

FEDERALISM AND INCENTIVES FOR SUCCESS OF DEMOCRACY

by Roger B. Myerson

March 2004

<http://home.uchicago.edu/~rmyerson/research/federal.pdf>

Abstract: Success or failure of democracy are interpreted as different equilibria of a dynamic political game. Democracy is frustrated and fails when voters do not replace corrupt leaders, given some transition cost of changing leadership, and any new leader would be likely to govern corruptly. Even when new politicians have a small probability of being virtuous, democracy can be frustrated in a unitary political system. But in a federal system, democracy cannot be consistently frustrated at both national and provincial levels, because provincial leaders who govern responsibly could build reputations to become contenders for higher national office. Similarly, democracy cannot be consistently frustrated in a democratization process that begins with decentralized provincial democracy and only later introduces nationally elected leadership.

Address: Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637. Email: myerson@uchicago.edu. URL: <http://home.uchicago.edu/~rmyerson/>

FEDERALISM AND INCENTIVES FOR SUCCESS OF DEMOCRACY

by Roger B. Myerson

"Countries in transition that have aimed for national elections as a first step (Bosnia for example) have bogged down and generally handed power over to avatars of the old regime. By contrast, Kosovo and East Timor began with local elections, with a far better result of bringing forward new talents and capabilities, and giving people a sense of empowerment."

Final Report on the Transition to Democracy in Iraq (Nov, 2002), page 24.

1. Introduction

In this paper we analyze some simple game models to develop new insights into the difficulties of establishing democracy in a nation where it has not existed before. In such a nation, past experience of authoritarian rule may have led citizens to expect very little from their political leaders. Before the transition from authoritarian rule, successful politicians had to develop a reputation for using their power to serve their superiors and to reward their supporters in the political hierarchy. So the absence of previous democracy may generally imply that there are no politicians who have developed reputations for responsibly using political power to serve the population at large. In such a situation, voters may naturally believe that the winner of the first election for control of the national government will use his power mainly to benefit himself and his active supporters. Worse, they may also naturally believe that any other politician would behave the same way if he could take the place of the incumbent leader. With such low expectations, voters are unlikely to support democratic challengers to the ruling party, and they

might see little reason to protest if the government suppressed its political opposition. This unfortunate equilibrium of low expectations is the central concern of this paper.

So our central question is, how can the chances of success for a new democracy be improved? It has been argued (at least since Lipset, 1959) that economic and social conditions may affect the chances of democracy. But economic and social conditions have little to do with the trap of low expectations described above. Furthermore, basic economic and social realities are hard to change by democratization policies in the short run. But the constitutional structures of political institutions are policy variables that are open to change in a democratization process, by the very definition of democratization. Our belief in democracy is based on an understanding that the structure of political institutions can have important consequences for the behavior of political leaders and the welfare-relevant performance of government. Different constitutional structures may change the political game in ways that can mitigate or eliminate this equilibrium of low expectations. So efforts to establish democracy where it has not existed before need to be informed by careful study of the ways that constitutional structures can affect the chances of success for new democracy.

A central focus of empirical research in this direction has been on the comparison of presidential and parliamentary structures. Stepan and Skach (1993) report evidence that chances for survival of democracy may be better with parliamentary institutions than presidential (see also Linz and Valenzuela, 1994). But Boix (2003, pp. 155-169) actually finds stronger evidence for federalism improving chances of democracy. This paper offers a theoretical perspective on the importance of federalism in the success of new democracy.

Most theoretical analyses of federalism begin with an assumption that the principal

consequence of federalism is that it limits the decisiveness of national majorities and allows some decisions about public goods and redistributive policies to be made by provincial electorates (e.g.: Persson and Tabellini, 1996; Dixit and Londregan, 1998; Oates, 1999; Boix, 2003). From this perspective, Boix argues that federalism should improve the chances for democratic success because provincial electorates generally have smaller internal differences of tastes and endowments than the national electorate, and thus reduces the stakes in redistributive politics. But Lipset, page 99, actually suggests the opposite argument, that federalism should improve the chances for democracy when provincial divisions cut across other economic and social divisions, so that the diversity in each province would be similar to the diversity in the nation.

In this paper, we focus on a different consequence of federalism: its effect on the structure of incentives for ambitious politicians. Under federalism, independent provincial leaders can hope to become candidates for national leadership by building a reputation for responsible democratic leadership at the provincial level. Thus, even when voters have no significant differences in policy preferences, federal separation of power can fundamentally change the nature of competition among politicians. Our shift of focus here can be described in terms of Riker's (1982) arguments for liberalism against populism in democratic theory: Instead of emphasizing the consequences of democracy for populist redistribution among citizens, here we evaluate the performance of democracy by its liberal effectiveness against corrupt abuse of power by political leaders.

