Berkeley Law

From the Selected Works of Robert Cooter

January, 2006

The Optimal Number of Governments for Economic Development

Robert D. Cooter

Available at: https://works.bepress.com/robert_cooter/105/
Market-Augmenting Government

The Institutional Foundations for Prosperity

Omar Azfar and Charles A. Cadwell, Editors

The University of Michigan Press
Ann Arbor
The Optimal Number of Governments for Economic Development

Robert D. Cooter

"Does government correct or create market failures?" For several decades, this question framed debate among economists about the state's role in the economy. The state corrects market failures by supplying public goods and suppressing external harms, whereas the state creates market failures by restricting competition and protecting monopoly power. Recent history in developing countries, however, points to a further role of the state: supplying the fundamental legal framework required by markets for private goods. A modern economy requires technical specialization and a fine division of labor, which in turn requires exchange among strangers at a distance in time and space. Clear property rights and enforceable agreements facilitate cooperation among strangers who do not share mutual trust. So the legal framework for markets especially entails property and contract law.

In view of this history, Mancur Olson reformulated the economic question of the state as, "How does government augment the economy?" According to Olson's formulation, government augments the economy by supplying both public goods and the legal framework for private goods. Stable governments and incorruptible officials augment markets best. Conversely, government instability and corruption undermine the supply of public goods and the security of property and contracts. This essay is concerned with the connection between instability, corruption, and democracy. Under certain circumstances, replacing broad, deep governments with shallow, narrow governments increases stability and reduces corruption. My general conclusion is that developing nations that are plagued by instability and corruption probably have too few elections and too few democratic governments. Under these circumstances, more extensive democracy augments markets better than less extensive democracy.
I. Hierarchies and Elections

An automobile manufacturer can choose to make tires for its cars in a subsidiary or buy tires from another firm. Making tires involves one firm using hierarchical organization (i.e., vertical integration of the tire-making function), whereas buying tires involves two firms transacting in a market.

According to the standard formulation, firms are hierarchies bounded by markets. Hierarchies work by means of orders, whereas markets work by means of bargains. At the point where an organization touches a market, administration ends and trade begins; vertical integration therefore subtracts markets, and vertical disintegration adds markets. Along the vertical dimension, then, the extent of hierarchy and the number of markets measure the same thing with different signs. Optimizing the extent of hierarchy is consequently the same as optimizing the number of markets.

Turning from the private to the public sector, the ministry of education can administer local schools or, alternatively, the citizens in each locality can elect a school board. The ministry of education is a single organization, and each school board is a small government. In this example, administration involves one hierarchical organization, while politics involves many governments. Just as the private sector consists of hierarchies and markets, so the public sector consists of hierarchies and governments. Orders direct hierarchies, and elections direct democratic governments. At the point where a public organization touches an election, administration ends and politics begins; vertical integration thus subtracts elections, and vertical disintegration adds elections. Along the vertical dimension, the extent of hierarchy and the number of elections measure the same thing with different signs. Optimizing the extent of hierarchy is consequently the same as optimizing the number of elections vertically.

My analogy between the firm and the state has focused on vertical integration or disintegration. Now I turn from vertical to horizontal organization. One firm can manufacture cars and another can grow cucumbers or a single firm can do both. Similarly, an elected town council can control police while a separately elected school board controls schools, or the town council can control both police and schools. In general, the scope of a government's power can encompass many issues or just a few. When one democratic government controls each issue, the citizens vote for many sets of officials. In a democracy, the citizens vote for as many sets of officials as there are governments. Horizontal integration subtracts elections and horizontal disintegration adds elections. Along the horizontal dimension, the scope of government and the number of elections measure the same thing with different signs. Optimizing the scope of government is consequently the same as optimizing the number of elections horizontally.

This essay concerns the optimal number of democratic governments. I search for the optimum along the vertical dimension in which hierarchy trades off with government, and I also search along the horizontal dimension for the optimal breadth of each government. This approach yields some novel conclusions about instability and corruption, which especially (but not uniquely) afflict developing nations. In the private sector, many small firms imply shallow hierarchy and narrow product lines, and in the public sector many small governments imply shallow hierarchy and narrow government. This essay explains under what conditions replacing broad, deep governments with shallow, narrow ones increases stability and reduces corruption.

Organizational Space

Figure 1 depicts organizational space with the vertical dimension representing the depth of hierarchy and the horizontal dimension representing the breadth of each government. A point in the organization space of figure 1 is specified by a given depth of hierarchy and breadth of government or, equivalently, by a given number of elections on the vertical and horizontal dimensions.

Different points in the organization space of figure 1 correspond roughly to countries with different political subdivisions. Deep hierarchy and broad government, as indicated by the origin of the graph, characterize unitary states like Japan or France. Moving vertically from the origin holds the breadth of government constant while hierarchy becomes shallow. This move roughly depicts the change from a unitary state to a federal system like that of Canada or Australia. The provinces of Canada and the states of Australia have broad powers subordinated in some respects to the federal government. Moving horizontally from the origin holds the depth of hierarchy constant while governments narrow. This move depicts the multiplication of special governments with single purposes. To illustrate, in the San Francisco area special district governments with separate elections provide regional parks, public transportation, water, and other local public goods. Finally, moving diagonally from the origin, governments...
The Optimal Number of Governments for Economic Development

sector has engaged the best minds in economics, and their findings contributed to global privatization in the 1990s. The boundary between the public and private sector, however, is not my topic. The problem of the optimal size of the state differs from the problem of the optimal number of governments. In effect, my analysis varies the number of governments while holding constant the size of the state.

Convention distinguishes the branches of government into the executive, legislature, and judiciary. Dividing power among the branches of government has engaged the best minds in political theory. Their findings contributed to the spread of stable democracies and the rule of law. The division of powers, however, is not my topic. The problem of the optimal division of powers within a state differs from the problem of the optimal number of governments. In effect, my analysis varies the number of governments while holding constant the division of powers.

Sometimes people can choose among alternative governments by moving or stipulating jurisdiction in a contract. Governments can compete to attract people, either by moving people to governments, as through migration, or by moving government jurisdiction to people, as through choice of laws in a contract. Economists have identified circumstances under which competition can cause a race to the top or a race to the bottom. Ideally, competition among governments can optimize their number just as competition among firms can optimize their number. Competition among governments, however, is not my topic.

An efficient economy provides people with the greatest feasible satisfaction of their preferences for private goods. Similarly, an efficient state provides citizens with the greatest feasible satisfaction of their preferences for public goods and the legal framework for private markets. Just as markets respond to consumers, so elections respond to citizens. Given favorable circumstances, the responsiveness of elections to citizens makes democracy the best form of government for satisfying preferences. Instead of asking, this essay assumes that democracy is the best form of government. Having assumed that some form of democracy is best, I ask how many elections are optimal.

Just as law constrains abuse by judges, so civil service rules limit corruption in state administration. The civil service is apparently the best form of administration to prevent corruption among officials. Instead of asking, this essay assumes that a civil service bureaucracy is better than politicized bureaucracy. Having made this assumption, I determine the optimal depth of bureaucracy.

narrow and hierarchies become relatively shallow. This move depicts the simultaneous decentralization and fragmentation of government.

This essay concerns the optimum in the organizational space represented by figure 1. The optimum occurs at the depth of hierarchy and breadth of government that maximizes the satisfaction of citizens. By inserting utility curves representing the satisfaction of citizens in figure 1, I could depict the optimum, but I will avoid this complication. Instead, I will analyze separately the optimum along the vertical and horizontal dimensions.

This Paper Is Not About...

Figure 1 depicts the two dimensions that characterize the problem that I call the "optimal number of governments." I will isolate and examine two variables—the horizontal and vertical number of elections—that affect the performance of government. Explaining variance in the performance of governments in general requires a model with many more dimensions. The omitted dimensions cause confusion between the optimal number of governments and related problems. To avoid confusion, I want to mention some problems that this essay is not about.

