Washington University
ERes Cover Sheet

Article Title: CHAPTERS 5 & 7

Author: SHEPSLE & BONCHEK

Source Title: ANALYZING POLITICS

Vol.: _____ Issue: _____ Date: 1997 Pages: 82-136 166-191

FROM SHEPSLE & BONCHEK. 1997. ANALYZING POLITICS. NEW YORK: NORTON

5

Spatial Models of Majority Rule

Our story line to this point has emphasized a tradeoff in group decision making between the coherence of group choices, on the one hand, and the fairness of the method of decision making, on the other. If we consider a limited set of group decision-making circumstances, then we may be able to avoid the pain of this tradeoff. Put somewhat differently, if individual preferences happen to arrange themselves in particular ways—that reflect a consensus of a specific sort—then group decisions (certainly those made by majority rule) work out quite nicely. In the last chapter, we described single-peaked preferences as one kind of consensus that facilitated coherence in majority-rule decision making. In this chapter we want to give an intuitive geometric characterization of this condition.

Frankly, however, all this gets pretty boring pretty quickly. The authors, and perhaps some of the readers, may enjoy technical riffs and philosophical discourses, but most readers are more impatient and anxious to see some payoff. We think this chapter constitutes an important investment. Once we give single peakedness a geometric representation we will be able to apply it to some interesting political situations—namely, two-party electoral competition and legislative committee decision making.

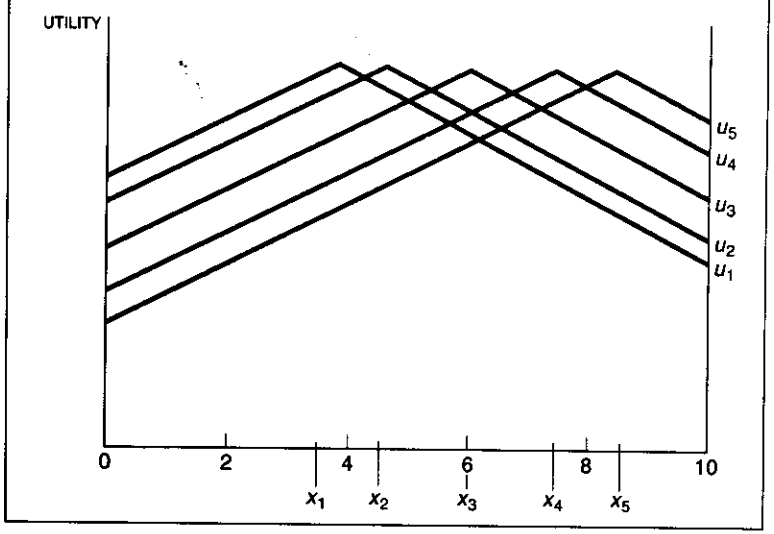
Spatial Formulation

The Simple Geometry of Majority Rule

Suppose a group's problem is, in effect, to pick a point on a line—the group must select some single numerical parameter. For example, a bank's board of directors must decide each week on the week's interest rate for 30-year home mortgages. In effect, the relevant interest rates are points on a line, one endpoint being 0 percent and the other being some positive number, say 10 percent. We write this interval as [0,10]. In this and other circumstances, we want the reader to imagine a group of individuals each of whom has a most-preferred point on the line, and preferences that decline as points further away in either direction are taken up.

In Figure 5.1 we display the preferences of the five-person board of bank directors, $G = \{1, 2, 3, 4, 5\}$. The board is meeting

Figure 5.1



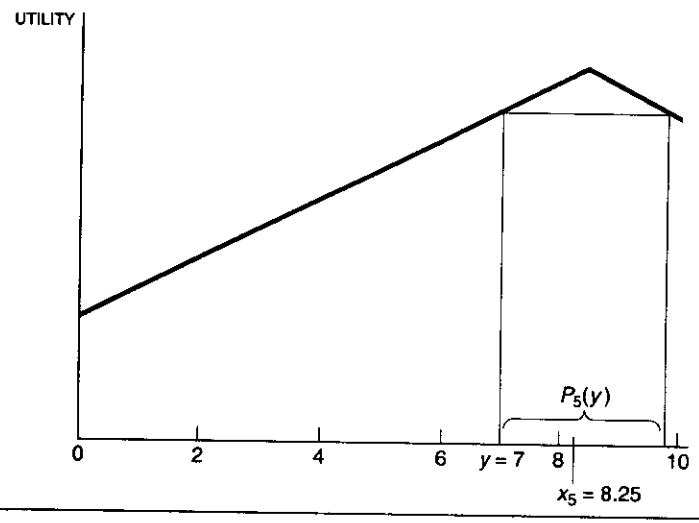
to decide the interest rate to charge for home mortgages this coming week. Each individual $i \in G$ has a most-preferred point (also called *bliss point* or *ideal point*), labeled x_i, located on the [0, 10] interval (drawn as the horizontal axis), representing his or her most-preferred interest rate.[1] Thus, director 1 has a most-preferred interest rate (x_1) of just less than 4 percent, director 2's (x_2) is just more than 4 percent, and so on. On the vertical axis we have written the label "utility" to measure preferences. For each individual we have graphed a *utility function* which represents the director's preferences for various interest rate levels in the [0, 10] interval. Naturally, the utility function, labeled u_i for Mr. or Ms. *i*, is highest for i's most-preferred alternative, x_i, and declines as more distant points are considered. Thus, Ms. 5 most prefers an interest rate a little higher than 8 percent, with her preference declining either for higher or lower rates. For obvious reasons (just look at the graphs) we say that the preferences of these individuals are *single-peaked*, which we define as follows:

Single-Peakedness Condition. The preferences of group members are said to be *single peaked* if the alternatives under consideration can be represented as points on a line, and each of the utility functions representing preferences over these alternatives has a maximum at some point on the line and slopes away from this maximum on either side.

Is there any connection between this definition of single peakedness and Black's definition given in the last chapter? That is, do utility functions with a single peak, as displayed in Figure 5.1, have anything to do with all voters agreeing that some alternative is "not worst"? You bet! Take any three interest rates displayed in Figure 5.1—say 3 percent, 5 percent, and 9 percent. It is pretty easy to see that 5 percent is not the

[1] Recall that "$i \in G$" means "the element i in the set G," where i stands for any one of the five bank directors.

FIGURE 5.2

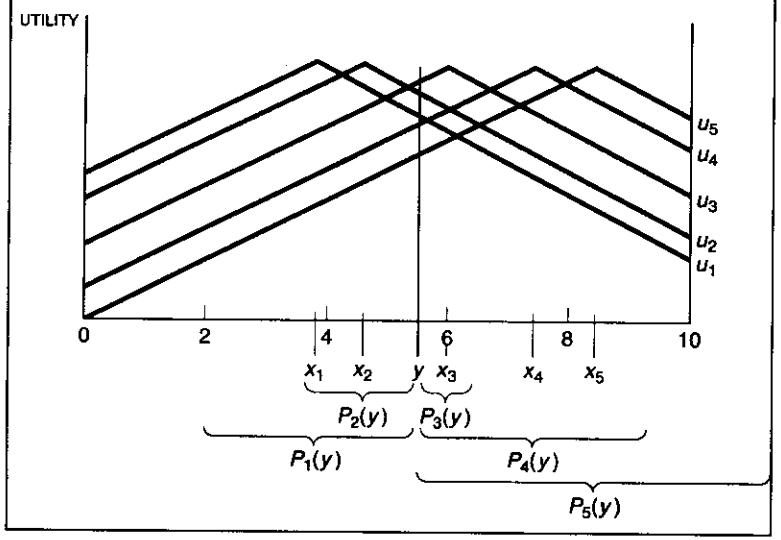


worst among these three rates for any of the five members of the group. Indeed, we claim that for any three interest rate levels the reader chooses, one of those is not worst for any of the five bankers. That's what single peakedness means!

In order to develop some tools that we will use in subsequent analysis, let's look at one of these individual bankers in isolation (by which we *really* mean let's look at an isolated utility function). In Figure 5.2 we show the most mean-spirited of the bank-board directors, Ms. 5, who most prefers a fairly high interest rate: $x_5 = 8.25\%$.[2] Consider an alternative rate, $y = 7\%$. The set of points Ms. 5 prefers to y is described by the set labeled $P_5(y)$ in Figure 5.2. This is Ms. 5's *preferred-to-y* set: if y were on offer, then $P_5(y)$ describes all the points she would prefer to it, given her preferences. As the figure shows, $P_5(y)$ is computed by determining the utility level for y and

[2] Obviously, everything is relative. This interest rate for a thirty-year fixed-rate mortgage was, for much of the early 1990s, a little on the high side. A decade earlier, when interest rates were in the mid-teens, a home buyer would have regarded a rate of 8.25 percent as a godsend.

FIGURE 5.3′



then identifying all the interest rates on the horizontal axis with utility levels greater than the utility for y.[3]

In Figure 5.3 we display the preferred-to-y sets of all five bank directors (note that y, in this figure, is just below 6 percent). Notice that these sets overlap to some degree—there are points in common to $P_4(y)$ and $P_5(y)$, for example. This means that there are specific points that *both* Mr. 4 and Ms. 5 prefer to y.[4]

Of great interest to us is the set of points a *majority* prefers to y. This is called the *winset of* y, written as $W(y)$. We define it as follows. Let M be the set of majorities in our group of

[3] We will include the endpoint of this set in the preferred-to set, even though, technically speaking, the end point is an alternative that the group member ranks at the *same* utility level as y.

[4] In set-theoretic notation, we can write these common points as the *intersection* of the two preferred-to-y sets: $P_4(y) \cap P_5(y)$. ($\cap$ is the intersection symbol, so that $A \cap B$ means "the points in both set A and set B.")