The concept of democratic failure here that we apply here may also be somewhat novel, because we do not equate failure of democracy with breakdown of democracy or with violent use of force. Other theoretical models have focused on important questions about when political

factions might rationally decide to threaten or actually use violent force against their rivals (Ellman and Wantchekon, 2000; Acemoglu and Robinson, 2000; Boix 2003). But here we consider failures of democracy that occur within the democratic system itself. Even when democratic elections are held with no political violence, we may say that democracy has failed the voters if elected leaders corruptly abuse their power and do not provide voters with any benefits of government beyond what they could have expected from an authoritarian regime.

We can see how such failures could occur within a democratic system. Suppose that voters perceive costs and risks of changing leaders in government. If voters expect that any future leader would govern corruptly, then voters could rationally re-elect corrupt leaders, thus giving leaders no incentive to eschew corruption. In such an equilibrium, voters would see no benefit from democracy. A subsequent termination of such dysfunctional democratic formalities might be a natural consequence, but we could appropriately say that the real failure occurred when the democratic system failed to deter politicians' corrupt abuse of power.

I have argued elsewhere that the effectiveness of democratic competition against corruption may depend critically on the details of the electoral system even when voters differ merely in their preferences about a simple public decision between two alternatives (Myerson 1993). So to keep the focus on federalism and avoid complicated questions about electoral systems, our simplifying assumption here is that voters have no differences at all about any public policy decisions. All voters simply prefer clean responsible government rather than corrupt government. The fundamental problem is that political leaders would generally prefer to govern corruptly. Candidates for high office cannot commit themselves to avoid the temptations of corruption after election unless they have a valuable reputation that would be lost by acting

corruptly.

So success of democracy requires a self-sustaining relationship between politicians and voters in which politicians cultivate reputations for good government and voters reject politicians who lack such a reputation. Section 2 here shows that, in a unitary democracy, the existence of reputations for a successful democracy may depend on which of several equilibria is expected by voters and politicians in a dynamic political game that has many equilibria. So with a unitary structure, the success or failure of democracy in different nations may depend on how their different national traditions have focused people's expectations on different equilibria, according to Schelling's (1960) focal-point effect.

The main results of this paper show how federal separation of power can change the dynamic political game in ways that eliminate the worst equilibria, in which politicians never use power responsibly and voters see no benefits of democracy. In Section 3 we consider a simple game model of federal democracy, in which voters elect both a national leader and local leaders for the various provinces, and we show that consistent frustration of democracy cannot occur in an equilibrium of a federal system. The basic idea here is that federal separation of powers changes the structure of democratic competition by allowing local politicians to cultivate reputations for good democratic government, when they serve as elected leaders of provincial government. Leaders of federal provinces have real power, and their hopes of rising to national leadership can give them a strong incentive to use this power responsibly. So from our perspective, the key attribute of federalism is that it creates a class of elected leaders who can exercise independent political power while aspiring to higher office. (For a related model of career incentives in a hierarchy of offices, see Harrington, 2000.) This advantage would vanish if

the provinces separated to become independent nations, because the leader of each former province would then have no hope of advancement.

Section 4 here considers a kind of sequential federalism in a transition to democracy, where elected provincial leaders hold all power initially in an interim provisional government, but later a unitary national democracy will be established. We show that the decentralized democratic structure during this interim phase can also strongly motivate provincial leaders to begin cultivating reputations for good government, so that they can be contenders for national power when direct national democracy is established. It is worth noting that the American transition to democracy had a just such a decentralized interim phase in 1777-1788 under the Articles of Confederation. Such a decentralized interim structure was also recommended for introducing democracy in American-occupied Iraq (Democratic Principles Work Group, 2002), but it has not been implemented there.

Section 5 concludes with a general discussion of our results.

2. A model of success and failure of a unitary democracy

Let us begin our analysis by considering a simple unitary democracy with an elected leader who serves a fixed term of one period and then must run for re-election again in each period until he is rejected by the voters. In each period, the leader must choose whether to serve responsibly or corruptly. (None of the results in this paper would change if we assumed that, once a leader acts corruptly, he must always be corrupt thereafter.) Let b denote the leader's benefit each period when he serves responsibly. Let c denote the leader's additional benefit from being corrupt instead of responsible in each period, so that a corrupt leader gets $b+c$ each period.