Locating the optimal boundary between the public and private
lit. Factor or Splice?

I begin with an analysis of political instability. The conventional analysis of unstable democracy focuses on the relationship between the legislature and the executive. The executive provides coherence through orders. Inflating in the legislature, however, can cause paralysis or chaos. Given these facts, increasing executive power at the expense of the legislature can increase stability. The gain in stability, however, creates a risk to democracy: the executive might suspend elections and rule by decree. In the conventional analysis, the executive provides stability and potentially endangers democracy, whereas the legislature provides democracy and potentially endangers stability.

Instead of the usual analysis, I view stability as depending on the number of governments. Under certain circumstances, single-purpose governments are more stable than comprehensive governments. Under these circumstances, increasing the number of governments and narrowing their scope, or organizing single-issue referenda, increase stability. I will explain the circumstances under which increasing the number of elections increases political stability.

Broad jurisdiction splices independent issues together like the strands of a rope. By splicing I mean combining issues and deciding them all at once. For example, the U.S. Congress often enacts omnibus legislation with extensive logrolling. To get spliced indirect democracy, citizens should elect representatives to assemblies with power over many different issues. In contrast, narrow jurisdiction factors politics into independent issues, like a mathematician dividing a large number into prime numbers. By factoring I mean separating issues and deciding them one at a time. For example, citizens may elect a town council to control the police and a school board to control schools. To factor indirect democracy, citizens should elect separate governments for separate issues. To factor direct democracy, citizens should decide each issue in a separate referendum. In figure 1, factoring represents a move to the right along the horizontal axis, whereas splicing represents a move to the left.

Sometimes a constitution factors, as when the town’s constitution establishes an elected council and a separately elected school board. Alternatively, a constitution may allow for factoring without requiring it. For example, the state constitutions in the United States prescribe procedures for establishing special governments for such activities as parks, transportation, and water. Citizens can establish or

abolish special governments by following the prescribed procedures. Alternatively, the constitution may limit or forbid factoring, as when it prevents a branch from delegating authority or a government from ceding authority.

I have discussed clear-cut cases of factoring, but unclear cases often occur. In the European Union, the selection of ministers forming the Council of Ministers differs for different issues. Thus, the council may consist of the national ministers of agriculture to decide a question about farm subsidies, whereas it may consist of the national ministers of transportation to decide a question about railroads. The national ministers of finance, however, often dictate to other national ministers. Does changing the council’s membership according to the issue in question amount to factoring the issues? As I will explain, the answer depends on transaction costs.

Transaction Costs of Bargaining

I will use a simple model of collective choice to analyze factoring and splicing in indirect and direct democracy. In the civil service, officials rank from low to high and administration proceeds by orders. In contrast, members of the town council hold the same rank and government often proceeds by bargains. In general, superiors give orders and equals bargain.

Instead of pure hierarchy or pure equality, government typically mixes orders and bargains. To illustrate, in some American cities the members of the town council bargain with each other in the shadow of the mayor. A strong mayor causes the town council to proceed relatively more by orders, whereas a weak mayor causes the town council to proceed relatively more by bargains. Similarly, the oxymoron “first among equals” traditionally describes the relationship between the British prime minister and his cabinet. As suggested by this phrase, British cabinet ministers bargain with each other in the prime minister’s shadow.

As explained, reality mixes orders and bargains, but my analysis separates them for clarity’s sake. Assume that each government consists of elected officials who bargain with each other. Spliced government lowers the transaction costs of bargaining across issues. To illustrate, the advocates of strong police may want to cut a deal with the advocates of rich schools. If the town council controls police and schools, the council members can bargain and agree to implement strong police and rich schools. Conversely, factored government
raises the transaction cost of bargaining across issues. To illustrate, if the town council controls police and the school board controls schools, the advocates of strong police may have difficulty cutting a deal with the advocates of rich schools.

I have explained that factoring raises, and splicing lowers, the transaction costs of bargaining across issues. Now I consider what determines whether citizens benefit more from low or high transaction costs.

Complements

Reasoning by analogy with the private sector, proponents justify horizontal mergers in private firms based on "complementarities" that create "economies of scope." To illustrate, automobiles and cucumbers are unrelated products. Manufacturers of automobiles and growers of cucumbers, however, may face comparable labor regulations. Consequently, merging the manufacturer of automobiles and the grower of cucumbers might save total costs of regulatory compliance.

Similarly, the citizens under the jurisdiction of a government might have complementary tastes in public goods. To illustrate, consider an example with two citizens, A and B, and two public goods, X and Y. If A intensely wants X and feels indifferent about Y, whereas B intensely wants Y and feels indifferent about X, then A and B have complementary tastes for X and Y. A and B can cut a deal to help satisfy their most intense desires. B supports A's efforts to obtain X, and A supports B's efforts to obtain Y.

The scope of complements determines the potential gain from political bargains. When different political factions have complementary tastes for public goods, splicing lowers the transaction costs of political bargains, thus increasing both the probability of a bargain and the size of the resulting surplus.

Conflict

Citizens, however, may have noncomplementary tastes. To illustrate, if A intensely likes X and B intensely dislikes X, then A and B have noncomplementary tastes for X. The differences in preferences of A and B for X provide no basis on which to cut a deal. Given purely noncomplementary preferences, politics becomes a game of pure conflict in which one player's win is another's loss.

Conflict in a democracy can cause instability for reasons clarified by some technical terminology. By definition, the core of a game is the set of unblocked distributions. Since every proposal in this case is blocked by an alternative, the game has an empty core. Majority rule games of distribution with symmetrical players generally have an empty core. To see why, assume that three voters, denoted A, B, and C, must distribute $100 among them by majority rule. Initially, someone proposes to divide the money equally: $(A,B,C) = (33,33,33)$. A's counterproposal is to share the surplus equally with B and give nothing to C: $(A,B,C) = (50,50,0)$. A and B can implement A's counterproposal under majority rule, and A's counterproposal makes A and B better off than the initial proposal, so A's counterproposal blocks the initial proposal. It is not hard to see that any proposal is blocked by another proposal. Thus A's proposal is blocked by B's counterproposal to distribute the surplus

$(A,B,C) = (0,75,25)$.

When the core is empty, each player can make credible demands whose satisfaction is infeasible. The contest for redistribution by majority rule destabilizes every possible coalition. Generalizing these results, Arrow proved that any democratic constitution can result in cyclical voting. Voting cycles, especially those provoked by a contest for redistribution, destabilize democracies, especially in developing countries.

Median Rule

Instead of cycling, many democracies produce stable government that pursues policies near the center of the political spectrum. Under certain conditions, voting among paired alternatives along a single dimension of choice yields an equilibrium at the point most preferred by the median voter. The median rule explains why the center dominates the politics of many democracies.

A problem of majority rule concerns intensity of sentiment. Democracy gives equal weight to all votes, regardless of how strongly the voter feels about the issues. From an efficiency perspective, however, more weight should be given to intensive preferences. To illustrate, assume that the chairman of a three-person committee asks each one to write his or her vote concerning a certain proposal on a slip of paper. When the slips are collected, the chairman reports, "I have two slips marked 'Yes' and one marked 'No, No, oh please,
Majority rule gives equal weight to a bland "Yes" and a passionate "No." The unresponsiveness of majority rule to the intensity of feeling about issues causes its inefficiency.

Being unresponsive to intensities, the median rule is not generally efficient by the cost-benefit standard. Under a very special assumption involving symmetrical preferences, however, the median rule is cost-benefit efficient. \(^\text{10}\)

If factoring results in asymmetrical distributions of preferences, then the median rule results in efficient outcomes. Splicing, however, enables the factions to bargain across issues and express their intensity of sentiment by trading votes. Asymmetrical distributions of preferences thus create complementarities and economies of scope that splicing potentially exploits. When the median rule is efficient, the voting equilibrium leaves no scope for bargaining. In the typical case, however, the median rule is inefficient, so successful bargaining creates a surplus.