DISPLAY 5.1
THE MAJORITY COALITIONS OF G={1,2,3,4,5}

Size of Coalition	Coalitions
3	{1,2,3}, {1,2,4}, {1,2,5} {1,3,4}, {1,3,5}, {1,4,5} {2,3,4}, {2,3,5}, {2,4,5} {3,4,5}
4	{1,2,3,4}, {1,2,3,5}, {1,2,4,5} {1,3,4,5}, {2,3,4,5}
5	{1,2,3,4,5}

bankers, G; it is the collection of three-person coalitions (there are ten such coalitions), four-person coalitions (there are five of these), and the coalition-of-the-whole. So, there are sixteen different majority coalitions in M; they are listed in Display 5.1. For each majority coalition, consider the common intersection of preferred-to-y sets (if there is any); these are the points that this particular majority prefers to y. Thus, the members of the majority {3,4,5} in Figure 5.3 share points each prefers to y. Take the union of these sixteen sets. This is $W(y)$.[5]

It is now rather straightforward to describe the coherent choices of groups. If some alternative, x, has an empty winset (written: $W(x) = \varnothing$, where $\varnothing$ means "empty" in set notation), then it is a clear candidate for the group choice. Why? Simply because $W(x) = \varnothing$ means there is no other alternative that any of the sixteen majority coalitions prefers to x. It's hard to deny choosing x if there is nothing any majority agrees on in its place. On the other hand, if the winset of x is not empty ($W(x) \neq \varnothing$), then it is hard to justify the choice of x. How can you

[5] In Figure 5.3 it turns out that members of only *one* of the sixteen majorities, {3,4,5}, has overlapping $P_i(y)$ sets. Members of the remaining fifteen majorities (listed in Display 5.1) cannot agree on any points they jointly prefer to y. Thus $W(y) = P_3(y) \cap P_4(y) \cap P_5(y)$.

choose x when some majority of the group clearly wants some other specific alternative? And, if the winset is nonempty for *every* alternative, then we have a problem: the group's preferences are incoherent since some majority prefers something to every alternative available.

The question of the moment is whether, or in what circumstances, an x possessing an empty winset exists. If any complete and transitive preferences may be held by the individuals in G—Arrow's "universal domain" condition—then, as we have seen, the answer is "not necessarily." Why? Because under Arrow's condition U, it is possible for majority preferences to cycle, in which case $W(x) = \varnothing$ for *no* alternative. But, if preferences are restricted then a different answer is possible.

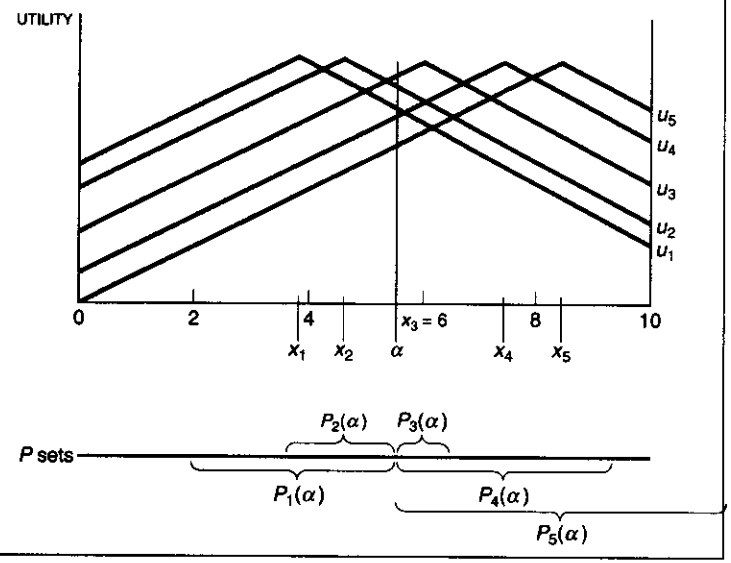
Black's Median Voter Theorem. If members of group G have single-peaked preferences, then the ideal point of the median voter has an empty winset.

One such group consisting of individuals with single-peaked preferences is pictured in Figure 5.3. The median voter ideal point in this group is x_3 of Mr. 3.[6] The claim of Black's Theorem (the same Duncan Black, by the way, as in the last chapter) is that $W(x_3) = \varnothing$, and that x_3 is the majority choice.

We can prove this theorem using the example of the five bank board members. Consider any arbitrary point in the feasible set of interest rates, [0, 10], to the left of x_3—say the point labeled α in Figure 5.3'. Notice that α is preferred to x_3 by members 1 and 2, but that x_3 is preferred to α by members 3, 4 and 5. Thus, x_3 is majority-preferred to α. But α is any arbitrary point to the left of x_3. For any such point we know at the very least that member 3, 4, and 5 will prefer x_3 to it. (It

[6] The median of a set ordered from left to right is the point such that at least half the points are at or to its right and at least half the points are at or to its left.

FIGURE 5.3'



is possible that some of the remaining members will share this preference, too.) Next, consider any arbitrary point to the right of x_3 (not pictured). Members 4 and 5 may prefer it to x_3, but members 1, 2, and 3 hold the opposite preference, so that x_3 is majority preferred. The argument is exactly the same as with α above, since we selected an arbitrary alternative to the right of x_3. To sum up, then, we now know that the ideal point of the median voter is preferred by a majority to *any* arbitrary point to the right or to the left of it, that is, to all remaining points. Hence, it has an empty winset and is the majority choice.

Before complicating this key result, let us mention that there are three hidden assumptions, and probably more besides, that warrant some discussion. First, in our example we proceeded with the group G that is *odd* in number. Thus, in Figure 5.3 we could display the five group members, with

member 3 the unique median. What if the size of the group were *even*? Suppose, for instance, that we ignored Ms. 5 in the figure and focused instead on the truncated group $G' = \{1, 2, 3, 4\}$. Now members 2 and 3 are both medians. Moreover, since it takes three votes to constitute a majority, it is true that $W(x_2) = \varnothing$ and that $W(x_3) = \varnothing$, too. Indeed, the winset of *any* point in the interval between the two, $[x_2, x_3]$, is empty. Technically, then, Black's Median Voter Theorem is true whether the group size is odd or even. But when a group is even in number, the fact that there may be *tie votes* means there may be more than one alternative with the property that it can't be beaten. This absence of a unique winner is a pain in the neck. It is a bit like more than one pretender to the crown, or more than one person claiming to be king of the mountain. It is for this reason that groups establish some procedure for breaking ties *well in advance* of any substantive deliberations or, better yet, they make sure that the group is odd in number.[7]

Second, we have assumed *full participation*. Everyone with the franchise is assumed to exercise it. Of course, in any particular instance of group choice this need not happen. If bank board members 4 and 5 oversleep one week, then Ms. 2 becomes the median voter of the now reduced three-person board; if members 1 and 3 are out of town the following week, then Mr. 4 becomes the median. In each of these cases, as well as in the case of the full board, the median voter result applies. *Who* the median is, however, depends upon who the participants in the group are. We may forecast the group decision if we make *assumptions* about participation (for example: assume everyone votes), or we may make forecasts *contingent* on participation (for example, if 4 and 5 oversleep, then we predict x_2 will be the group's choice; but if everyone votes, then x_3 is the predicted choice).

[7] For example, the U.S. Constitution *requires* the Senate to be even-numbered, but establishes a tiebreaking procedure. The Vice President of the United States, sitting as the president of the Senate, may (only) vote in case of a tie. Likewise, the standing rules of the House of Representatives provide a tiebreaking rule, asserting that a motion *fails* if it obtain no more yeas than nays—it fails on a tie.

Third, we have assumed that those exercising the franchise do so *sincerely*. But as we have seen in earlier chapters, group members will have occasion and incentive to misrepresent their preferences and not reveal them honestly. This is a subject of great interest that we take up on its own in Chapter 6.

The (Slightly) More Complicated Geometry of Majority Rule

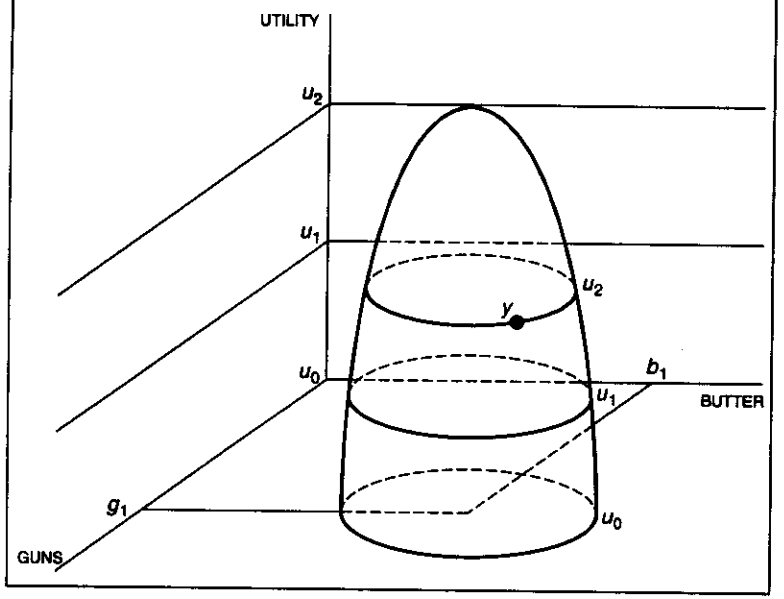
As we shall demonstrate in the next section, one-dimensional models of choice with the single-peakedness condition permit rather sophisticated ways to think about real politics. They generate very crisp expectations about how politics in these settings gets played out. But so many social situations cannot be reduced to one-dimensional affairs.

Recall the game of "divide the dollars." If the game were played by a group of three individuals, then it is necessary to have *two* dimensions in which to represent outcomes. The first dimension gives the amount that player 1 receives, while the second dimension gives the amount that player 2 receives. (Subtract the sum of these two numbers from the total number of dollars to be divided and you get the amount that player 3 receives.[8]) We hope we convinced the reader earlier that games of division, like divide-the-dollars, are commonplace in political life. So, it must be conceded that, as crisp and as sophisticated as the one-dimensional models are, they are *special cases* of a more general multidimensional arrangement. We need to see what this more general arrangement is like.

We can say most of what we need to by focusing on a two-

[8] Generally, when dividing a fixed pie among n categories (or people), we need only n-1 dimensions to display all outcomes.