A politician out of office gets payoff 0. Each politician wants to maximize the expected discounted value of his payoffs, where payoffs in future periods are discounted by some discount factor ρ per period. (In our game model, each "politician" could also be interpreted as a political faction that acts in unity as a team.)

In any period, each voter gets welfare w from the government if its leader acts responsibly, but each voter gets 0 from the government if its leader acts corruptly, except that a voter's payoff in either case is reduced by a transition cost x in a new leader's first period in office. Voters also discount their future benefits by the discount factor ρ per period. Since voters all get the same payoffs, we assume that elections are determined by the preferred choice of any voter. These parameters (b, c, w, x, ρ) are assumed to all be positive numbers with $\rho < 1$.

In an equilibrium of this game, we say that democracy succeeds if every new leader is always expected to serve responsibly. It is not difficult to construct an equilibrium in which democracy succeeds. In this scenario, each new leader is expected to govern responsibly, but if he ever deviated to corruption then he would be expected to continue corruptly until replaced; and voters are expected to re-elect a leader who has always been responsible but to replace a leader who has ever been corrupt. A responsible leader's expected total benefit to a voter is

$$w(1 + \rho + \rho^2 + \dots) = w/(1-\rho).$$

Thus, for such success of democracy to be an equilibrium, we only require

$$w/(1-\rho) \geq x \text{ and } b/(1-\rho) \geq b+c,$$

so that voters want to replace corrupt leaders, and leaders prefer a long responsible career over a short corrupt career.

In an equilibrium, we may say democracy fails if new leaders are always expected to

serve corruptly. We may also say that democracy is frustrated if it does not succeed and voters would never replace a corrupt leader. It is straightforward to construct a bad equilibrium where democracy is frustrated and fails, with any positive (w,x,b,c,ρ) . With a positive transition cost x , voters do not want to reject a corrupt incumbent if his replacement would be equally corrupt, and no leader has any incentive to forfeit the positive benefits of corruption c if he will be re-elected in any case.

But now let us perturb and extend our model by allowing that there is a small probability $\epsilon > 0$ that any randomly sampled politician may be intrinsically virtuous, and thus can only serve responsibly. This possibility of virtue may be attributed to an ideological education in which future politicians are taught that they should serve the people responsibly. In the tradition of economic analysis, we are not accepting Plato's assumption that education can reliably give leaders ideal social preferences, but we are allowing that such ideological education may change a young politician's actual preferences with some very small probability. A politician who is not intrinsically virtuous may be called normal. That is, normal politicians maximize their expected payoff as described above, and voters understand that any politician has probability $1 - \epsilon$ of being normal.

With this perturbation, our definition of failure of democracy must be modified. We say democracy fails if any normal leader is always expected to serve corruptly. That is, democracy fails if the only type of leader who would serve responsibly is the rare virtuous type, who would have been just as benevolent if he were a authoritarian monarch.

A bad equilibrium where democracy is frustrated and fails still exists if

$$\epsilon \leq x(1 - \rho)/w,$$

because then voters' cost x of rejecting an incumbent is greater than their expected gain $\varepsilon w/(1-\rho)$ from better government in the unlikely ε -probability event of getting a new leader who is virtuous. Thus, if the probability of intrinsic virtue is small enough, then the problem of democratic failure remains.

We may summarize our results for the unitary government as follows.

Theorem 1. Suppose $\varepsilon < x(1-\rho)/w < 1$ and $b+c < b/(1-\rho)$. Then there exists an equilibrium where unitary democracy succeeds (leaders act responsibly), but there also exists an equilibrium where unitary democracy is frustrated and fails (corrupt leaders are re-elected, each leader serves corruptly unless he is intrinsically virtuous).

In Theorem 1, the formula $x(1-\rho)/w$ computes the lowest probability of a new leader serving responsibly such that national voters would be willing to replace a corrupt leader in a scenario where all leaders are expected to be either always-responsible or always-corrupt. The first inequality in Theorem 1 says the probability ε of virtuous types is too small for voters to want to replace a corrupt leader if only intrinsically virtuous types would serve responsibly. The second inequality says that voters would want to replace a corrupt leader if any new leader would be sure to serve responsibly forever. The third inequality says that a politician would prefer an unbounded term of responsible service over one period of corrupt service.