City Council and School Board

I have explained that splicing lowers the transaction cost of bargaining across issues and successful bargaining across issues can increase the satisfaction of voters with complementary tastes. Splicing also increases the risk of failed bargains and circular votes. When spliced voting causes intransitivity, factored voting may improve the outcome by allowing the median voter to prevail on separate dimensions of choice. Median rule on separate dimensions of choice often satisfies the preferences of voters more efficiently than an unstable contest of distribution. Intransitive preferences in multidimensional choice may factor into single-peaked preferences on each single dimension of choice. In general, single-purpose government is like a safe stock with a modest yield, whereas multipurpose government is like a risky stock that pays a lot or nothing.

To develop these points, I turn to a quantitative example. Assume that expenditures on police and schools are the two majority political issues in a small town. Police and schools especially represent the two ways in which government augments markets. Police are essential to secure property rights and enforceable contracts. Schools are essential to preparing people to participate in a modern, technological economy.

First consider splicing the issues of police and schools. A town council that decides both issues provides a forum for bargaining. If bargaining succeeds, council members who care intensely about police may trade votes with council members who care intensely about schools, so that each one gets what he or she wants most. If bargaining fails, the council members may waste resources in an unstable contest of distribution. Second, consider factoring the issues. A town council that controls police and a separately elected school board that controls schools denies a forum for bargaining over the two issues. With bargaining obstructed and assuming single-peaked preferences, the median voter prevails on each dimension of choice.

Figure 2 sharpens the example with numbers. Assume that voters in a town are divided into equal numbers of liberals, conservatives, and moderates. Expenditure can be high or low for schools and police, with the resulting net benefits for each group of voters indicated in figure 2. \(^\text{11}\) The liberals intensely prefer high expenditures on schools and mildly prefer the savings in taxes from low expenditures on police. The opposite is true of conservatives, who intensely prefer high expenditures on police and mildly prefer the savings in taxes from low expenditures on schools. The moderates mildly prefer the tax savings from low expenditures on police and schools. The row labeled "total" indicates the sum of net benefits to the three groups.

Assuming majority rule, contrast the consequences of splicing and factoring issues in figure 2. If the issues are factored, then two out of three voters (conservatives and moderates) vote for low expenditures on schools, so factoring results in low expenditures on schools. Furthermore, two out of three voters (liberals and moderates) also vote for low expenditures on police, so factoring results in low expenditures on police. Thus, factoring results in low expenditures on schools and police.

If issues are spliced, the voters must choose among four combinations of public goods depicted in the columns of figure 3. The net
Expenditures on Schools and Police, Respectively

<table>
<thead>
<tr>
<th></th>
<th>(1) high, high</th>
<th>(2) low, low</th>
<th>(3) high, low</th>
<th>(4) low, high</th>
</tr>
</thead>
<tbody>
<tr>
<td>liberal</td>
<td>11</td>
<td>1</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>conservative</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>moderate</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>total</td>
<td>22</td>
<td>7</td>
<td>15</td>
<td>14</td>
</tr>
</tbody>
</table>

Fig. 3. Voter’s net benefits from combinations of public goods

benefits to voters depicted in figure 3 are calculated from the numbers in figure 2. For example, (low, high) indicates low expenditures on schools and high expenditures on police, which results in a payoff of zero for liberals, 12 for conservatives, and 2 for moderates.

The numbers in figure 3 can be used to deduce the winner in a vote between any two alternatives. If voters simply vote their preferences in figure 3, without bargaining or trading, then an intransitive cycle results. Specifically, two of three voters (liberal and conservative) prefer (high, high) rather than (low, low). Two of three voters (conservative and moderate) prefer (low, low) rather than (high, low). Two of three voters (liberal and moderate) prefer (high, low) rather than (low, high). And, finally, two of three voters (conservative and moderate) prefer (low, high) rather than (high, high). As explained, column 1 beats column 2, 2 beats 3, 3 beats 4, and 4 beats 1. Thus, voting in figure 3 results in an intransitive cycle.

Figures 2 and 3 illustrate the general principle that splicing dimensions of choice can cause intransitivity where none exists on any single dimension of choice. Instead of simply voting their preferences, however, splicing may cause the voters to bargain with each other and cooperate. Since liberals care more about schools than police, whereas conservatives care more about police than schools, they have complementary tastes and could profitably trade votes. A platform calling for high expenditures on schools and police allows the liberals and conservatives to get what they want on the issue that each one cares the most about, as required for efficiency. Stabilizing such an agreement requires the parties to abandon the majority rule game of distribution, which has no core, and cooperate with each other.

Whether comprehensive government or single-purpose governments satisfy the preferences of political factions better depends on the ability of politicians to cooperate. In general, splicing issues increases the gains from cooperation and factoring issues decreases the losses from conflict. Finding the optimal number of governments requires balancing these considerations. These facts suggest the prescription, “Splice when cooperation is likely and factor when conflict is likely.”

Factoring by Referenda, Splicing by Legislation

Most constitutions that permit referenda restrict them to a yes or no vote on a single issue. To illustrate, Californians might be asked to vote yes or no on restricting abortions and yes or no on capital punishment, but the law precludes Californians from being asked to vote yes or no on restricting abortion and restricting capital punishment. A practical reason compels restricting each ballot initiative to a single issue. Logrolling, which combines issues in a single vote, requires bargaining. Bargaining among different groups requires representation. Ballot initiatives bypass elected representatives. Thus a multiple-purpose ballot initiative would invite bargaining without any framework for it.

In legislatures, the members often bargain, compromise, and draft a single bill that combines different issues. In contrast, rules restricting ballot initiatives to a single issue prevent logrolling, so different groups have little incentive to bargain or vote strategically. When citizens vote their preferences on a single dimension of choice, the median usually prevails. In general, direct democracy factors the issues, so the median voter should prevail. In contrast, members of legislatures bargain, compromise, and roll logs. In general, indirect democracy splices issues, which should result in bargains or cycles.

The contrast between splicing and factoring predicts some consequences of a shift from indirect to direct democracy. A change from indirect to direct democracy often replaces cycles or bargains among representatives with the preference of the median voter on each dimension of choice. Is this change better or worse? That depends on how well indirect democracy works. Given informed voters and competitive elections, indirect democracy produces effective representation of political interests. If representatives bargain successfully and cooperate with each other, then citizens get their way on their preferred issues. Under these circumstances, indirect democracy satisfies the preferences of voters better than direct democracy does.

Indirect democracy, however, can also create a political cartel whose members conspire to blunt electoral competition. For example, the spectacular disclosure of corruption among leading Italian politicians in the 1990s suggests that citizens had little influence over
deals struck by their representatives. An opaque political process and proportional representation made Italian electoral competition relatively ineffective. Under these circumstances, a change to direct democracy can break the political cartel.

In addition, indirect democracy can cause an unstable contest of redistribution among interest groups. Changing to direct democracy can increase stability, which should increase the satisfaction of citizens with politics.

I have explained that direct democracy causes the median voter to prevail on each dimension of choice, which is better than a cycle or a political cartel and worse than perfect bargaining by elected representatives. This proposition summarizes the main difference in theory between direct and indirect democracy. Besides this large difference, some small differences are sometimes important.

First, direct democracy gives more weight to those citizens who actually vote, whereas indirect democracy gives more weight to the number of citizens living in a district. To illustrate, assume that poor people, who vote at relatively low rates, live in poor districts. Indirect democracy apportions representatives by population, so the number of representatives from poor districts reflects the number of poor citizens, including those who do not vote. In contrast, direct democracy responds to the citizens who actually vote. Thus, in the preceding example, in which rich people vote at higher rates than poor people, direct democracy gives more weight to the opinions of rich people. This phenomenon may tilt California ballot initiatives in favor of older, conservative, white citizens.