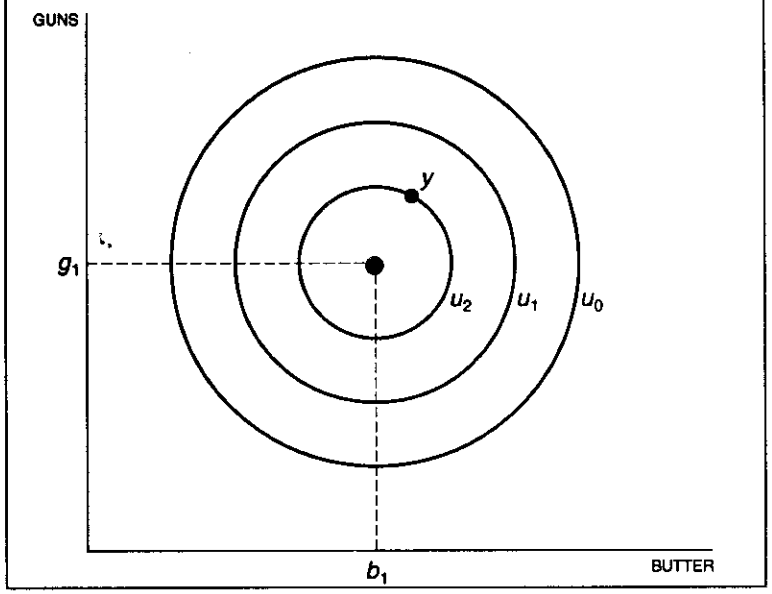
FIGURE 5.4



dimensional circumstance, like that pictured in Figure 5.4. (There are actually *three* dimensions in this figure, but this will be clarified shortly.) We consider a problem in budgeting, in which a group of legislators, perhaps an appropriations committee, must decide how to divide expenditures between "guns" and "butter" (symbolizing the competition between defense and other domestic programs). Outcomes, then, are described by two numbers: dollars spent on butter and dollars spent on guns. The set of outcomes, or simply the "policy space," is two-dimensional, and this is the domain over which preferences are expressed.[9] The third dimension of Figure 5.4,

[9] If the size of the budget were fixed in advance, then by the argument in the previous paragraph and footnote, we would need only *one* dimension, say dollars for defense; once this is established, dollars for domestic programs is strictly determined—it's what is left over. When the budget is *not* determined in advance, then we need *both* dimensions to display all outcomes.

FIGURE 5.5



marked "utility," permits us to draw three-dimensional graphs of legislator preferences. As earlier, a legislator is assumed to have an ideal point in the policy space. His preference function, or utility function, is at a maximum over this point. We assume further (in most of the applications we pursue) that preferences decline with "distance" from the legislator's ideal point. A typical legislator, with typical ideal point and preference function, is displayed in Figure 5.4. The legislator's preference function is a "hump" that reaches its highest utility level just over his ideal point in which b_1 dollars are spent on butter and g_1 dollars on guns. This ideal point, (b_1, g_1), is located in the plane of the butter-guns policy space.

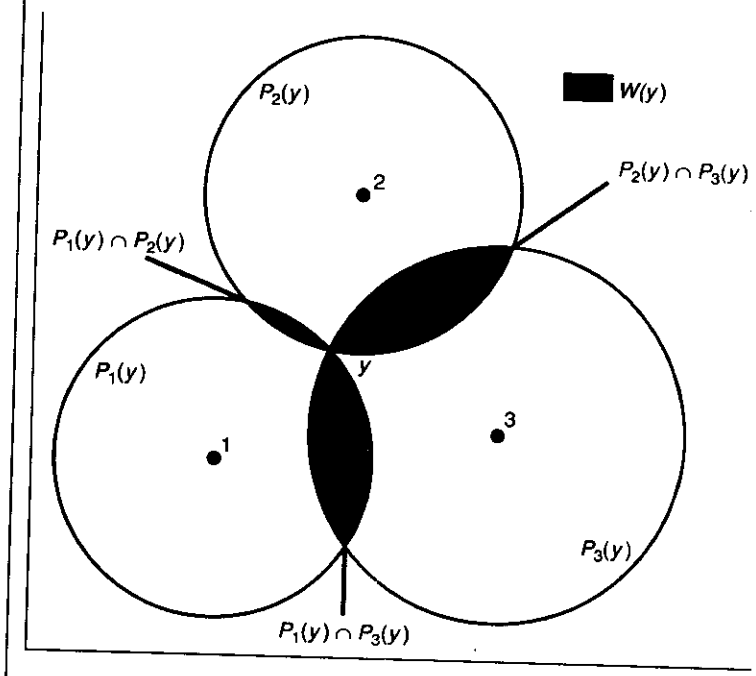
A more convenient way to represent precisely this same information, however, is given in Figure 5.5. In this figure the reader is looking down directly onto the plane of the butter-guns policy space. It is as though you are hovering in a heli-

copter above the peak of the preference hump in Figure 5.4. The location of our typical legislator's ideal point is exactly the same as in Figure 5.4. But instead of adding a third dimension (coming out of the page toward you) in order to graph his preference function, we instead overlay "slices" of his utility function onto the policy space, producing the set of nested circles called *indifference curves*. Each circle is a slice of the policy hump in Figure 5.4. It is a locus of policy outcomes among which the legislator is indifferent (since all the points on a circle lie on the same slice and hence at the same height on the utility function of Figure 5.4). Since distance from an ideal point is a measure of preference, points on a circle centered on her ideal, being equidistant from that ideal, are equally preferred by her. The logic is the same in comparing a point on one circle to that on another. A legislator prefers a point on a circle with a *smaller* radius to one on a circle with a larger radius, because this means the former point is closer to her ideal than is the latter point.[10]

Notice the point labeled y in Figure 5.5. The circle through y centered on our legislator's ideal point, as we just determined above, contains all the points of legislator indifference to y. This means that all the points *inside* the circle, being closer to her ideal, are actually *preferred* by her to y. That

[10] In this simplest of multidimensional setups, in which the policy space is two-dimensional and preference is measured by distance, indifference curves will be circles centered on the legislator's ideal point. In more than two dimensions, the indifference "contours" will be spheres or (in four or more dimensions) hyperspheres. A second sort of complication, which applies in the (simplest) two-dimensional as well as higher dimensional situations, is to allow preferences to be related to distance, but in a more complicated way. One dimension of policy may be "more important" to a legislator than another dimension. Thus, movement away from her ideal point along one dimension will have a greater impact on preferences than an identical movement along the other dimension. Put differently, preference is said in this instance to decline with *weighted* distance from the ideal (where the weights reflect the salience of each dimension to the legislator). In this instance, indifference contours will no longer be circles, but will be *ellipses* instead. We will stick with the most basic formulation.

FIGURE 5.6

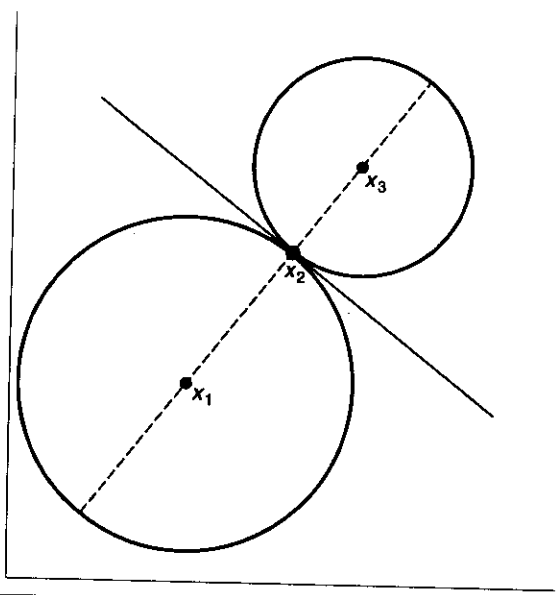


is, we can call the points inside the circle our legislator's *preferred-to-y* set, a natural generalization of that same concept in the one-dimensional development earlier in this chapter. Figure 5.6 displays three legislator ideal points and each legislator's indifference curve through y (the curve plus all points inside it comprising each legislator's preferred-to-y set,

labeled $P_i(y)$). The shaded intersection $P_1(y) \cap P_2(y)$ gives the points preferred by *both* legislators 1 and 2 to y; $P_1(y) \cap P_3(y)$ are the points preferred by 1 and 3 to y; and, finally, $P_2(y) \cap P_3(y)$ give those points for 2 and 3. Since two out of three is a majority, the union of these three "petals" is the winset of y, $W(y)$. Each petal gives the points that a specific majority coalition prefers to y.

If we move the ideal points of the three legislators in Figure 5.6 around, so that they line up in a row, we have a situation like that depicted in Figure 5.7. Can you determine the winset of the middle legislator's ideal, $W(x_2)$? The steps are laid out in that figure. Mr. 2's preferred-to-x_2 set is empty (how could he prefer anything to his most-preferred point?). Ms. 1's and Ms. 3's indifference curves through x_2 and centered on

FIGURE 5.7



their respective ideal points are *tangent* to one another. They do not overlap at all. Hence $W(x_2) = \varnothing$, since there are no points preferred to x_2 by majority {1,2}, {1,3}, {2,3}, or {1,2,3}. That is, the members of no majority coalition have preferred-to-x_2 sets that intersect. Thus, x_2 is the majority choice.

Another look at Figure 5.7 should show why this happened. Consider the bold line through x_2 perpendicular to the dashed line. On the x_1 side of this bold line, for any selected point off the dashed line, x_2 is closer to x_3 than the selected point is. So a majority, {2,3}, prefers x_2 to any such point. From a precisely parallel argument, a majority, {1,2}, prefers x_2 to any point off the dashed line on the x_3 side of the bold line. So, the only points that remain are those *on* the dashed line. That is, even though the group choice problem is *actually* two-dimensional, individual preferences line up so as to make the problem *effectively* one-dimensional. On this line individual legislators have single-peaked preferences (the reader should convince herself of this), with x_2 the median ideal point. Hence, Black's Median Voter Theorem applies, which is precisely what Figure 5.7 demonstrates.