Under the conditions of Theorem 1, there can also exist other equilibria between the extremes of complete failure and success. For example, consider a randomized equilibrium where any new leader is randomly responsible with probability $\sigma = x(1-\rho)/w$, and then is always expected repeat his previous choice, and voters keep a responsible leader but randomly

reject a corrupt leader with probability $\tau = (1-\rho)c/(b\rho)$. To make σ the overall probability of responsible leadership, a normal leader must be responsible with probability $\hat{\sigma} = (\sigma-\varepsilon)/(1-\varepsilon)$, so that $\sigma = (1-\varepsilon)\hat{\sigma}+\varepsilon$. The conditions in Theorem 1 imply that $\hat{\sigma}$ and τ are between 0 and 1. These probabilities form an equilibrium because τ satisfies the equation

$$b/(1-\rho) = (b+c)/(1-(1-\tau)\rho),$$

which implies that a new leader is willing to randomize between responsibility and corruption, and σ satisfies the equation

$$\sigma w/(1-\rho) - x = 0,$$

which implies that voters are willing to randomize between rejecting and re-electing a corrupt incumbent leader.

This randomized equilibrium and the previously discussed pure-strategy equilibria have the property that there is a constant probability of voters rejecting a leader immediately after any period in which he has acted corruptly. (Other equilibria can be constructed with and without this property.) With this property of stationary reactions to corruption, if democracy is not frustrated then, with probability one, the voters will eventually get a leader who always serves responsibly. (See the proof of Theorem 3.)

So unitary democracy, as a dynamic political game, can have multiple equilibria. Thus the difference between success or failure of democracy may depend only on the beliefs that voters and politicians have about each others' future behavior. (For other game-theoretic formulations of this important idea, see Weingast, 1995, 1997.)

3. Federal democracy with national and provincial governments

Now let us consider a federal democracy where each of N provinces has an elected leader called a governor, and there is also an elected leader of the nation called the president. (No distinction between presidential or parliamentary system is intended here, and we could have equally well called the national leader a prime minister, provided that his selection is understood as the primary function of the parliament.) The sequence of decisions in each period is as follows. At the beginning of each period, national voters first choose a president, and then provincial voters choose a governor in each province. Then each leader chooses whether to serve responsibly or corruptly, unless he is intrinsically virtuous, in which case he must serve responsibly.

As before, each politician has probability ε of being the virtuous type, otherwise he is normal. A president gets payoff b_1 or b_1+c_1 each period, depending on whether he is responsible or corrupt. A national voter gets payoff w_1 when president serves responsibly, but gets 0 from corruption, except that his payoff is reduced by transition cost x_1 in a new president's first term. Similarly, a governor gets payoff b_0 or b_0+c_0 each period, depending on whether responsible or corrupt, where we may assume that b_0 and c_0 are much less than b_1 and c_1 . A provincial voter gets w_0 when governor serves responsibly, but gets 0 from corruption, except that payoff is reduced by cost x_0 in a new governor's first term. At any period, voters and politicians discount future benefits by the factor ρ .

For simplicity, let us assume that, when a governor is elected president, the voters in his province get local benefits that are just equal to their cost x_0 of replacing him, so they cannot lose by their governor's good reputation. We also assume elections at each level are determined by

voters' expected payoffs from this level of government alone, ignoring any effects from the other level of government. That is, national elections are not influenced by the local effects on the voters in one province when its governor becomes president; and voters in provincial elections are not influenced by the national benefits of searching for better presidential candidates. (The latter effect could add a small incentive for replacing a corrupt governor, and so would not affect our positive results below.) So at each level, all voters have the same effective preferences, and we can assume that the winner of each election is chosen according to the preferences of any voter.

In describing a federal equilibrium, we say democracy at either level (national or provincial) succeeds if every new leader at this level is always expected to serve responsibly. We say that democracy at this level fails if normal leaders at this level are always expected to serve corruptly. And we say that democracy at this level is frustrated if it does not succeed at this level and voters at this level would never reject a leader's re-election bid.

To make the analysis interesting, let us maintain as a basic assumption that the parameters at each level are such that, if this level were a unitary democracy, then it would have multiple strict equilibria where democracy can succeed or fail. That is,

$$(1) \quad \varepsilon < x_0(1-\rho)/w_0 < 1, \quad b_0+c_0 < b_0/(1-\rho), \quad \varepsilon < x_1(1-\rho)/w_1 < 1, \quad b_1+c_1 < b_1/(1-\rho).$$

Also, we assume that a politician would always prefer being president over being governor, no matter how he would serve in either role:

$$(2) \quad b_1 > b_0 + c_0.$$

Our analysis of federal equilibria begins with two straightforward observations about the implications of frustration and failure of democracy at the national level.