Second, critics of direct democracy allege that the majority of citizens will vote to redistribute wealth from the few to the many. For example, if most citizens buy auto insurance, they will vote to cut its price. Or, if most citizens rent houses, they will vote for rent control. More generally, critics of direct democracy allege that the majority of citizens will vote to undermine the rights of the minority.

This criticism, however, has a weak foundation in theory. From the viewpoint of theory, direct democracy favors voting, which does not necessarily harm minorities more than splinter voting. Splitting voting encourages citizens to coalesce into blocs in order to bargain with each other. While a system of proportional representation can guarantee representation in political bargaining to every minority group, two-party competition contains no such guarantees. When groups coalesce, some minorities may suffer permanent exclusion from the ruling coalition.

In contrast, after factoring the issues, the minority on one dimension of choice is seldom the same group of people as the minority on another dimension of choice. Any single person with complicated political views wins on some dimensions of choice and loses on others. In general, factoring issues can dissolve large blocs of citizens and ensure that everyone wins some of the time. In addition, all the nonmedian voters participate in determining the median voter. Thus everyone's preferences have an effect on voter equilibrium.

Any democratic system of politics, whether direct or indirect, requires protection of minorities, such as ethnic groups and wealthy people. Forms of protection include bicameralism and constitutional rights. Thus, the Bill of Rights in the U.S. Constitution constrains the states, so a federal judge would nullify a California referendum that violates the Bill of Rights. This fact imposes an essential constraint on California's referenda. Furthermore, on may political issues, the bicameral U.S. Congress can preempt states by enacting federal legislation.

Besides the legal obstacles, transaction costs currently limit the frequency of referenda. Specifically, electoral rules usually require the costly gathering of signatures to create a ballot initiative, thus limiting the number of referenda placed on the ballot. In the future, however, technological developments such as collection of signatures over the Internet and electronic voting could dramatically lower the transaction cost of direct democracy. With lower costs, the pace of referenda will accelerate, thus forcing citizens to vote on a barrage of hopeless proposals and to decide close votes over and over again.

Is there a better means of rationing referenda? Bonding offers an attractive alternative. According to this approach, supporters could place a proposition on the ballot by posting a money bond with the electoral commission. If the proposition performs well in the election, the bond will be returned. Conversely, if the proposition performs poorly in the election, the bond will be forfeited to the state. For example, in lieu of 100,000 signatures, supporters of an initiative might post $100,000, which they would forfeit unless the initiative won, say, at least 45 percent of the votes.

Compared to collecting signatures, bonding reduces the transaction costs of direct democracy. Compared to cheap collection of signatures by Internet, bonding discourages frivolous or previously defeated initiatives. By bonding ballot initiatives, constitutional law could reduce the velocity of direct democracy without stopping it or imposing unnecessary costs. Note that some countries, notably New
Zealand and the United Kingdom, already require candidates for Parliament to post a bond that they forfeit for poor performance in elections. Also note that a market for bonds would allow poor groups to mount an initiative for a popular issue.  

Optimal Breadth of Government

Citizens need an elected legislature so that their representatives can bargain together; as discussed earlier, increasing the breadth of government creates economies of scope by allowing vote trading among people with complementary tastes. But broad government can also cycle, or engage in redistributive contests. Under these circumstances, citizens would benefit from narrower government, which substitutes the median rule for bargains. Figure 4 shows social value increasing and then decreasing as the breadth of government increases. Unstable democracies are presumably to the left of the optimum, $x^\star$. To increase stability, these democracies should factor political issues. Beyond $x^\star$, however, too many elections cause the loss of too much surplus. At the optimum $x^\star$, the gain from more stability exactly offsets the loss from the foregone surplus as the number of elections increases.

III. Deepen or Devolve?

Now I turn from the problem of stability to the problem of corrupt administration. The usual analysis of corrupt administration focuses on the depth of political appointments in the state bureaucracy. Political appointees escape the rules imposed on civil servants to constrain corruption. As political appointments go deeper, more officials take bribes. Consequently, increasing civil service appointments at the expense of political appointments can reduce corruption in administration.

Instead of the usual analysis, I view corruption as depending on the number of governments. Under certain circumstances, reducing the height of the bureaucracy reduces corruption. I will explain the circumstances under which increasing the number of elections along the vertical dimension decreases political corruption.

I begin by summarizing the arguments that I will develop in detail. Centralized government deepens hierarchy by requiring administration to reach down to localities. As the chain of command lengths, officials can divert more resources away from their intended purposes. Deep hierarchy, consequently, promotes corruption. To reduce the diversion of resources, officials at the top impose rigid rules on administrators. Given the fast pace of social and economic change, however, rigid rules impose heavy costs. To reduce rigidity, a constitution can devolve power and shorten the chain of hierarchy. As power devolves in a democracy, however, the number of elections increases. Too many elections can drain the reservoir of civic spirit that animates voters, causing a decline in the quality of participation. The optimal number of elections balances the decline in the quality of participation against less diversion and more flexibility obtained by adding more elections. Now I will demonstrate these facts more formally.

Rule Game

I begin by depicting a simple agency game of administration to show how deeper hierarchy causes more resource diversion and less flexibility. After delegating responsibility to an agent, should the principal give the agent discretion or require the agent to follow a rule? Principals impose rules on agents for a variety of reasons, such as reducing transaction costs, improving coordination, increasing predictability, reducing disparity, and facilitating transparency. Instead of discussing many reasons, I will reduce the problem of imposing rules to its simplest elements and analyze one fundamental trade-off. Imposing rules
on agents reduces their opportunities to divert resources, whereas giving discretion to agents allows them to respond flexibly to changing circumstances. Diversion of resources is the cost of flexibility in an organization.

I will formulate the rule game in order to analyze the trade-off between diversion and flexibility. In the rule game, nature chooses a state and then the agent acts. Knowing nature’s state, the agent who enjoys discretionary power can respond flexibly to events as they develop. The principal wants the agent to reallocate resources when unexpected events occur, and the principal does not want the agent to divert resources when events occur as expected. Discretion gives the agent control over the decision, whereas a rule requires the agent to implement the principal’s plan under all circumstances. The principal must decide whether to give the agent discretion or impose a rule.

Figure 5 depicts the rule game concretely as a tree. First, the principal decides whether to give the agent discretion or impose a rule. Second, nature chooses a good or bad state. Third, if the agent has discretion, the agent decides whether to follow the principal’s plan or divert resources. Alternatively, if the principal imposes a rule, the agent must follow the rule regardless of the state of nature.

The payoffs from different paths in the game tree appear in parentheses at the right side of figure 5, with the principal’s payoff written first and the agent’s written second. Relative payoffs illustrate important facts, whereas absolute payoffs signify nothing. The principal’s plan is designed for a good state. If a good state materializes, the payoff to the principal is higher when the agent implements the principal’s plan (1) rather than diverting resources to an alternative project (.5). If a bad state materializes, however, the payoff to the principal is higher when the agent reallocates some resources to the alternative project (.5) instead of implementing the principal’s plan (0). So a loyal agent with discretion implements the principal’s plan in a good state and reallocates resources to an alternative project in a bad state.

The agent’s interests do not coincide perfectly with the principal’s. In a good state, the agent’s payoff is higher when he or she diverts resources to his or her preferred project (1.2) rather than implementing the principal’s plan (1). In a bad state, the agent’s payoff is also higher when he or she reallocates resources to his or her preferred project (.5) rather than implementing the principal’s plan (0). The agent’s dominant strategy is to divert resources, which serves the principal in a bad state and diserves the principal in a good state.

Now I turn from what the actors do to what they know. As in the delegation game, the rule game assumes that the principal who delegates a task to the agent knows the entire payoff matrix and observes his or her own payoff but does not observe the state of nature or the agent’s act. Figure 6 summarizes what the principal can infer from what he or she observes. When the payoff equals 1, the principal can infer from figure 6 both the state of nature (good) and the agent’s act (implement). Similarly, when the payoff equals 0, the principal can infer from figure 6 the state of nature (bad) and the agent’s act (implement). When the payoff equals .5, however, the principal cannot infer whether the agent’s reallocation was loyal (bad state) or disloyal (good state).