Having the legislator ideal points line up is pretty convenient, isn't it? Pretty unlikely, too. Certainly it seems more unlikely than arbitrary configurations such as that in Figure 5.6. Thus, while we have tools like Black's Median Voter Theorem with which to analyze majority rule in one-dimensional settings, it is probably fair to say that many interesting political circumstances are genuinely multidimensional. Can we say anything about the prospects for a majority choice in multiple dimensions? The answer is yes, but the news is not very good.

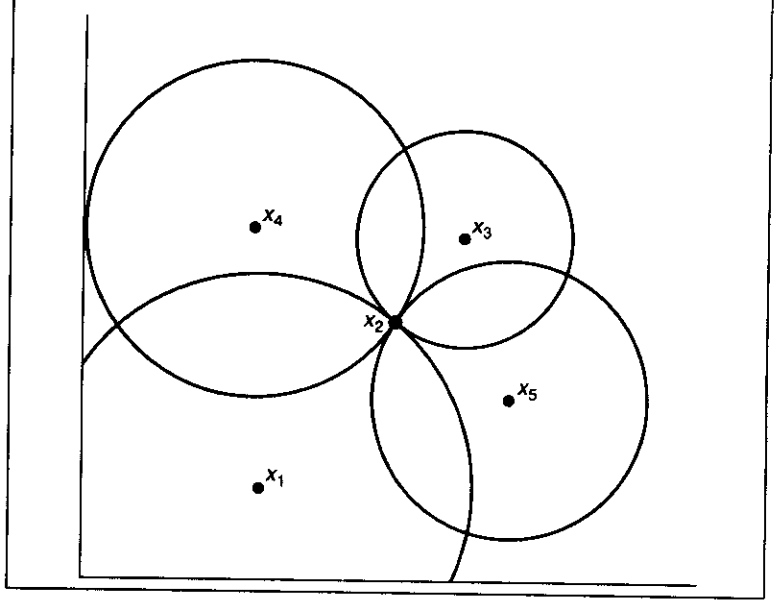
The highly unlikely distribution of individual preferences in Figure 5.7 provides a basis for generalization. What allows the ideal point of Mr. 2 to emerge as the majority choice is the fact that the ideals of the others are "symmetrically" distributed about 2's ideal. From Mr. 2's ideal any movement

away from it is obviously opposed by Mr. 2 himself; but it is also always opposed by at least one of the other guys. In fact, as we saw when considering the bold line through x_2 perpendicular to the dashed line containing all three ideal points, any point on Ms. 3's side of that line is less preferred than x_2 by 1 and 2, and any point on Ms. 1's side of the line is less preferred than x_2 by 2 and 3.

Now let's add two more voters to the picture that are symmetrical in precisely the same way (Figure 5.8). Voters 2, 4, and 5 lie on a line, just as 1, 2, and 3 do. It is still the case that $W(x_2) = \varnothing$, because any departure from x_2 is opposed by at least three of the five voters. You may test this proposition out for yourself by laying a straight edge through x_2 at any angle. There are always two voters who would like to move to some point on one side of the straight edge, two who would like to move to points on the other side of the straight edge, and one (Mr. 2) perfectly content to stay at x_2. Since no majority favors moving in any direction (there are always three votes against), the winset of x_2 is empty. Something about distributing voters symmetrically around a common point seems to be producing a coherent majority choice.

Indeed, we can be very specific here. Let us consider a set of n voters (where n is any number, which we will take to be odd to simplify the presentation), whose ideal points are $x_1, x_2, \ldots, x_n$. These n ideal points are in a multidimensional policy space, like the one pictured in Figure 5.8 (although the results we present below apply to policy spaces of more than two dimensions, as well). These ideal points are distributed in a *radially symmetric* fashion if the following conditions hold: (1) There is a distinguished ideal point, labeled x^*; (2) the $n-1$ remaining ideal points can be divided into pairs (since n is odd, $n-1$ is even and this is possible); and (3) the two ideal points in any pair, say x_i and x_j, plus x^* all lie on a line with x^* "between" x_i and x_j. In Figure 5.8, x_2 is the distinguished point, x_1–x_3 and x_4–x_5 are the pairs of remaining ideal points, and x_2

FIGURE 5.8



lies on a line "between" the ideal points in each pair. Notice that radial symmetry does not require the two ideal points of a pair to be equidistant from the distinguished point (x_3 is closer to x_2 than x_1 is); they must simply line up.

The economist Charles Plott noticed that radial symmetry of ideal points captured in higher dimensions a property that single-peaked preferences possess in one dimensional policy spaces. In a famous paper in 1967,[11] he established the following result:

Plott's Theorem. If voters possess distance-based spatial preferences, and if their ideal points are dis-

[11] Charles R. Plott, "A Notion of Equilibrium and Its Possibility under Majority Rule," *American Economic Review* 57 (1967): 787-806.

tributed in a radially symmetric fashion with x^* the distinguished ideal point and the number of voters odd, then $W(x^*) = \varnothing$.

Plott's Theorem generalizes Black's Median Voter Theorem to more than one dimension.[12] It provides the conditions for the centripetal ("center-seeking") tendency we observed in the one-dimensional setting studied by Black. Thus, in Figure 5.8 displaying the five legislators whose ideal points satisfy radial symmetry, x_2 is the majority choice. It is preferred by a majority to any other point in the policy space.

That's the good news. The bad news is that Plott's Theorem does not provide very general conditions. Return to Figure 5.8 and, in your mind's eye, move x_4 a small distance to the northeast (say half an inch). Now lay your straight edge along a line through x_2 but just below x_5. Since x_3, x_4, and x_5 now all lie above this line, this means that there is a majority—{3,4,5}—that prefers moving away from x_2. Specifically, there are points on the northeast side of this straight edge through x_2 that voters 3, 4, and 5 prefer to it, so that $W(x_2) \neq \varnothing$. That is, Plott's condition of radial symmetry is highly sensitive to small perturbations of voter ideal point locations. This can be put in an especially dramatic form. Imagine the ideal points of 1,000,000 voters radially distributed around the ideal point of the 1,000,001th voter. So, $W(x_{1,000,001}) = \varnothing$, in accord with Plott's Theorem. Now suppose two new voters move into the community, *and their ideal points are not radial symmetric about* $x_{1,000,001}$. This small perturbation in the voting situation—after all, how much effect can the introduction of two new voters

[12] When we say that one result "generalizes" another, we mean that the latter is a special case of the former. Thus, Black's median voter result is the one-dimensional version of Plott's Theorem, in which the median voter's ideal is the distinguished point and pairs of voter ideal points, one from each side of the median, are distributed around it in a radially symmetric fashion.

have in a voting population of more than one million?—completely destroys the previous equilibrium.[13]

If departures from radial symmetry were relatively unusual events, then this sensitivity to ideal point distributions in Plott's Theorem would not really be bad news. But, as the reader may grasp intuitively, the requirement of radial symmetry is actually quite restrictive; one would not expect groups "naturally" to have their preferences distributed in so even a manner as this. So departures from this condition take on a greater significance. In what is one of the most remarkable theoretical statements in this entire field, Richard McKelvey demonstrated exactly how significant these departures from radial symmetry are.

McKelvey's Chaos Theorem.[14] In multidimensional spatial settings, except in the case of a rare distribution of ideal points (like radial symmetry) that hardly ever occurs naturally, there will be no majority rule empty-winset point. Instead there will be chaos—no Condorcet winner, anything can happen, and whoever controls the order of voting can determine the final outcome.

We started out by seeking ways to restrict Arrow's universal domain condition to see if there were narrower domains in which majority rule worked tolerably well. In one-dimensional choice situations, we saw that single-peakedness is sufficient. In multidimensional situations a radially symmetric distribution of ideal points is sufficient. But, small departures from the latter throw everything into chaos. No point is the "king

[13] The sensitivity is not quite so severe when the number of voters is even. In this case the distinguished point is *not* a voter ideal point. Some shifts in voter ideal points are possible without disturbing the empty winset property of this distinguished point.

[14] Richard D. McKelvey, "Intransitivities in Multidimensional Voting Models," *Journal of Economic Theory* 12 (1976): 472-482.

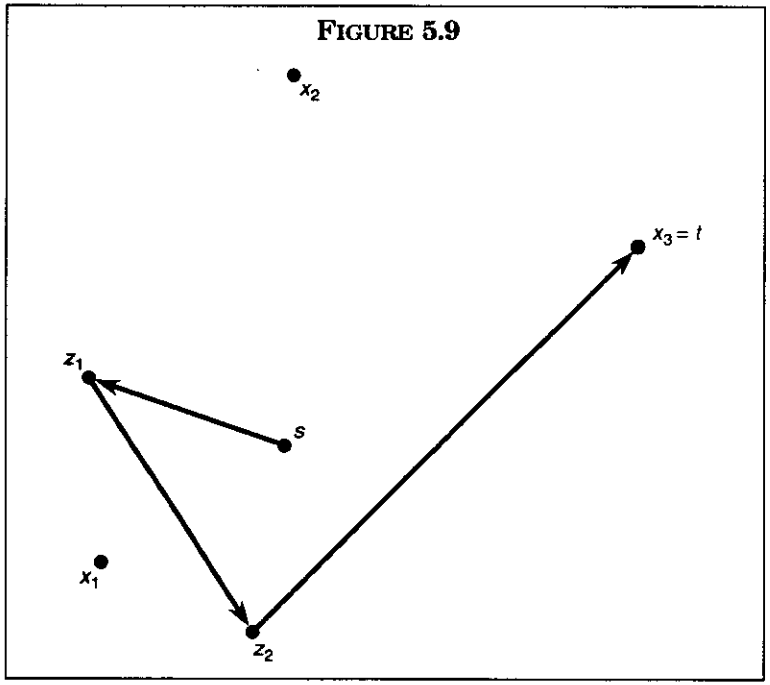
of the mountain" in the sense that it is preferred by a majority to all contenders, so it is difficult to justify any particular group choice (since for any proposed choice there is some alternative a majority prefers to it). This, in turn, means that there will always be majority cycles.