If democracy is expected to fail at the national level, then national voters would prefer to a candidate who has the highest probability of being virtuous, based on his record, because only a virtuous president would serve responsibly.

If democracy is frustrated at the national level in an equilibrium, then it also must fail at the national level, because a normal president has no incentive to serve responsibly if national voters would always re-elect him even when he acts corruptly.

As in the unitary case, it is easy to construct multiple equilibria of this federal democracy.

There exists an equilibrium where provincial democracy succeeds but national democracy fails. In this equilibrium, corrupt governors would not be re-elected, and so all governors act responsibly. But national voters understand that any governor (or any other politician) would become corrupt with probability $1 - \epsilon$ after election to the presidency, and so a corrupt president is re-elected.

There also exists an equilibrium where provincial democracy fails but national democracy succeeds. In this equilibrium, a rare governor who serves responsibly can be identified as intrinsically virtuous, but that does not make him more attractive to national voters, because they are confident any that president would serve responsibly so as to win subsequent re-election. Thus, a good reputation would not increase a governor's probability of winning higher national office, and so governors have no incentive to serve responsibly in this equilibrium.

But notice that these two mixed equilibria require voters to have inconsistent expectations about the functioning of democracy at different levels. Success (or failure) of democracy at one level may lead voters to focus on the possibility of getting similar performance from government at the other level as well. The inconsistency across levels of the above equilibria may make them

less likely to be the focal equilibrium that people expect in such a federal democracy.

There exists an equilibrium where democracy consistently succeeds at both the provincial and national levels. In this equilibrium, the president and the governors always act responsibly, because they would not be re-elected otherwise. But our main result is that democracy cannot consistently be frustrated and fail at both levels in a federal system.

Theorem 2. In any federal equilibrium, if national democracy is frustrated and fails, then provincial democracy cannot fail or be frustrated. With frustration of national democracy, when the president has not served responsibly, any new governor's probability of serving responsibly must be at least $\epsilon w_1 / ((1-\rho)x_1)$.

Proof of Theorem 2 When national democracy fails, only a virtuous president will serve responsibly. Now consider the game at the first period, or at any other point in the game when the incumbent president is not virtuous and some province has elected a new governor. Let q denote the smallest probability that the new governor will always serve responsibly. So this governor's responsible service would make voters think that his probability of being virtuous is ϵ/q . Then national voters' expected gain from replacing the corrupt president by this governor would be

$$(\epsilon/q)w_1/(1-\rho) - x_1.$$

Frustration of national democracy means that national voters would never choose to replace corrupt president. Thus, $(\epsilon/q)w_1/(1-\rho) - x_1 \leq 0$, and so

$$\epsilon w_1 / ((1-\rho)x_1) \leq q.$$

With $w_1/(1-\rho) > x_1$, we get $q > \epsilon$. Thus, every normal governor must have a positive

probability of acting responsibly. That is, provincial democracy cannot fail (at least as long as the president is not the virtuous type). With frustration of national democracy, the lack of turnover at the national level means that the motivation for a normal governor to serve responsibly cannot come from any hope of gaining national power, and so it must come from a threat that provincial voters would not re-elect him if he serves corruptly. Thus, provincial democracy cannot be frustrated. Q.E.D.

When national democracy is frustrated, governors see no chance of being elected president, and so they anticipate careers entirely contained within their current province. With such isolation of provincial politics, it may be reasonable for provincial voters to satisfy the property of stationary reactions to corruption: that there is a constant probability of provincial voters rejecting a governor in the next election after any period in which he has acted corruptly. This property enables us to strengthen the conclusions of Theorem 2 as follows.

Theorem 3. In any federal equilibrium, if national democracy is frustrated and provincial voters have stationary reactions to corruption then, with probability one, each province will eventually get a governor who always serves responsibly.

Proof of Theorem 3. Recall first our observation that frustration of national democracy implies failure of national democracy. So by Theorem 2, frustration of national democracy implies that provincial voters must have a positive probability of rejecting a corrupt governor. With stationary reactions to corruption, this positive probability must be same after each period of corrupt government in the province.

If Theorem 3 were false then, in some positive-probability event, there would be an

infinite number of periods in which a governor serves corruptly, and after each of these periods there would be a constant positive probability that the incumbent would be replaced by a new governor, and each replacement has a positive ε probability of being the virtuous type who always serves responsibly. But then the probability of never getting any virtuous governor must be zero. Q.E.D.