Solution

The rule game’s solution is a pair of strategies that maximize each player’s expected payoff, given the strategy of the other player. To
solve the game recursively, assume that the principal gives discretion to the agent. The last decision in time is the agent’s choice between implementing the principal’s policy or reallocating resources. As depicted in figure 6, the agent’s payoff from reallocating exceeds his or her payoff from implementing, regardless of the state of nature, so the agent has a dominant strategy.\textsuperscript{18} Knowing this, the principal computes his or her best strategy by assuming that the agent will use discretion to reallocate resources. The rule in question is designed for good states. As depicted in figure 5, imposing a rule on the agent yields a higher payoff to the principal in a good state, whereas giving discretion to the agent yields a higher payoff to the principal in a bad state. In this example, the rational principal imposes a rule when the probability of a good state, \( p \), exceeds .5, and, otherwise, the rational principal gives the agent discretion.\textsuperscript{19} The game’s solution can be summarized as follows:

\[
p \geq .5 \quad \rightarrow \quad \text{principal imposes rule, agent implements,}
\]

\[
p < .5 \quad \rightarrow \quad \text{principal gives agent discretion, agent diverts.}
\]

I mention in passing several more special assumptions in my formulation of the rule game. First, my “solution” solves the problem of delegating power for a given contract between the principal and agent. Computing the optimal contract for the principal and agent requires another formulation of the problem.\textsuperscript{20} Second, I computed the game’s solution when rationally self-interested actors play it once. In reality, the actors may repeat the game, which gives the agent more reason to cooperate. Third, I implicitly assumed that the principal cannot invest in monitoring the agent. In reality, monitoring increases the risk of punishment, which deters diversion by agents. Finally, I assume that agents are self-interested, whereas some agents may remain loyal due to moral commitment.

Graph

Figure 7 graphs the trade-off between diversion and flexibility characterized by the rule game. The horizontal axis in figure 7 represents constraint of the agent by rules, which increases when moving to the right. The rule of law implies that officials follow rules rather than exercising discretion. Consequently, the horizontal axis in figure 7 characterizes more constraint by rules as an increase in “legality.” Conversely, the horizontal axis in figure 7 also represents the agent’s discretionary power, which increases when moving to the left. In figure 7 legality and discretionary power are polar opposites.

The vertical axis in figure 7 depicts the principal’s marginal costs. Moving from left to right, the principal imposes more rules and allows less discretion to the agent, so diversion costs typically decrease and inflexibility costs typically increase at the margin.\textsuperscript{21} The intersection of the marginal cost curves in Figure 7 corresponds to the level of legality that minimizes the principal’s total costs.

The costs of inflexibility and diversion depend on the environment’s predictability. Good luck reduces the cost of inflexibility, so an increase in the probability of a good state causes the “inflexibility curve” to shift down in figure 7. Conversely, good luck increases the diversion of resources by agents, so an increase in \( p \) causes the “diversion” curve to shift up in figure 7. Combining these effects, an increase in the probability of good luck from \( p_{\text{low}} \) to \( p_{\text{high}} \) causes the principal’s preferred level of legality to shift up from \( L_{\text{low}}^* \) to \( L_{\text{high}}^* \) in figure 7.

In general, predictability makes rules more attractive to principals, whereas unpredictability makes discretionary power more necessary.
Examples

To illustrate the rule game, assume that the minister of health constructs a plan to maximize the number of kidney transplants. Implementation of the plan requires the work of an administrator and cooperation from the nurses. If the nurses cooperate, the minister's payoff (1) comes from the administrator implementing the plan. If the nurses resist, however, the minister's highest payoff is higher when, instead of implementing the plan (0), the administrator reallocates some funds to another program (.5). The minister must decide whether to impose rules that enforce the plan or give the administrator discretionary power.

The minister cannot observe the behavior of the nurses or the administrator. A high payoff (1) enables the minister to infer that the administrator implemented the plan and the nurses assisted, and a low payoff (0) enables the minister to infer that the administrator implemented the plan and the nurses resisted. In contrast, with an intermediate payoff (.5) the minister cannot infer whether the administrator reallocated funds in response to the nurses' resistance or diverted funds even though the nurses cooperated. If the nurses are more likely to cooperate than resist, the minister's payoff is higher from imposing the rule. Conversely, if the nurses are more likely to resist than cooperate, the minister's payoff is higher from giving discretion to the administrator.

A second example concerns procurement by the state. In many state universities, a professor who wants to purchase a computer has to follow prescribed procedures that constrain the choice of sellers and the terms of the contract. Procurement rules typically reduce the purchaser's discretion in order to avoid kickbacks or bribes.

A third example concerns challenges to the legality of actions by state agencies. Assume that the court interprets a statute and imposes a rule on a state agency. Individuals harmed by departures from the rule have the right to sue the agency, thus alerting the court concerning the agency's misbehavior. To illustrate concretely, federal courts interpreted the U.S. Constitution as requiring the police to recite a list of procedural rights when charging a person with a crime ("Miranda warnings"). If police obtain evidence about a crime by failing to recite these procedural rights, the courts exclude the illegally obtained evidence from trial. Like all rules, the procedures do not fit every case. Even so, the courts apparently prefer to prescribe the rules for all cases rather than giving discretion to the police.

The Optimal Number ofGovernments for Economic Development

Significance of the Rule Game

Having developed the model of rules, I next consider its significance. The Constitution or other fundamental law sometimes requires officials to make rules and follow them. The rule game predicts some consequences of the constraints of legality. Requiring more legality than the principal prefers imposes costs on him or her. Specifically, the principal loses to the extent that the cost of agent's inflexibility exceeds the reduction in diversion costs. The magnitude of the principal's loss depends on the environment's predictability. The harm is greater when the environment becomes less predictable and bad luck becomes more probable.

Figure 7 illustrates these facts. To be concrete, assume that the probability of good luck equals $p_{low}$, so the principal prefers $L^*_{low}$. Now assume that the principal is forced to increase legality to $L_{max}$. The resulting loss to the principal equals the amount by which the cost of inflexibility exceeds the marginal cost of diversion in the interval $[L^*_{low}, L_{max}]$, as indicated by the area $A + B + C + D + E$ in figure 7. If the probability of a good state rises from $p_{low}$ to $p_{high}$ in figure 7, the principal's loss from a requirement of maximum legality $L_{max}$ shrinks from the area ABCDE to the area A.

Politics, Civil Service, and Courts

In many state bureaucracies, politicians occupy the top offices and civil servants occupy the subordinate offices. To illustrate, the U.S. president appoints the head of most agencies, each head chooses a personal staff, and the civil service fills most jobs below the head's personal staff. Alternatively, political appointment can go deep into administration. In a patronage system, the winners in the game of politics distribute state jobs to loyal followers as the spoils of victory. To illustrate, patronage operates deep in administration in the city of Chicago and many developing countries.

Administration by civil servants suffers from inflexibility, whereas administration by political appointees suffers from corruption. The best system apparently provides for political appointment at the top level in the bureaucracy and civil service control below the top. The rule game can explain why patronage produces more efficient government at high levels of administration, and civil service rules produce more efficient government at low levels of administration.
Think of the state as a chain of relationships in which each official is an agent relative to those above him or her. In the typical state bureaucracy, civil servants are agents relative to the political appointees heading the organization, political appointees heading the organization are agents relative to elected officials, and elected officials are agents relative to the citizens who vote. In each of the chain’s links, a combination of discretion and legality orders the relationship with the agent. Now I explain why efficiency requires discretion to dominate legality at the top of the chain and efficiency requires legality to dominate discretion at the bottom of the chain.

The closer they are to the top of the chain, the more citizens know about officials. In U.S. foreign affairs, the communications media scrutinize the president, monitor the secretary of state, occasionally notice an ambassador, and mostly ignore civil servants in the State Department. When the principal has more information, the agent has less scope for undetected diversion of resources. In terms of figure 7, more information for the principal causes diversion costs to rise more slowly, as the agent receives more discretion.