Indeed, McKelvey establishes that all the points are in one great big cycle. What this means, practically speaking, is that the situation is ripe for manipulation by whomever controls the agenda. What McKelvey shows is this: Pick any two points in the policy space—call them s (starting point) and t (terminating point). Then there is a sequence of points—$z_1, z_2, \ldots, z_k$ (for some finite number, k) such that $z_1 P_G s$, $z_2 P_G z_1$, $z_3 P_G z_2, \ldots, z_k P_G z_{k-1}$, and $t P_G z_k$. That is, from any starting point, there is a sequence of votes by which a majority will move the outcome to *any* terminal point (including, say, the ideal point of the agenda setter).

This is illustrated in Figure 5.9 for a three-person legislature. The ideal points of the three legislators are x_1, x_2, and x_3. The point s is the status quo ante. If Mr. 3 were the agenda setter empowered to make motions and order them in a voting sequence, then he could, in a small number of steps—in fact, in only *three* steps—drive the outcome to x_3, his ideal point. First he proposes z_1, which both Mr. 1 and Ms. 2 prefer to s. So $z_1 P_G s$. Then he proposes z_2 which both he and Mr. 1 prefer to z_1; so $z_2 P_G z_1$. Then, in the final step, he proposes his ideal point, x_3, which both he and Ms. 2 prefer to z_2. Voila! He has driven the legislative process, by artfully choosing the alternatives upon which to vote, to a terminal outcome located at his ideal policy: $t = x_3$.[15]

[15] It should be noted that the members of each majority coalition in this example blindly vote their preferences, like lambs following the judas goat to slaughter. The legislators seem like passive putty in the hands of the wily agenda setter, Mr. 3. In the next chapter, we will endow "followers" with some sophistication by which they might be able to control their "leader."

FIGURE 5.9



APPLICATIONS

Applications of the spatial model are so plentiful and rich that it is hard to know where to start. Since we do not have time to dally, and since we shall make subsequent use of the spatial model in a variety of contexts, we will content ourselves here with a fairly limited set of applications. We begin at the beginning, so to speak, with Downs's model of electoral competition. Anthony Downs was one of the first scholars to use the spatial model for political analysis. This application also demonstrates both the strengths and weaknesses of the simplify-

ing assumption that the political world can be modeled as one-dimensional. Then we will turn our spotlight on institutional analysis, looking at both a one-dimensional and multidimensional analysis of legislative politics.

Spatial Elections

The real origins of the spatial model are found in a famous paper written in 1929 by Harold Hotelling.[16] An economist interested in the locational decisions made by firms, Hotelling was especially fascinated by the stylized fact—true then, and still true today—that competitor firms regularly locate their retail shops next door to or just across the street from one another. Gasoline stations are found on opposing corners of an intersection, a pair of major department stores "anchor" a suburban shopping mall, and, in small-town America in the good 'ol days, competing "five and dime" stores like Woolworth's and Kresge's located just opposite one another. Why would nominal competitors, who have a great big geographic market to divide up between themselves, locate in such close proximity?[17]

We leave the economic location question to one side, but not without noting that Hotelling mused that political parties seemed to behave in much the same fashion as economic competitors. This musing became the major focus of the now-classic study by Anthony Downs, *An Economic Theory of Democracy*, where he gave the "spatial model of electoral competition" its fullest development and exposure.[18]

The "spatial" part of Downs's spatial model consists of a one-dimensional ideological continuum, [0,100]. The continuum is scaled by the proportion of economic activity left in the

[16] "Stability in Competition," *Economic Journal* 39 (1929): 41-57.

[17] Even more spectacular in many cities is a small stretch of a major highway along which dozens of automobile dealerships locate (in Boston, for example, "automile" holds more than thirty competitors.)

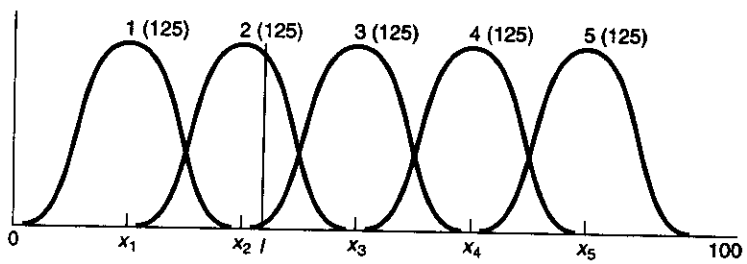
[18] Anthony Downs, *An Economic Theory of Democracy* (New York: Harper and Row, 1957).

hands of the private sector, so that the left endpoint reflects a fully socialized economy while the right endpoint is identified with a totally private-enterprise economy. While political competition in real life consists of taking positions on and articulating visions about a host of political issues, Downs supposes that, when all is said and done, political debate boils down to ideology—do you want some good, service, or purpose provided by government or by the private sector? Political competition, then, is a contest between politicians intent upon capturing control of government by appealing to voters with offers of alternative plans, platforms, programs—indeed visions. These appeals are identified with different points on the left-right ideological continuum.

As a first approximation for the hurly-burly of campaigning, electioneering, and voting, this is not a bad one. Politicians are conceived of as single-minded seekers of election. They are graduates, so to speak, of the Vince Lombardi School of Politics, whose motto is, "Winning isn't everything; it's the only thing."[19] Downs assumed politicians seek to maximize *votes*, although in variations on his model, politicians alternatively maximize their vote plurality (the difference between their vote and that of their closest competitor) or their probability of winning. In any event, most early spatial models of electoral competition took votes to be the coin of the realm, regarded politicians as focused exclusively on winning elections, and suggested that they did so by promising policies, platforms, and programs that attracted voters. In this spatial context, a candidate is represented by some location on the ideological continuum, some point in the [0, 100] interval.

[19] For those too young to remember, Vince Lombardi was the legendary football coach of the Green Bay Packers and, at the end of his career, the Washington Redskins. To Lombardi, nothing was more important than winning. While not exactly an uplifting imperative, Lombardi's maxim is a pretty good first approximation for what it takes to succeed in the national politics of most countries, the business world, and the National Football League.

FIGURE 5.10



Voters, Downs assumed, were single-mindedly interested in policy: the goods and services produced by government (or left to the private sector); the form and content of government regulation of the privte sector; the distribution of tax, unemployment, and inflation burdens; government policies on social issues like abortion and divorce; and matters of war and peace. Voters care mightily about these matters and base their assessments of candidates accordingly. Voters, however, are heterogeneous in their tastes so that, just as there are left-wing and right-wing politicians, so there are left-wing and right-wing voters. Specifically, each voter is identified with some point in the [0, 100] ideological space—the voter's ideal point—and his or her preferences are assumed to decline for points more and more distant from this ideal. That is, the set of voters may be represented by single-peaked preferences. Figure 5.10 displays an electorate of 625 voters (actually, five different voter "types" with 125 voters of each type). A voter of type i (i = 1,2,3,4,5) has ideal point x_i, and preferences declining in distance from x_i.

The most famous version of the Downsian model involves two-candidate competition. The question Downs asks is, given a distribution of voters like that in Figure 5.10, where will two single-minded seekers of election locate themselves? We can

gain some intuition on this question by fixing the position of one of the candidates. Let's fix the position of L, the leftie candidate, at l as shown in the figure. What position, r, should R, the rightie candidate, adopt so as to maximize his votes? To answer this question we need a rule of calculation. The Downsian rule is that each voter votes for the candidate whose location is closest to his or her ideal point.[20]

We can now answer the questions posed in the previous paragraph. Candidate R should snuggle up infinitesimally close to the right-hand side of L. That way, R gets all the votes to the right of l and, since l is to the left of the midpoint of the voter distribution, that means that R gets more than half of all the votes.[21] That is, R gets the 375 votes from voters of type 3, 4, and 5; L gets the 250 votes of types 1 and 2. Put more generally, L's location divides the electorate into two groups: those with ideals less than l and those with ideals greater than l. R's optimal response is an r just next to L on the side of the larger group. We have thus figured out how R will respond to any move made by L. L thus knows that she will get the *smaller* group and, given that she too wants to maximize votes, she should try to make this smaller group as large as possible. She can do this, the reader may have guessed, by setting l equal to the ideal point of the median voter, since then the groups to the left and right are equal in size. If l and r just straddle the ideal point of the median voter, x_3, then each location is optimal against the other's and the election ends in a virtual tie.

[20] For any two candidate positions, say α and β in [0, 100], where α is to the left of β, the midpoint is $(\alpha + \beta)/2$. The candidate located at α receives the votes of all voters with ideals to the left of this midpoint, whereas the candidate located at β gets the votes of all voters with ideals to the right of this midpoint. Voters at the midpoint are indifferent between the two candidates since their positions are an identical distance from these voters' ideals; these voters flip coins to decide for whom to vote.

[21] If l happened to be to the right of the midpoint of the voter distribution, then R would maximize his votes by squeezing up against L on its left side, thereby getting a majority of the votes.

We draw precisely the same conclusion if we fix *R*'s position first and let *L* optimally respond. For any *r* chosen by *R*, *L* will set *l* just next to *r* on the side of the larger group. Under these circumstances, the best R can do is to "move to the median."

Finally, suppose *L* and *R* must announce their policy platforms simultaneously. If once announced, a candidate is "stuck" with the position for the duration of the campaign, then a candidate is likely to worry that his or her position is vulnerable. A position is vulnerable if the opponent's position lies between it and the median of the voter distribution since, by the Downsian rule of calculation, the opponent will then get more than half the votes.[22] The only position that *cannot* be vulnerable is one that actually is at the median ideal. If, on the other hand, candidates are not stuck with their announced positions, but can revise their policy platform during the course of a campaign, then one will observe one of two patterns. If both initial announcements are on the same side of the median ideal, then there will be a "hopscotching" converging pattern as the vulnerable position (as just defined) hopscotches over her opponent's position in order to be closer to the median, that position in turn is hopscotched over by the now-vulnerable opponent, and so on until there is no more hopscotching to do—namely when both positions have converged upon the median. If, on the other hand, initial announcements are on opposite sides of the median ideal point, then there will be a homing in on the median from each side as the one more distant moves closer.[23]

In all of these circumstances, each a slightly different modeling assumption about the sequence in which various events

[22] Indeed, a position is vulnerable if the opponent's position is *closer* to the median. The position closest to the median wins more than half the votes.