4. Decentralized provisional government in a process of transition to unitary democracy

In the preceding sections we considered the effects of federal decentralization in a democracy that has a fixed permanent constitutional structure. But it may be a mistake to assume that a new democracy must begin with the adoption of a permanent constitution. Given the fragile vulnerability of new democracies, the optimal constitutional structure for the first difficult years of democratic politics may be quite different from the optimal constitutional structure for governing the nation in the long run. Even in a nation where a unitary democratic structure is considered preferable for the long run, there may be advantages of adopting a more decentralized constitutional structure during an interim phase of transition to democracy.

So let us now consider a process of transition to democracy in which there is an initial phase of T periods when there will be only local democracy in the N provinces. The central government could be kept weak in this transitional period by constituting national authority only by a council of representatives from the provincial governments, each subject to recall at any time, as in the American Articles of Confederation. But suppose then that a unitary national democracy will be established and the first president will be elected in period $T+1$. All parameters from the previous model ($N, b_0, c_0, x_0, w_0, b_1, c_1, x_1, w_1, \rho, \varepsilon$) have the same

interpretations as before, except that now the provincial parameters (with 0-subscripts) apply only in the first T periods, and the national parameters (with 1-subscripts) apply after period T . Each politician initially has a small probability ε of being the virtuous type. The basic parametric assumptions (1) and (2) are still applied in this section.

Theorem 4. Suppose $\rho^T(b_0 + c_0) > (1 - \rho^T)c_0$, $b_1 + c_1 \geq N(b_0 + c_0)$, and $w_0 > x_0$. In any equilibrium where national democracy is expected to fail after period T , decentralized democracy must succeed until period T , and any corrupt governor would be replaced by provincial voters. So there cannot be consistent frustration or failure of democracy in any equilibrium of this transitional process. But there is an equilibrium in which democracy consistently succeeds at all periods.

The first inequality in Theorem 4 holds if $\rho^T \geq 0.5$, which allows a transition period of up to 13 years when the annual discount factor is 0.95. The second inequality says that the president of the unitary national government after time T will get all the opportunities for corruption that are available to the governors of the N provinces in the initial phase. The third inequality says that the welfare gain that provincial voters get from responsible government in one period exceeds their cost of changing leadership.

Proof of Theorem 4. Assuming that national democracy will fail after period T , the national voters at period $T+1$ will elect a president with highest probability of being virtuous, given his record. If some governors had any positive probability of acting corruptly, then by acting responsibly they could make voters believe that their probability of being virtuous was more than ε , and so one of them would be elected president. There can be at most N such

governors alive with good reputations at period $T+1$, and so some of them must expect at least $1/N$ probability of being elected president. A governor's expected cost of governing responsibly for T periods is

$$c_0(1 + \rho + \dots + \rho^{T-1}) = c_0(1 - \rho^T)/(1 - \rho),$$

but his expected gain from being a candidate for president after T periods is at least

$\rho^T(1/N)(b_1+c_1)/(1-\rho)$. The inequalities in the theorem imply that

$$\rho^T(1/N)(b_1+c_1)/(1-\rho) \geq \rho^T(b_0 + c_0)/(1-\rho) > c_0(1 - \rho^T)/(1 - \rho),$$

So the gain from governing responsibly is strictly greater than the cost, and so no governor would choose to behave corruptly in the first T periods.

A governor who already had a corrupt record would have no incentive to be responsible at T , so (with $w_0 > x_0$) provincial voters would replace him at T . By backwards induction, we can extend this argument to show that any governor with a corrupt record would be rejected by the voters during the decentralized initial phase of T periods.

In the equilibrium where democracy consistently succeeds, provincial voters at periods $2, \dots, T$ and national voters after period $T+1$ would reject any incumbent who has acted corruptly, and national voters at period $T+1$ will select a president at random from among the governors (if any) who served responsibly at period T . Governors are motivated to behave responsibly at any period t before $T+1$, because the basic parametric assumptions (1) and the inequalities in the theorem together imply that

$$\begin{aligned} b_0(1+\rho+\dots+\rho^{T-t}) + (1/N)\rho^{T+1-t} b_1/(1-\rho) &= b_0(1-\rho^{T+1-t})/(1-\rho) + (1/N)\rho^{T+1-t} b_1/(1-\rho) \\ &\geq (1-\rho^{T+1-t})(b_0+c_0) + \rho^{T+1-t} (b_1+c_1)/N \geq b_0+c_0. \end{aligned} \quad \text{Q.E.D.}$$

5. Discussion

By comparing the sets of equilibria in three simple models of democratic competition, we have tried to show how federal separation of powers or a decentralized provisional government can improve incentives for politicians to make democracy succeed.