While voters have relatively good information about top officials, the environment of high politics is also unpredictable. In figure 7, low predictability increases the costs of inflexibility. To illustrate, unpredictable diplomatic crises require a flexible response by the secretary of state.

Extensive monitoring and an unpredictable environment tip the balance in favor of giving broad discretion to officials at the top of agencies. Broad discretion requires politics, not the civil service. Rather than imposing rules, voters communicate goals to top officials. So efficient administration in a democracy requires political control over top officials in state agencies.

Conversely, the public cannot scrutinize lower levels of administration. Consequently, the public holds top officials responsible for any diversion of resources detected in the lower levels of administration. To discharge their responsibility, high officials impose rules to reduce diversion by low officials. In terms of figure 7, less information for the principal causes diversion costs to rise more quickly as the agent receives more discretion. Rules constrain such abuses. So efficiency in a democracy requires civil service rules to control employment at less visible levels of administration. (High officials also have other reasons to make rules for a complex bureaucracy.)

The problem of monitoring also arises in a judicial hierarchy. When faced with disputes, courts sometimes can choose between deciding each case on its own merits or developing general rules that apply to all cases. Case by case adjudication retains flexibility for lower courts and permits them to diverge from the preferences of higher courts. In contrast, rules reduce flexibility in lower courts and compel lower courts to conform more to the preferences of higher courts.

My discussion of politics, administration, and courts suggests three vague boundaries that demarcate significant changes in discretionary power. First, officials enjoy strong discretion when the law leaves them free to pursue political goals. To illustrate, legislators have strong discretion in proposing legislation, and the executive has strong discretion when selecting a cabinet. Second, officials enjoy weak discretion when the law prescribes goals and leaves officials free to choose the means. To illustrate, a civil engineer in the ministry of roads can decide how to build a road required by an executive order, and the ministry of education can design a program to improve literacy as prescribed by legislation. Third, pure legality leaves officials without any discretion, which results in mechanical decision making. To illustrate, a table that prescribes an exact punishment for each crime or the exact division of assets on divorce leaves little discretion to judges.

Legislators and the executive typically have political discretion, and civil servants typically have technical discretion. The situation of judges is more complicated. Common law systems give judges discretion to make some kinds of law, whereas civil law systems sometimes aspire to eliminate the discretionary power of judges. Philosophers of law disagree about the ideal mix of politics, technique, and legality in judging. In any case, pure legality, or the mechanical application of law, fails for most decisions. British unions periodically paralyzed the railroads by means of a tactic called “work to rule,” which means that the workers implemented all rules literally. Like the railroads, courts that apply rules mechanically cannot do justice.

Voting

I have explained that a deeper hierarchy better disguises the diversion of resources, whereas more elections and a more shallow hierarchy lowers the cost of monitoring by citizens. Thus, more elections can reduce corruption and increase flexibility. Now I want to explain how holding too many elections dilutes the civic spirit required for effective elections. Citizens must decide whether to vote and how to
vote. First, I will analyze how a rational citizen will vote, and, second, I will analyze whether a rational citizen will vote.

Self-Interest or Public Interest?

I like ice cream better than cabbage because of the taste, he likes San Diego better than Seattle because of the weather, and she likes the Republicans better than the Democrats because she is conservative. Among the many reasons that people have for their preferences, I will contrast two broad types. On the one hand, a citizen can vote based on material self-interest. A narrowly self-interested voter asks, "Which outcome will do more to increase my own wealth and power?" On the other hand, a public-spirited voter asks, "Which outcome will benefit the country more according to my political philosophy?"

To supply efficient quantities of public goods, officials need information about the policy preferences of citizens. By supplying this information, self-interested voting sometimes promotes efficiency in the supply of public goods. All too often, however, citizens use politics to obtain advantages for themselves at the expense of others. Banks want loan guarantees, farmers want price supports, unions want tariffs, artists want subsidies, taxis want fewer licenses for cabs, the elderly want property tax exemptions, and so forth. This kind of self-seeking wastes resources and oppress the powerless. While people seldom criticize a consumer in the grocery store for following self-interest when filling his or her shopping cart, people often criticize citizens for voting their self-interest.

Do most citizens vote their self-interest or public interest? The determinants of voting behavior have been studied for many years. Survey research reveals that voters know little about issues or candidates, so they typically rely on guidance from political parties, ideology, and informed friends or associates. In spite of their ignorance, however, citizens tend to vote for candidates who promote the interests of the groups to which they belong. For example, farmers tend to vote for candidates who subsidize agriculture, ethnic groups tend to vote for candidates who benefit minorities, and investment bankers tend to vote for candidates who liberalize finance.

Supporting candidates who advance a group's interests can benefit a person by showing solidarity with its members. To illustrate, dairy farmers in a rural community may be more willing to cooperate with other dairy farmers who endorse milk subsidies. Conversely, an ethnic group may censure members who oppose preferential treat-

ment for minorities. In general, groups develop ideologies that advance a self-serving conception of the public interest, like the automobile executive who believes that "what's good for General Motors is good for America."

Why Vote?

Journalists often deplore the fact that only about half of the eligible citizens vote in major U.S. elections and participation has fallen since the nineteenth century. Voter participation rates are similar in other countries, except where democracy is new; the law compels citizens to vote, as in Australia and Argentina; or the names of nonvoting citizens are posted in public, as in some Italian towns. Unlike journalists, however, economists find voter participation rates mysteriously high. Models of self-interest predict much lower voter participation rates than actually occur, and here is why. A self-interested citizen will decide whether or not to vote by comparing the cost of voting and the expected benefit. Given current rates of voter participation, the probability is historically negligible that a single voter in a large election will affect the outcome. So the effort required to vote exceeds the expected benefit for voters in large elections.

Some notation clarifies this point. The value of the time required to vote usually measures its opportunity cost, which I denote $C_i$ for citizen $i$. For simplicity, assume that the citizen cares about who wins the election, not the margin of victory. Let $p_i$ denote the probability that citizen $i$'s vote decides the election's outcome. Let $B_i$ denote the increase in citizen $i$'s wealth or power obtained by getting his or her preferred outcome in the election. Thus, the expected benefit from voting equals $p_iB_i$. According to the self-interested theory of voter participation, a citizen votes when $p_iB_i > C_i$ and a citizen does not vote when $p_iB_i < C_i$. The self-interested theory of voting predicts that voter participation rates will fall until $p_iB_i$ approximately equals $C_i$.

The paradox of voting refers to the fact that current levels of voter participation far exceed the rate at which $p_iB_i$ equals $C_i$. If the self-interested theory of voting accurately described the behavior of most citizens, voter participation rates would fall far below current levels.

To illustrate, assume that having your preferred candidate win the election is worth $1,000 to you. Assume that voting requires one hour of your time, which you value at $10. Self-interest prompts you to vote if $p_i \geq 100$, which implies that $p_i \geq 1/100$. In large elections, the probability of any one vote being decisive is much smaller than 1/100.
The Optimal Number of Governments for Economic Development

Fig. 8. Optimal height of administration

$b_{2i}$ denote the increase in the decisive voter’s benefit from knowing more about how to vote. According to the civic virtue theory, person $i$ gathers political information until $v_{2i} = c_{2i} - p_{2i}b_{2i}$.

Democracy is a great motivator. Perhaps civic virtue $v_1$ and $v_2$ increases with the initial growth of democracy. Increasing the number of elections, however, eventually begins to dilute civic virtue. When the total number of elections is large, the value $v_1$ or $v_2$ for any particular election decreases with a further increase in the number of elections. By these facts and the preceding formulas, too many elections results in a fall in voter participation and in voter information about each election.