[23] Of course if politicians are perfectly informed about voter preferences, they won't need to proceed tentatively toward the median, whether by hopscotching or homing in; instead they can move directly to the median.

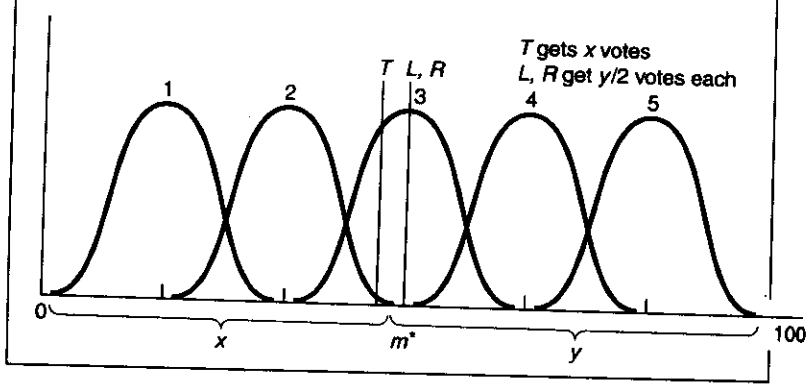
take place in the course of a campaign, there is a common convergence on the ideal point of the median voter. And this *centripetal* tendency is precisely what is predicted by Black's Median Voter Theorem. In effect, Downs's model provides a rationale for why majoritarian politics is centripetal.[24]

The logic of Black's theorem, as elaborated in the electoral context by Downs, reminds one of those occasions when someone says something very intelligent and quite obvious (once it is said!), causing you to reflect, "Now why didn't I think of that?" Downs was motivated by the fact that so many foreign observers of American life had, since practically the beginning of the Republic, noted how similar America's political parties were: "Tweedledum and Tweedledee," empty bottles differing only in their labels.[25] More recently, observers of British politics have begun to notice that the losing party ultimately transforms itself to look at least a bit like its more successful opponent. Thus, in the postwar period British Tories have accepted a good deal of the welfare state championed by the Labor Party, whereas, in recent days, the Labor party has trimmed its more socialist sails in order to look to voters a bit more like Margaret Thatcher's and John Major's Conservative Party.

[24] Notice that the rationale is not that the middle is "where the votes are." Certainly this may be true; in many circumstances the middle of the spectrum is where most persons' preferences lie, with the numbers getting smaller as one moves toward the more extremist tails of the distribution. But go back to Figure 5.10 and suppose that the extremists are the more plentiful. That is, suppose types 1 and 5 have 250 voters each, types 2 and 4 have 62 voters each, and type 3 consists of a single voter. Will the Downsian logic we have recounted above be any different here? We think not. The centripetal pull is the same, even though the "center" is *least populated* with voters!

[25] It might interest the reader to know that Downs's book originated as a doctoral dissertation in economics at Stanford University, where a member of Downs's dissertaion committee was Kenneth Arrow. So, Downs had both cycles and instability à la Arrow's theorem on one side and their opposite —stylized facts about stable party configurations— on the other. His research sought to make sense of these seemingly incompatible matters. Single-peakedness did the job.

FIGURE 5.11



The centripetal forces Downs identified are certainly plausible, yet it is clear that parties do not converge all the time. Why might this be? Downs's spatial model is quite user-friendly as a "discovery tool," so we can vary some of its assumptions and see what happens. Suppose, for example, we do not foreordain that there are two candidates. What if Leftie (L) and Rightie (R) are not the only two kids on the block? There is a third candidate, call her Trey (T), who may enter the race if she thinks she has a chance. Well, if L and R locate at the median (call it m^*)—$l = r = m^*$—and if, when there are more than two candidates the one with the *most* votes (not necessarily a majority) wins the election, then T certainly does have a chance. She can locate close on one side or the other of the median, win nearly all the votes on that side, and thus defeat L and R who end up splitting the remaining votes (Figure 5.11). On the other hand, if the positions of L and R are sufficiently widely dispersed, then T can enter *between* them at some position t. She will get the votes of voters whose ideal points lie in the interval $[(l + t)/2, (t + r)/2]$. The left boundary of this interval is the midpoint between the positions of L and T, whereas the right boundary is the midpoint between the positions of T and R. By the same Downsian rule of

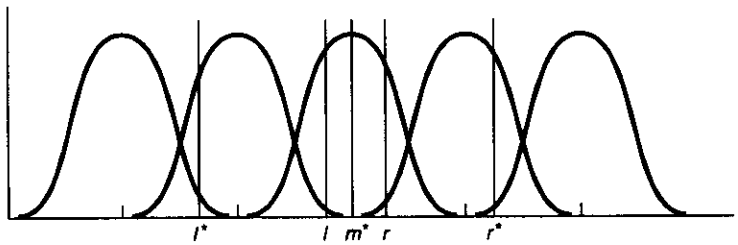
calculation, L gets all the voters in the interval $[0, (l + t)/2]$, and R gets all the voters in $[(t + r)/2, 100]$. If there are more voters in the first interval than in the second or in the third, then T wins. So, when there is the possibility of entry L and R can locate neither too closely together nor too far apart.

In fact, there may be a set of *entry-deterrence* locations for L and R, with these two getting roughly the same number of votes, and no third candidate able to locate in any place that would give her a victory (thereby discouraging her from entering at all). The point here is that when we broaden Downs's initial model to take account of some factor he had omitted—the possibility of entry by a third candidate—we discover that there may be occasions and circumstances in which the established parties (L and R) are ill-advised to converge toward the median.

Research has, in fact, been conducted on precisely the issue of Downsian candidate competition with (prospective) entry.[26] As noted, it is clearly one extension of the original Downsian assumptions that produces the possibility of nonconvergent candidate locations. But there are other possibilities. Candidates, for instance, may have their own policy preferences, ones often known to the voters. Thus, suppose L and R have their own policy ideal points at l^* and r^*, respectively (shown in Figure 5.12). They may declare policy programs at other locations, say $l \neq l^*$ and $r \neq r^*$. But why should the voters believe these policy declarations? One doesn't need to be altogether cynical to believe that once one of them wins it will be sorely tempted to implement its preferred policy (l^* or r^*), not its declared policy (l or r); politicians cannot be trusted to do what they say when they have preferences of their own. Effectively, then, candidates once again will not converge, this time because there is no point to doing so (they won't be be-

[26] The interested reader may consult Kenneth A. Shepsle, *Models of Multiparty Electoral Competition* (Chur, Switzerland: Harwood, 1991) for a summary of much of this research.

FIGURE 5.12



lieved by voters), even if they were willing to implement what they promised. The "commitment technology" is simply not up to the task.

What if it were? What if candidates had policy preferences, as in the previous paragraph, but had available to them means of making promises stick. Perhaps all they need to say is "Cross my heart," and the voters will believe them. Perhaps voters believe policy promises because they know that politicians know that a reputation for deception and misrepresentation is a serious electoral obstacle in future electoral campaigns. So, for any of a number of reasons, suppose that candidate promises are credible, on the one hand, but that candidates still care about what policies are implemented, on the other hand. What will the candidates do in this circumstance? In a lovely paper, Randall Calvert,[27] demonstrates that, just as in the case where candidates didn't care a whit about policy, these two candidates will converge to the median voter's ideal. Referring again to Figure 5.12, L wants an outcome closest to l^* and R wants the final policy to be closest to r^*. If these two points happen to be equidistant from m*, and if each candidate (credibly) announced his or her ideal policy,

[27] Randall Calvert, "Robustness of Multidimensional Voting Models: Candidate Motivations, Uncertainty, and Convergence," *American Journal of Political Science* 29 (1985): 69-95.

respectively, then the election would end in a tie (and, presumably, the winner would be determined by something like a flip of a fair coin). But L, by moving just a tad toward the center, could win the election outright at a very small cost to herself in terms of policy. But this would be terrible for R—not that he lost the election but that a policy near l^* is so awful. He could avoid all this by moving in toward the center a bit more than L had, which, in turn, encourages L to move in a bit more, and so on and so forth. In the end, *even though both candidates had policy preferences and, in fact, did not care at all about who won the election but only about what policy would be implemented,* they converge to the median voter's ideal anyhow.

Needless to say, we could play with Downs's model in a variety of interesting ways. Many have.[28] What we have shown in this section is that the stripped-down spatial model of Downs, with competition on an ideological dimension between two election-oriented candidates, leads to policy convergence. The policy that emerges from the competitive forces captured by this model is the ideal point of the median voter. This result, to the casual observer, describes what often happens in real elections, as candidates try to smooth down their more extremist edges in order to curry favor with voters in the center of things. Thus, once Bill Clinton vanquished his liberal opponents within the Democratic Party in 1992 (Jesse Jackson and Mario Cuomo), he headed toward the ideological center, running in the general election as a more conservative "new Democrat." Incumbent president George Bush, on the other hand, tried to shed some of his hardline conservative attributes, also moving toward the center as he compromised on his

[28] For a summary of the extensive literature the interested reader may turn to the following two volumes: James M. Enelow and Melvin J. Hinich, *The Spatial Theory of Voting* (New York: Cambridge, 1984); and James M. Enelow and Melvin J. Hinich, eds., *Advances in the Spatial Theory of Voting* (New York: Cambridge, 1990).

"no new taxes" pledge. In many other elections one sees a similar dynamic—partisan candidates of the left and the right hedging, qualifying, and compromising in order to appear more centrist.

