example, suppose that half of the electorate consists of affirmative voters who give approval votes to two affirmative parties, and the other half consists of negative voters who give approval votes to one negative party. Then the probability of an affirmative party getting the seat would be 1/2 if a randomly-selected voter breaks the tie, because half the voters would choose the negative party; but it would be 2/3 if a tied party is selected at random, because two of the three tied parties are affirmative. Thus, under the random-party tie-breaking scheme, the affirmative voters could have an incentive to support a second affirmative party, even if it were slightly corrupt.

4. Conclusions

To analyze the effect of elections on political corruption, a more general model would also include party leaders (or candidates) as players, each of whom would choose his party's corruption level endogenously. In analogy to the Bertrand model of oligopolistic price competition, we might suppose that each party leader's objective is to maximize his party's expected corrupt profit, which is his party's expected share of legislature multiplied by his party's corruption level. The voters would vote in the election after getting public signals that depend on the corruption levels that the party leaders have chosen. In such a game, two conditions would be necessary for electoral competition to deter corruption: the voters' signals should enable them to identify corrupt parties with reasonably high accuracy; and, in the subgame after these signals have been received, the voters' equilibrium behavior should lead to the electoral defeat of parties which have been identified as more corrupt. The focus in this paper has been on the second of these two conditions, because it is the condition that depends on the electoral system. To learn when this condition may be satisfied while keeping our model as simple as possible, we have studied here only the subgame that begins after the politicians have chosen their corruption levels and the voters have received signals that reliably predict each party's corruption level. Our results have direct implications for more general game models, however. If an electoral system is ineffective in this subgame then more general models cannot have subgame-perfect equilibria in which all parties choose to not be corrupt, because a party could get positive expected profits by deviating to a positive corruption level.

In the proof of our theorem, plurality voting fails to be fully effective because it can create a strong incentive for voters on each side of the policy question to coordinate and vote for the same party (see also Myerson and Weber 1991). In an equilibrium under plurality voting, if an affirmative voter switched his vote from a corrupt affirmative party to a noncorrupt affirmative party that no one else is supporting, then his switch may be more likely to change the government policy decision from Yes to No (by transferring the district's seat from a corrupt affirmative party to a corrupt negative party) than to reduce the level of government corruption (by transferring the district's seat to the noncorrupt party). Thus, under plurality voting, the incentive to coordinate with other like-minded voters may be more important, in each individual's voting decision, than the incentive to oppose corruption.

On the other hand, Borda voting is ineffective because it may disproportionately favor a block of like-minded voters who can divide their support among many parties that advocate their preferred policy. So under Borda voting, the affirmative voters may want to divide their support among several affirmative parties, even if some of these parties are corrupt.

Thus, we find two opposite reasons why an electoral system may fail to be fully effective. An electoral system like plurality voting may give too much incentive for a block of voters to coordinate and concentrate their support on one party. Such coordination incentives create a bandwagon effect (in the sense of Simon 1954) which becomes a barrier to entry against noncorrupt new parties. On the other hand, an electoral system like Borda voting may give too much incentive for a block of voters to spread their support over many parties, rather than consolidating their support behind one party. Such divisiveness can create an underdog effect (in the sense of Simon 1954) which protects the vote share of corrupt parties.

Approval voting is fully effective because it avoids both the strong coordination incentives of plurality voting and the divisiveness of Borda voting. Under approval voting, each voter can freely give full support to any noncorrupt party that advocates his preferred policy, without worrying about whether others are voting for this party, because giving such support would not prevent him from also joining like-minded voters in support of other parties. But the nondivisiveness of approval voting means that, in equilibrium, a voter would have no incentive to support a corrupt party that advocates his preferred policy, once he realizes that all other voters who would consider supporting this corrupt party are (also) giving their support to the noncorrupt parties that advocate the same policy.

Similarly, proportional representation minimizes the incentives for a block of voters to coordinate, but it does not create any positive incentives for them to divide their support among several parties. So under proportional representation, each voter should always vote for one of his most-preferred parties, which (when there is only one Yes/No policy question to be decided in the legislature) will be a noncorrupt party that advocates his preferred policy.

These results suggest that we should search for other electoral systems that are similar to approval voting and proportional representation in these two properties. Myerson (1991) has formalized an axiom of low incentives to coordinate or coalitional-straightforwardness, and an axiom of nondivisiveness, for general legislative electoral systems. The main result of Myerson (1991) is that approval voting and proportional representation systems are the only electoral systems that satisfy these two axioms together with three other natural properties (responsiveness, neutrality, and homogeneity). However, these axioms do not correspond exactly to the concept of full effectiveness as defined in this paper, and other good electoral systems may also be fully effective. (In particular, rules like single transferable vote and the Simpson-Kramer minmax rule should be studied in future research.)

Of course, the model considered here is a highly simplified abstraction. In particular, we have ignored problems of adverse selection and moral hazard, by simply assuming that every politician's level of corruption is given as a publicly known function of his party affiliation. Banks and Sundaram (1991) have studied a dynamic model which confronts the problem of monitoring candidates whose types are unknown and whose corruption (or effort) levels are endogenously chosen after winning elected office. However, it is more difficult to completely characterize the set of all equilibria of such a model. Banks and Sundaram simplify their analysis by considering an electorate that consists of only one voter, but this assumption precludes any comparison of different electoral systems.

More generally, the electoral system is only one of many structural factors that can affect the level of political corruption in a country. Independence of the judiciary, freedom of the press, separation of powers between national and local governments, competitive nomination procedures in parties, campaign financing systems, and voters' political consciousness can also be significant determinants of the corruption level. When we use a simple model to argue that the rules of the electoral system can affect the amount of political corruption, we do not deny the importance of such other factors. The analysis in this paper is intended to show that, when other factors are held constant, electoral systems may differ in their effectiveness for inhibiting corruption. This difference may be important to consider in the comparison of electoral systems and the constitutional design of democratic institutions.

REFERENCES

Banks, J. S., and Sundaram, R. K. (1991). "Adverse Selection and Moral Hazard in a Repeated Elections Game," University of Rochester discussion paper.

Myerson, R. B. (1991). "Proportional Representation, Approval Voting, and Coalitionally Straightforward Elections," Northwestern University discussion paper. To appear in Political Economy: Institutions, Information, Competition and Representation, edited by N. Schofield, Cambridge: Cambridge University Press, forthcoming.

Myerson, R. B., and Weber, R. J. (1991). "A Theory of Voting Equilibria," Northwestern University discussion paper.

Riker, W. H. (1982). Liberalism Against Populism. San Francisco: Freeman.

Simon, H. A. (1954). "Bandwagon and Underdog Effects and the Possibility of Equilibrium Predictions," Public Opinion Quarterly 18, 245-253.