**Optimal Height of Administration**

Too deep administration dilutes democratic purposes and gives too much discretion to administrators, in which case the constitution should replace *hierarchies* with *elected governments*. Conversely, too many elections can drain the reservoir of civic spirit that animates voters, in which case the Constitution should replace *governments* with *hierarchies*. Figure 8 depicts social value as increasing and then decreasing as the height of hierarchy increases. Corrupt democracies are presumably below the optimum $y^*$. To reduce corruption, these democracies should devolve power.
IV. Conclusion

Government augments the private economy by providing public goods and the legal framework for private markets. Instability and corruption undermine the ability of government to augment the economy. The usual analysis commends centralizing authority until the threat of dictatorship offsets the gain in stability and efficiency. In other words, the idea is to create a government strong enough to protect property rights and personal safety but so circumscribed that it does not indulge in public predation.

I propose a nonconventional solution to this problem. A democracy plagued with political instability should reduce the scope of unstructured bargaining among politicians and increase the scope of the median rule. To achieve this substitution, political issues should be factored. To factor, replace comprehensive government with single-purpose governments and substitute indirect democracy with direct democracy. Each single-purpose government and referendum requires a separate election, so the number of elections must increase.

Each link in the chain of a deep bureaucracy provides an opportunity to divert resources. Imposing rules on administrators reduces corruption at the cost of less flexibility. In a period of rapid social and economic change, strict rules meant to reduce corruption impose an especially high cost in terms of administrative inflexibility. Instead of more rules, a democracy plagued with corruption should reduce the depth of its bureaucracy. Each additional level of government in a federal system requires a separate election, so the number of elections must increase.

At the beginning of this essay, I explained that I would isolate and examine two variables—the horizontal and vertical number of elections—that affect the performance of government. My analysis omits consideration of many other variables. Under some circumstances, interaction with additional variables could nullify my conclusions. To illustrate, in some postcommunist countries without effective local elections, the devolution of power might increase corruption. Fiscal feudalism is a situation in which small localities insulate themselves from world trade by creating regulatory and tax barriers, often in order to sustain corrupt bureaucrats and hegemonic politicians. Instead of federalism, devolution might produce feudalism. In spite of this qualification, my model identifies real forces at work in political economy that conventional theories neglect. Determining the relative power of these forces requires empirical research combining laboratory experiments, econometric research comparing states, and case studies.

Notes

I would like to thank Martin McGuire (and other participants in the IRIS conference on market-augmenting government) for comments on the first draft of this essay. It draws upon Cooter 2000, pt. 2.

1. For example, contrast Schultz 1977 and Stigler 1975.
3. While the feasible points are probably discrete (you cannot hold half an election), the space is continuous.
4. This exercise requires a social welfare function. The philosophical questions posed by this controversial idea deflect from this essay’s point.
5. For example, see Tiebout 1956 and Romano 1987. I review these theories in chapter 5 of The Strategic Constitution (2000).
6. Since Riker 1962, the economic theory of political coalition formation focuses upon the “minimal winning set” or a similar idea. Another possibility is to focus on the most complementary coalition. The most complementary coalition maximizes the gains from trading votes. See chapter 3 of Cooter 2000.
7. Each member of a potential coalition may demand his or her marginal contribution to it as the price of joining. A member’s marginal contribution to the coalition may be computed as the fall in the coalition’s total value caused by the member quitting. Here I apply the shapley value of a coalition member (see Luce and Raiffa 1967, 249). With increasing returns to scale (superadditivity), however, cooperation does not create enough value for each member to receive the marginal product of membership, so paying the marginal product of membership to everyone is unfeasible. To illustrate concretely, consider a coalition formed by A and B that distributes the surplus equally between them: \((A,B,C) = ($50, $50, $0)\). If either member of the coalition were to leave it, the payoff to the coalition would fall from $100 to zero. By this logic, the marginal product of each of the two members of the coalition equals $100 but the total product of the coalition also equals $100. Consequently, paying $100 to each member of the coalition is unfeasible. Unfeasible demands may be credible. A threat by a member of a majority coalition is credible, according to one definition, if another coalition could satisfy the demand without worsening its own position. To illustrate by the preceding example, consider the coalition formed by A and B that distributes the surplus equally between them: \((A,B,C) = ($50, $50, $0)\). If B were to withdraw from the coalition, the coalition’s payoff would fall from $100 to zero. Noting this fact, assume that B demands a payoff of $75 to remain in the coalition. The threat is credible because B could leave the coalition and
form a new coalition with C, distributing the surplus \((A,B,C) = (0,75,25)\), which makes B and C better off. A, however, can make the same demand as B. So A and B can each make a credible demand for $75. Both demands cannot be satisfied because there is only $100 to distribute. So each demand is credible and both demands are unfeasible.


9. The crucial condition for this result is that each voter has single-peaked preferences. With single-peaked preferences, a voter’s satisfaction always increases when moving toward the voter’s most preferred point along the single dimension of choice. With multiple-peaked preferences, a voter’s satisfaction increases at some point when moving away from the voter’s most preferred point. To illustrate, some voters—call them YUPPIES—prefer a high level of expenditure on public schools, in which case they will send their children to public school, but if the level is not high they would prefer it to be low so they can send their children to private school and save on taxes. The worst alternative for the YUPPIES is a moderate level of expenditure on public schools. I review these results in chapter 2 of *The Strategic Constitution* (2000).

10. Majority rule counts voters, whereas cost-benefit analysis adds individual values. Counting voters gives the same result as adding individual values under the assumption of “strong symmetry.” Under strong symmetry, each nonmedian voter who gains from a change away from the median can be matched with at least one voter who loses, and the loser loses no less than the winner wins. In notation, let \(x^*_m\) denote the point most preferred by the median voter. Consider any alternative \(x^*\). Let \(J\) denote the set of individuals that (strongly) prefers \(x^*_m\) to \(x^*\) and let \(K\) denote the set of individuals that (weakly) prefers \(x^*\) to \(x^*_m\). By strong symmetry, for each \(k\) in \(K\) there exists a \(j\) in \(J\) such that \(u_i(x^*_k) - u_i(x^*) \geq u_k(x^*) - u_k(x^*_m)\). This fact implies

\[
\sum_{i \in J \cup K} u_i(x^*) \geq \sum_{i \in J \cup K} u_i(x^*_m).
\]

11. I implicitly assume additively separable utility functions for each group, so any group’s total utility equals the sum of its utility on each of the two issues.

12. Cost-benefit efficiency requires choosing the level of expenditures that maximizes the sum of net benefits, which occurs with high expenditures on schools and high expenditures on police.

13. Since the voters’ preferences form an intransitive cycle, any coalition formed simply by trading votes in figure 3 is dominated by another coalition (empty core). For example, a liberal-conservative coalition to obtain (high, high) is dominated by a liberal-moderate coalition to obtain (high, low), a liberal-moderate coalition to obtain (high, low) is dominated by a conservative-moderate coalition to obtain (low, low), and so on. Thus, the liberal-conservative coalition might not prove to be stable. To guarantee its stability, the parties would need the ability to make sidepayments. With sidepayments, the liberal-conservative coalition dominates other possible coalitions and no possible coalition dominates the liberal-conservative coalition.


15. The median rules can, however, fail when voting on a single dimension of choice when preferences are not single peaked. Furthermore, with factoring and single-peaked preferences, nonseparable utility functions in multidimensional space can destabilize the median rule on each separate dimension of choice.


17. Also note that people accused of crimes in the United States must post bail to avoid jail while awaiting trial. The person who appears for trial recovers the bail, whereas the person who fails to appear for trial forfeits the bail. In reality, most people borrow money for bail from a professional bail bondsman, who charges a rate based on his or her assessment of the risk. Similarly, with ballot initiatives a market should develop allowing supporters to borrow the bond. Lenders would charge low rates for promising ballot initiatives that carry low risk, and lenders would charge high rates for unpromising initiatives that carry high risk.