This convergence is not always complete, however. Sometimes a candidate applies brakes on convergence for fear of alienating his or her base, or even stimulating a third-party entrant. Thus, civil rights activists, unions, and government workers—elements of the Democratic base—made it virtually impossible for candidate Walter Mondale to converge toward the center in the 1984 presidential election. Elements of the conservative movement kept Ronald Reagan ideologically true in that same election. Third-party candidates entered the presidential races of 1968, 1980, and 1992 (George Wallace, John Anderson, and Ross Perot, respectively), sometimes because the candidates were thought to have converged too much,[29] sometimes because they were thought to have stayed too close to their more extremist supporters.[30] Thus, both too much convergence and too little convergence may provide the impetus for a third-party challenge.

We have clearly only scratched the surface of Downs's spatial model of party competition, and only covered some of the many mechanisms and rationales according to which competitors converge toward the median voter's ideal policy, on the one hand, or maintain distinctive policy profiles, on the other. This, in sum, suggests the richness of Downs's approach.

Electoral phenomena, however, are not the only focuses of the spatial model. A twin enterprise, a kind of "elections writ small," has employed the spatial model to study the selection

[29] In 1968, Wallace entered on the right, thinking "there's not a dime's worth of difference" between Democrat Hubert Humphrey and Republican Richard Nixon.

[30] Both Anderson and Perot sought to capture the center, which they believed had been conceded by Carter and Reagan in 1980 and Clinton and Bush in 1992, respectively.

of policy in legislative settings. We turn to those now, and examine both one-dimensional and multidimensional versions of the spatial model.

Spatial Models of Legislatures

We shall have a much more thorough look at legislatures in Part IV, so here we are primarily interested in seeing what the spatial model can do. We shall see that it is a quite powerful analytical tool for representing the ways in which preference-based (rational) behavior and structural features of institutions interact to produce final outcomes. It suggests that legislative outcomes depend in essential ways not only on what legislators want, but also on how they conduct business in the legislature.

To keep things as simple as possible, we take the legislature to be a set of n individuals, where n is an odd number, and where everyone casts a vote. It makes decisions by majority rule. The most elementary situation, one we examine first, is the unidimensional case in which the legislature must choose a point on a line. Each legislator, i, has an ideal point x_i, and single-peaked preferences. The median voter is legislator m with ideal point x_m. We know in this circumstance that x_m can defeat any other point on the dimension in a majority contest (Black's Theorem). Perhaps more amazing is the fact that the median preferences prevail in a comparison between *any* two alternatives, so that if m prefers x to y then so does a majority.[31]

[31] This may be proved as follows. Suppose x_m, x and y are all points in the dimension, and that $x_m \leq x \leq y$. Legislator m clearly prefers x to y. But then so does every legislator to the left of m. Together these legislators constitute a majority, so x is preferred by a majority to y. Likewise, by the same reasoning, if $y \leq x \leq x_m$, then both m and a majority (all the legislators to the right of m) prefer x to y. So, we have shown that whenever x and y are on the *same* side of the median, a majority always agrees with the preferences

In addition to the preferences of the median legislator, x_m, we identify two other distinguishing features of the situation. Whenever a legislature faces a decision-making opportunity, there is always a *status quo* in place, labeled x^0. This is the current policy at the time of legislative choice. We assume it remains in place if the legislature chooses not to change it.[32] The second feature of interest common to most legislatures is a division-of-labor arrangement known as a *committee system*. In such a system, a committee is a subset of the n legislators (momentarily, we describe some of its specific powers). The median ideal point of the committee members is labeled x_c. Just as majority preferences in the entire legislature are identical to the preferences of the legislature's median voter, majority preferences inside a committee are a copy of the preferences of the committee's median member. Because of these identities, much of our analysis need only consider x^0, x_m, and x_c. In what follows, then, we put the spatial model through its paces in examining the making of policy choices by an n-member legislature possessing a committee system.

We consider here three decision making regimes, or institutional arrangements. The first is ***pure majority rule.*** There is a status quo and any legislator can offer a motion to change it. A motion, once proposed, is pitted against the status quo. If it wins it becomes the new status quo; if it loses it goes to the place where all losing proposals go (a sort of elephant's burial ground). The floor is once again open for some new motion (against the old status quo, if it survived, or the new status quo, if the previous proposal prevailed). This procedure of motion-making and voting continues until no member of the legislature wishes to make a new motion.[33]

The second regime is the ***closed-rule committee system.*** In this system, a (previously appointed) committee first gets to decide whether the legislature will consider changes in the status quo; that is, it has *gatekeeping agenda power*, and can decide whether to open the gates to enable policy change or not. Second, if the gates are opened only it gets to make a proposal (*monopoly proposal power*). Third, the parent legislature may vote the committee's proposal either up or down; the proposal is closed to amendments. Hence, the proposal is said to be considered under a *closed rule*, and the committee is said to offer its parent body a *take-it-or-leave-it-proposal*.

The third regime we examine is the ***open-rule committee system.*** This system is identical to the one described in the previous paragraph, except for the third feature. Under an open rule, once the committee has made a proposal, the parent legislature may open the floor to *amendments* to the committee's proposal. Once the committee has opened the gates and made a proposal, it concedes its monopoly access to the agenda.

We will examine each of these systems in both the one-dimensional and the multidimensional setting. We want to know if there is anything regular or routine that we can expect from these alternative majority-rule regimes. In conclusion, we will offer some brief comparative observations on these regimes, leaving a full-blown consideration for Chapters 11 and 12, where we will take up institutions more systematically.

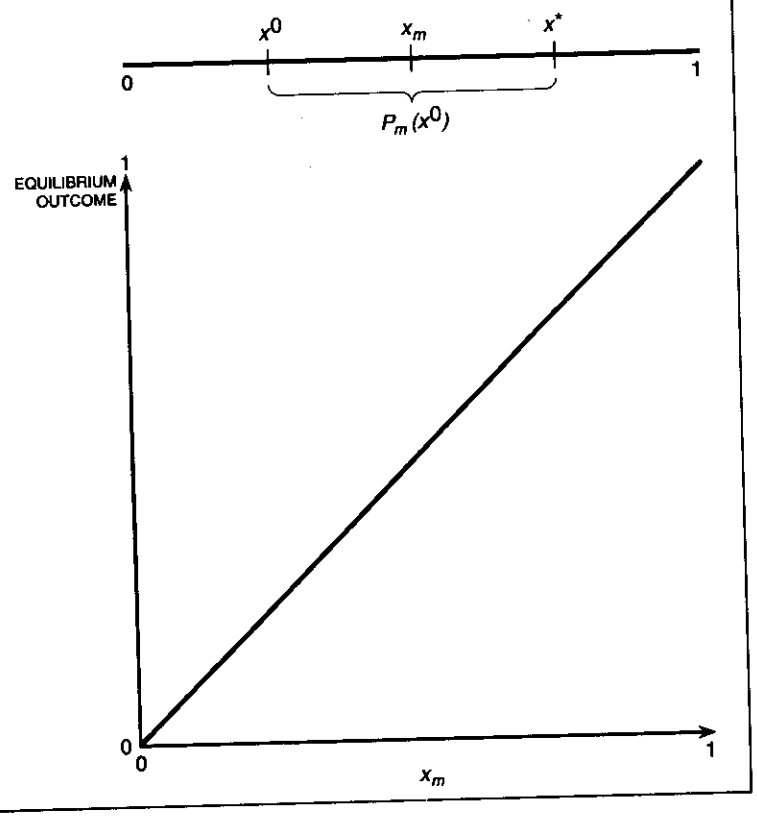
PURE MAJORITY RULE We start with a legislative choice in a one-dimensional spatial setting. In the pure majority-rule re-

of the median voter. Suppose, then, that $x \leq x_m \leq y$, and that m prefers x to y. Consider the legislator just to the left of m. Her ideal is closer to x and farther from y than was m's ideal; so if m prefers x to y, surely she does, too. But, then so do all the other voters to the left of m's ideal, and once again they jointly comprise a majority. This establishes that a majority *always* agrees with the preferences of the median voter.

[32] In some circumstances, the status quo policy does *not* remain in place, unless the legislature takes positive action to keep it in place. If the legislature fails to do anything, then the status quo reverts to some specific policy (known, naturally enough, as the *reversion point*). This is true, for instance, in statutes that possess *sunset provisions*, an example we explore further in Case 5.1.

[33] A variation on this "stopping rule" is to allow a motion to be in order at any time to close the floor to new motions (in effect, a motion to take a final vote and then to adjourn the legislature, at least on the subject matter at hand).

FIGURE 5.13



gime there are no committees, so we need to know only the locations of x_m and x^0. A typical situation is given in the top panel of Figure 5.13. From the median ideal and the status quo, we determine the median legislator's preferred-to-x^0 set, $P_m(x^0)$[34]. Since we've established that the median's preferences

[34] Since we assume that legislative preferences are distance-based, we know that legislator m prefers to x^0 all points closer than it to x_m. This determines the set $P_m(x^0)$ pictured in Figure 5.13.

are the same as a majority's preferences, this interval is the set of motions that would prevail over x^0 in a majority contest. So, if someone is recognized and makes a motion *outside* this set, it will go down in flames, whereas any motion *in* this set will be victorious and become the new status quo. It is evident that the political process defined this way will produce outcomes that either leave the status quo unchanged or move it closer to x_m (since every point in $P_m(x^0)$ is closer than x^0 to x_m). Ultimately it will converge on x_m. Moreover, it will not depart once it reaches x_m (since, as we showed above, the status quo cannot move further from x_m in any vote). So, just as in the Downsian model of electoral competition, there is a centripetal tendency in the pure majority rule legislative regime. It is for this reason that we think of pure majority-rule legislative choice as an "election writ small."

The great utility of these spatial models of legislative choice is they permit the analyst to do what in economics is called *comparative statics*—we can ask "what if" questions. Having derived an equilibrium outcome from our basic setup, as we did in the previous paragraph, we may now ask how that equilibrium changes as relevant parameters change. We have already seen that there are really only two relevant parameters—x^0 and x_m. Holding the latter fixed, we first ask what happens if the former changes—that is, what would happen to the final outcome if the status quo were closer to or further from the median legislator's ideal? The answer is: nothing! Although different locations for x^0 will affect $P_m(x^0)$, and hence what motions can succeed at any point in time, we know that ultimately the result converges to x_m, no matter what x^0 is. So one interesting conclusion we draw from the pure majority rule model is that *it does not possess a conservative bias*, weighting past decisions unduly. The past (as reflected in the status quo) will influence the "path" (by restricting what motions can succeed at different stages of the process), but will not affect the ultimate destination.