We argued first that, under a unitary political system, the success or failure of democracy may depend on the equilibrium beliefs of voters and politicians. Confronted with such a multiple equilibrium problem in an abstract game-theoretic model, it is tempting to argue that voters simply need to be focused on one of the good equilibria, where successful politicians expect to hold power only as long as they maintain reputations for responsibly serving the voters. But in a nation where democracy has not previously existed, such a good equilibrium requires that good democratic reputations be attributed to political leaders who actually have no prior history of using power responsibly for voters' benefit. Under the previous nondemocratic political system, successful politicians would have had to cultivate very different kinds of reputations: for loyally serving the authoritarian leader and for reliably rewarding active supporters. So in a new democracy, we may generally expect to find mid-career politicians who have invested years in cultivating reputations for using public funds principally to benefit others in the power elite itself. If such a leader suddenly begins allocating more resources to serve the voters, then he may risk losing the trust of others in the power elite, with no political gain for himself when the voters are not confident of getting such service in the future. Thus, voters and politicians in a new democracy may be strongly focused on the bad equilibrium of democratic failure.

We then argued that, under a federal political system, an anticipated failure of democracy at the national level would generate incentives for local politicians to serve responsibly in

provincial governments. Thus a federal system can effectively offer an insurance policy against general failure of democracy. In any equilibrium under federal democracy, voters must see benefits of active democracy at some level of government.

In claiming that federalism must sharpen the incentives for responsible government at one level or the other, we have never claimed or predicted that provincial government would be more or less corrupt than national government. Indeed, if the long-run survival of democracy ultimately depends on its success at the national level, then survival selection might actually generate a population of democracies where federal countries have statistically more corruption than unitary, precisely because all surviving democracies have fairly responsible national governments but some federal democracies have corrupt provincial governments. Thus, our model is compatible with Treisman's (2000) findings of greater corruption in federal systems. But if provincial democratic success can teach voters to expect national democratic success, then federalism should also yield statistically higher survival rates, as Boix (2003) observes.

Our theoretical argument could be undermined by strong regional identities, if the ideology that molds the ϵ behavioral types involves strong local identity. That is, if the most likely behavioral-type that would abstain from corruption in a provincial government was not generally virtuous but only locally chauvinistic, then responsible local service would not be so effective for building a reputation that could appeal to national voters.

To refocus the analysis of democratic failure, we have been assuming that voters' constitutional power to replace an incumbent leader is not in question. By classic Madisonian arguments (from the Federalist Papers 47-51), any constitutional constraint on an incumbent leader must be enforceable by other leaders with appropriate power and motivation, and so the

protection of any constitutional limits requires a system of separation of powers. In a very general sense, the argument in this paper may be viewed as an extension or modification of this classic argument for how federal separation of powers can support democratic survival. In our model of federalism, provincial leaders do not force national leaders to hold free national elections, but by offering themselves as potential candidates they guarantee that voters will have a good reason to continue demanding free national elections.

Under any system of separation of powers, agency problems can increase political corruption at the boundaries where mixed effects of different branches of government make responsibility unclear (Treisman, 1999). But regional separation of powers in federalism may yield clearer boundaries between independently elected officials' domains of control than other functional ways to separate powers. Otherwise, the role of governors in our analysis, as elected officials with aspirations to higher office, might be taken just as well by directly elected cabinet ministers of lower ranks.

The competitive analysis here can be understood with parallels to oligopoly theory, where profit-taking in equilibrium is reduced by higher elasticity of demand and lower barriers to entry. Political corruption may be seen as an analogue of oligopolistic profit. In our argument, federalism lowers entry-barriers into national politics when it gives independent provincial leaders an opportunity to prove their qualifications for national leadership. And the possibility of advancement to greater national office gives provincial leaders a higher elasticity of demand for political leadership with respect their corruption-price. Such political elasticity can also be created in a federal system by Tiebout effects. The essence of the Tiebout argument is that, with national mobility of people and resources, local corruption erodes its own tax base.

The observation that federalism sharpens political competition can offer insights into the tensions of the federal bargain between national and provincial leaders (Riker, 1964).

Understanding that governors are potential rivals for national power, national leaders may prefer a system of national criminal sanctions against provincial corruption as a way to prevent governors from building virtuous reputations without habituating voters to reject corrupt incumbents. The appeal of secession for governors (especially for governors who have been corrupt) is increased when local rivals' national ambitions make local politics more competitive. Bilevel political parties may serve politicians by moderating this competitive tension.