18. The following summarizes the agent’s payoffs.

<table>
<thead>
<tr>
<th>Agent’s Act</th>
<th>Implement</th>
<th>Reallocating</th>
</tr>
</thead>
<tbody>
<tr>
<td>State of Nature</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>good</td>
<td>0</td>
<td>.5</td>
</tr>
</tbody>
</table>

19. If \(p\) denotes the probability that the state of nature is good, imposing a rule and giving discretion to the agent yield the same expected payoff to the principal when \(p\) solves equation

\[
1p + 0(1 - p) = .5p + .5(1 - p).
\]

Solving this equation yields \(p = .5\), which is the tipping point.

20. In a general game of contracting, the parties could adjust the payoffs by making sidepayments, which could improve their incentives. To illustrate, if \(p < .5\), instead of retaining the contract resulting in the payoffs in figure 5, the principal and agent both prefer a contract in which the principal promises to pay the agent a bonus of .3 conditional on the agent receiving a payoff of 1. This contract, like any optimal contract, induces the agent to maximize the joint payoffs.

21. *Marginal* diversion costs typically decrease (and marginal inflexibility
costs typically increase) because the principal typically imposes rules first on those activities for which diversions cost most and inflexibility costs least.

22. As the state bureaucracy grows, regulatory agencies pose obstacles to citizens, who turn to elected officials for help. Providing help requires knowledge that increases by interacting with the state bureaucracy over many years (see Fiorina 1977). In doing such "casework" for constituents, the incumbent in the legislature has the advantage of experience over a challenger. Following the principle "The best guide to a maze is its architect," legislators have an incentive to create a bureaucratic maze so that voters reject challengers and rely on incumbents as guides. Thus, incumbent politicians sometimes seek an electoral advantage by increasing the complexity of administration faced by citizens and retaining control over it.

23. In common law systems, trial courts decide facts and appeals courts decide law. In these systems, case by case adjudication allows lower courts to control more outcomes by making them turn on facts. Consequently, general rules allow higher courts to control more outcomes by making them turn on law.

24. Ronald Dworkin has argued, for example, that each legal dispute has one right answer, thus suggesting that judges have little discretion (1977). These views evolved somewhat in Dworkin's Law's Empire (1986). Empirical studies often conclude that judges on high courts implement their own political philosophies (Brenner 1982).


27. For the dynamics of "ethnification," see Kuran 1998.

28. A bumper sticker observed on a pickup truck in Berkeley read "If God had intended us to vote, He would have given us candidates."

29. To illustrate, in a vote between a Republican and Democratic candidate the benefit \( B_i \) of a Republican voter \( i \) equals \( u'(x) - u'(x) \).

30. See the discussion in Hasen 1996. Using a different method of calculation, Romer (1996, 200) concludes that the probability of a tie in a U.S. presidential election in which 50 million people vote is approximately \( 10^{-4} \).


32. Let \( f(v,b) \) denote the density function representing the distribution of social value \( v \) and material benefit \( b \) among citizens. The total number of voters in an election, according to this theory, equals the sum of all the voters for whom \( v \) exceeds \( C - pB \), or

\[
\text{voter participation } = \int_{C-pB} f(v) dv.
\]

33. As with the self-interested theory, the mixed motive theory predicts that voter participation should increase when: the power of a vote \( p \) increases, the private material benefit \( B_i \) from winning the election increases, or the opportunity cost of voting \( C \) decreases. In addition, the mixed motive theory predicts that voter participation increases when the value of conforming to the social norm \( v \) increases. Increases in \( v \) might occur because more people internalize civic virtue, the social advantage from political participation increases, or the social cost of not voting decreases.

References


Comment

Martin C. McGuire

Regarding Robert Cooter’s essay and its special relevance in honoring the work of Mancur Olson, two events of the year 1956 are worthy of report. The first is a conversation I had with Mancur Olson concerning the distinctions between Joan Robinson’s Economics of Imperfect Competition and Edward Chamberlain’s Theory of Monopolistic Competition. Well do I remember Mancur’s earnest and passionate preoccupation with that debate; this was possibly the first I knew of his interest in the question of how the size of a group influences its effectiveness in providing itself with a collective good. The second event worth reporting is the publication of Charles Tiebout’s “A Pure Theory of Local Expenditures” (1956), which launched the gargantuan and flourishing literature on the theory of local public goods, local government jurisdictions, and clubs.

Both Tiebout jurisdictions and Buchanan clubs are based on three basic principles of government organization.

1. Each government jurisdiction should be sized so that average cost per person is minimized. This requires that $MC_N = TC/N$, where $N$ is the number of people in a jurisdiction, $TC$ is the total cost of provision, and $MC_N$ is the marginal cost of adding another person.

2. The population of each jurisdiction should be as homogeneous as possible with respect to tastes (demand prices, really) for the public good.

3. The amount, $X$, of public good supplied in each jurisdiction should satisfy the Samuelson efficiency condition $MC_x = MRS_j$, where $MC_x$ represents the incremental cost of an incremental increase in $X$ and $MRS_j$ represents the $j$th person’s marginal valuation of $X$ relative to money.

Central to the Tiebout-Buchanan prescription for government organization is the conclusion that jurisdictions should be as homoge-
neous as possible because each person by assumption can locate in only one jurisdiction. The crucial idea for normative design of a government or jurisdictional structure then becomes the trade-off between decreasing the number of jurisdictions and thus increasing size (so as to gain from economies of scale) and the increase in diversity that accompanies larger size and requires costly compromises among citizens in the provision of local public goods.

Into this picture, then, place Mancur Olson. Writing a decade later in “The Principle of Fiscal Equivalence” (1969), he called our attention to the existence, at that time, of over 75,000 jurisdictions and suggested that each should incorporate as citizen-voters (who collectively make local tax and expenditure decisions) only those individuals who are affected by the decisions. The political or constitutional extent of a jurisdiction, in other words, should be congruent with the population affected by its decisions. Olson recognized, of course, that this might well lead to a significant proliferation of jurisdictions. How can this be reconciled with the economies of scale addressed by Tiebout and Buchanan? Olson’s answer was simple: by allowing individuals to be included in more than one jurisdiction.

Using a concrete example drawn from Robert Cooter’s essay, table 1 indicates this difference between the Tiebout-Buchanan and Olson approaches. The table shows a society with four types of people: those with high demand for both schools and police, those with low demand for both, those with high demand for schools and low demand for police, and those with just the opposite tastes. Tiebout suggests as an ideal four jurisdictions—one for each police-school combination—with jurisdictions homogeneous and populated by one person each (one individual here stands for an efficient-sized group of identical individuals. I am ignoring the so-called “integer problem”). Olson, on the other hand, implicitly recommends four jurisdictions, each populated by two individuals with each individual located in two different jurisdictions.

<table>
<thead>
<tr>
<th>Preferences for Schools</th>
<th>Preferences for Police</th>
</tr>
</thead>
<tbody>
<tr>
<td>high 2</td>
<td>high 2</td>
</tr>
<tr>
<td>low 2</td>
<td>low 2</td>
</tr>
</tbody>
</table>

Which of these two camps does Cooter endorse? His essay can in fact be used most effectively to consolidate the two approaches. Consider his figure 1 (reproduced here in slightly modified form). The vertical dimension is Tiebout’s. Implicitly, at least, governments are multipurpose, providing an entire package of local public goods. Tiebout is concerned with movements up and down the y axis, assuming that each government is “unitary” in Cooter’s terms. As fewer and fewer governments are allowed, cities/counties/states/regions become larger and larger. More people are included because of cost savings from scale economies, but greater compromise is required because, as more people are included, the dispersion of preferences increases.

The horizontal dimension is Olson’s. Moving to the right, each person belongs to an increasing number of special jurisdictions. The lower right corner is the limiting case of one jurisdiction for each person for each good. The upper left is one monolithic, all-providing government, for which the term local is meaningless.
Notes

1. This took place in our University College, Oxford, Junior Common Room, in the spring term of that year.
2. In this case, the collective good is the price everyone gets to charge to the market consumers.

References