Our analysis should have implications for current efforts to cultivate democracy in countries where it has not existed before. From our perspective, the danger of the democratic-failure equilibrium is greatest where historical experience suggests that citizens should expect little from their leaders. In American-occupied Iraq, defeat of Ba'athist totalitarianism left a political vacuum. The first leader to control the national government there could build a patronage network far stronger than any opposition group. Nobody would be surprised if this leader then suppressed political opposition and abused power for the benefit of himself and his supporters. Indeed, citizens might reasonably expect that any further political change would just install another leader who would behave the same way. This is exactly the logical structure of democratic failure that we saw in Section 2. But if local elections were held first in a decentralized provisional government, then local leaders with national ambitions would have a positive incentive to begin cultivating a reputation for responsible democratic leadership, as we saw in Section 6. Although the Democratic Principles Work Group (2002) suggested just such a plan, the strongest emigré leaders had no incentive to recommend a decentralized political system

designed to decrease the entry-barriers against new political rivals.

Such a decentralized transitional structure would seem less likely to succeed in Afghanistan, where provincial governments have been dominated by local warlords. In a new democracy, old leaders with reputations adapted to the old system may expect to be replaced by new leaders who can build reputations for democratic responsibility. So provincial warlords who have already violated norms of democratic responsibility would rationally use their substantial power to impede the establishment of new democracy at the provincial level. Thus, problems of co-opting, pensioning, or defeating such established local leaders must be understood as central to the problem of democratization in a nation like Afghanistan. In such a situation, local democracy might well have to develop after national democracy.

Democracy is worth cultivating because the structure of political institutions makes a difference to the performance of government. But new democracies are vulnerable to failure and breakdown, and a democratization process needs every advantage that can be derived from careful analysis of different democratic institutions. At a time when great armies have been sent across the world with an announced goal of cultivating democracy, the scarcity of interest such analysis seems particularly disappointing.

REFERENCES

- Daron Acemoglu and James A. Robinson, "Why did the West extend the franchise? democracy, inequality and growth in historical perspective" Quarterly Journal of Economics 115:1167-1199 (2000).
- Carles Boix, Democracy and Redistribution, Cambridge U Press (2003).

- Avinash Dixit and John Londregan, "Fiscal federalism and redistributive politics," Journal of Public Economics 68:153-180 (1998).
- Matthew Ellman and Leonard Wantchekon, "Electoral competition under the threat of political unrest," Quarterly Journal of Economics 115:499-531 (2000).
- Democratic Principles Work Group (Ghassan Atiyah and Kanan Makiya, coordinators), Final report on the transition to democracy in Iraq (November 2002)
http://www.journalofdemocracy.org/Articles%5CIraq_Report.pdf.
- Joseph E. Harrington, "Progressive ambition, electoral selection, and the creation of ideologues." Economics of Governance 1:13-24 (2000).
- Juan J. Linz and Arturo Valenzuela, eds., The Failure of Presidential Democracy: Comparative Perspectives, Johns Hopkins U Press (1994).
- Seymour M. Lipset, "Some social requisites of democracy: economic development and political legitimacy," American Political Science Review 53:69-105 (1959).
- Roger B. Myerson, "Effectiveness of electoral systems for reducing government corruption," Games and Economic Behavior 5:118-132 (1993).
- Wallace E. Oates, "An essay on fiscal federalism," Journal of Economic Literature 37:1120-1149 (1999).
- Torsten Persson and Guido Tabellini, "Federal fiscal constitutions: risk sharing and redistribution," Journal of Political Economy 104:979-1009 (1996).
- William Riker, Federalism: Origin, Operation, Significance, Little Brown (1964).
- William Riker, Liberalism Against Populism, Freeman (1982).
- Thomas C. Schelling, The Strategy of Conflict, Harvard U. Press (1960).

Alfred Stepan and Cindy Skach, "Constitutional frameworks and democratic consolidation: parliamentarism versus presidentialism," World Politics 46:1-22 (1993).

Charles Tiebout, "A pure theory of local expenditures," Journal of Political Economy 64:416-424 (1956).

Daniel Treisman, "Political decentralization and economic reform: a game-theoretic analysis," American Journal of Political Science 43:488-517 (1999).

Daniel Treisman, "Decentralization and the quality of government," UCLA working paper (2000).

Barry Weingast, "The economic role of political institutions: market-preserving federalism and economic growth," Journal of Law, Economics, and Organization 11:1-31 (1995).

Barry Weingast, "The political foundations of democracy and the rule of law," American Political Science Review 91:245-263 (1997).