Reversing emphasis and holding x^0 fixed, we now ask what happens if x_m were different. The answer is graphed in the lower panel of Figure 5.13 for x_m located on a line between 0 and 1. The location of x_m in the [0, 1] interval is given on the horizontal axis and the equilibrium outcome in this interval is given on the vertical axis. The graph is a 45° line showing (for any x^0) that the equilibrium outcome *perfectly tracks* the identity of the median ideal point. This is centripetality in the extreme! Not only does pure majority rule legislative choice converge to the median ideal; but also, if that median should change, then so will the equilibrium outcome. So a second interesting conclusion we can draw from pure majority rule is that it is perfectly responsive to central tendencies: The median legislator's ideal is, by definition, the central point in the distribution of preferences; pure majority rule produces an outcome at this point; and, were this point to change, the legislative outcome would "track" it.

CLOSED-RULE COMMITTEE SYSTEM Most legislatures are not pure majority-rule institutions. Even town meetings and other approximations to pure majority rule about which observers occasionally wax romantic require some mechanism to determine the content of agenda items and the order in which they will be taken up. Some legislatures establish a single agenda committee to decide these matters. However, most legislatures (certainly in the United States) employ a division-of-labor committee system that divides up agenda power by policy area. Subsets of legislators have disproportionate influence over the agenda in specific policy jurisdictions. The committee serves, in its jurisdiction, as an agenda agent for its parent legislature.

We will have much more to say about these things in Part IV. For now, we need focus only on the fact that what distinguishes the closed-rule regime from pure majority rule is that there is, in addition to x^0 and x_m, a third parameter of

interest, namely the median ideal point of an agenda-setting committee, x_c.[35] Many of the conclusions we draw about this regime depend upon the relative locations of x^0, x_c, and x_m.

The decision-making procedure, as we suggested earlier, is for the committee either to make no proposal at all, in which case x^0 remains in place, or to make a motion to change the status quo, which the parent body must accept or reject as is. What will such a committee do? To answer this question, we once again determine $P_m(x^0)$, as in the top panel of Figure 5.13. This is a set whose boundary points are x^0 itself and x^*; it contains the only points a legislative majority prefers to x^0. The committee, as personified by its median voter, c, treats these points as its "opportunity set," picking its favorite as the motion it makes (if it makes any motion at all). We look at three orderings of the relevant parameters (there are six orderings in all, but the omitted ones are simply mirror images of the ones we consider):

CASE 1 ($x^0 \leq x_m \leq x_c$). Here the median legislator is between the status quo and the median committee member. In this case x^* is the *right* boundary of $P_m(x^0)$, just as shown in Figure 5.13. If $x_c \leq x^*$, then the committee will propose its median ideal point which then will be approved by a legislative majority (since it lies inside $P_m(x^0)$). If, on the other hand, $x^* < x_c$, then the best the committee can do is to propose x^*, which is approved by a legislative majority.[36] In either case, both committee and parent legislature wish to

[35] Since we are elaborating the one-dimensional model here, we are only interested in the median ideal of a single committee. In multidimensional contexts, where there are many jurisdictions into which the dimensions of the policy space are arranged, we will need to know the policy preferences of different committees, each responsible for its own bundle of policy dimensions. More on this will be developed in Part IV.

[36] Actually a legislative majority is indifferent between x^0 and x^*. We assume that an indifferent voter votes *for* the motion on the floor. (Alternatively, the committee could propose a point just to the left of x^*, which secures a majority outright.)

move away from the status quo *in the same direction*. The final outcome will move x^0 in that direction, further than the median voter would want, but not always as far as the committee median wants.

CASE 2 ($x^0 \leq x_c \leq x_m$). Here the median committee member is between the status quo and the median of the whole legislature. In this case $x_c \in P_m(x^0)$ automatically. So the committee can get majority legislative approval for x_c.

CASE 3 ($x_m \leq x^0 \leq x_c$). In this last setting the status quo is between the two medians. This is a particularly interesting case, because committee and legislative majority are at loggerheads. The committee wishes to move right, while a majority of the parent legislature wants to move left. The committee's gatekeeping authority pays off for it in a big way here, because it will choose simply to keep the gates closed.[37]

So, the first thing we learn about the closed-rule regime is that only a very limited number of things can happen—three things, in particular. If x_c is interior to the legislative median's preferred-to-x^0 set, then the outcome is x_c. If it is not, then either of the two end-points of $P_m(x^0)$ are possible—x^0 if committee and legislative median are at loggerheads; x^* otherwise. In the pure majority rule regime, in contrast, only *one* thing can happen—x_m—something that *never* happens under the closed rule regime (unless, by coincidence, $x_m = x_c$ or $x_m = x^0$). This suggests that endowing a privileged group with agenda power is not without its consequences: *agenda power discourages centripetal outcomes as it tugs the process in the direction of the privileged group.*

[37] The committee could move a proposal (some point to the right of x^0), but it would be defeated. So it might as well not bother and simply keep the gates closed (especially if the bother were at all costly). On the other hand, if outside interests took heart in the fact that the committee was at least putting up the good fight, and rewarded the committee accordingly, then the committee might wish to "bother" (though the result would be unchanged—x^0 would stay in place).

There are a variety of comparative statics exercises one might do. We focus on one: for a fixed legislative median and committee median, what happens as x^0 changes? (When we ask this question, we don't literally mean that the status quo suddenly changes. Rather, we are asking what would happen if the status quo were more or less extreme.) In case 1, for example, if x^0 were further to the left, then $P_m(x^0)$ would get bigger (x^* moves to the right). At some point it contains x_c (if it doesn't already). So, as x^0 moves away from the chamber median, there will be a discontinuity when x_c jumps from being outside m's preferred-to-x^0 set to inside that set. Put crudely, the worse the status quo, from m's perspective, the more likely c can get her way.[38] The same pattern prevails as x^0 moves to the right. At first it moves toward x_m, so *m*'s preferred-to-x^0 set contracts. Once it "passes" x_m the preferred-to set begins expanding again.[39]

CASE 5.1
SUNSET PROVISIONS AND ZERO-BASED BUDGETING

In the 1970s public policy analysts developed two ideas as an attempt to counter rising budget pressures. The first idea was called a sunset provision. The identified problem was the persistence of expenditures that may have outlived their usefulness. It seemed that once a project was on the

[38] The classic statement of this result, plus a derivation of some of the political consequences of it, is found in Thomas Romer and Howard Rosenthal, "Political Resource Allocation, Controlled Agendas, and the Status Quo," *Public Choice* 33 (1978): 27-43.

[39] The reader might try to see what happens as x^0 changes in cases 2 and 3.

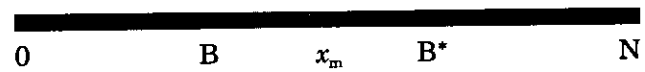
books, it never went away. With a sunset provision as part of the enabling legislation, the project would have to be renewed after a specified time period in order to extend its life. In other words, the sun would automatically set on a project, unless the legislature took further action.

The second idea was called zero-based budgeting. Also associated with the problem of expenditures that were growing out of control, this concept required bureaucratic agencies to justify a project budget "from zero," rather than merely justifying the *growth* in proposed expenditures over the previous year's budget. It was alleged that this procedure would reduce accumulating and persisting inefficiencies in agency budgets.

For our purposes, sunset provisions and zero-based budgeting are similar because they create situations in which the status quo alternative to a proposal is zero. Consider the case of zero-based budgeting for an agency. The legislative median is at x_m and last year's agency budget is at B. Now we let the agency make a proposal for next year's budget, a proposal which the legislature may accept or reject by majority rule. Under ordinary procedures, we assume that legislative rejection of the agency proposal results in last year's budget, B, continuing in place next year. Under zero-based budgeting, on the other hand, we assume that legislative rejection leads to a zeroing out of the agency budget altogether.

At first glance, it would seem that the zero-based budgeting procedure is pretty tough on the agency. That's the whole idea, since this method was designed to limit an agency's power in budget negotiations. However, zero-based budgeting actually *increases* agency discretion. This is seen in the figure below. If, under ordinary procedures, B is the reversion outcome (the outcome if the legislature rejects the agency proposal), then as long as the agency proposes a

budget in the gray region, the legislature will approve it (any such proposal is in $P_m(B)$); that is, under ordinary procedures, the agency could extract a budget as large as B^*. If, on the other hand, the zero-based budgeting procedure were in effect, then the agency could get a budget as large as N (since N is in $P_m(0)$). If only the policy wonks who invented zero-based budgeting had an analytical model of the legislative process, they would have appreciated the *perverse* incentives such an arrangement provides.



In concluding this brief treatment of the closed-rule regime, let us reemphasize the fact that the key parameters are x^0, x_c, and x_m. An electoral earthquake that fails to change relationships among these parameters will not change policy outcomes (a fact that may puzzle those not equipped with the theory we have been developing here). If, for instance, a legislative election caused massive turnover in incumbents, but did so symmetrically so as to leave x_m unchanged, then "the more things change, the more they stay the same." Likewise, if before an election legislator c and m are at loggerheads, as defined in case 3 above, then electoral change, no matter how massive, that leaves the (possibly newly determined) c and m at loggerheads, will simply maintain the status quo ante. The institutional impediments implicit in the closed-rule regime stand in stark contrast to the hypersensitivity of pure majority rule.

OPEN-RULE COMMITTEE SYSTEM We've seen thus far that, although there is an entire continuum of possible final outcomes, only one thing (x_m) can occur under the pure majority rule

regime, and only one of three things (x_c, x^0, or x^*) can possibly happen under the closed-rule regime. In our treatment of the open-rule regime, we will discover that only two possibilities exist. Either the gates remain closed and x^0 prevails or the gates are opened and x_m is the final outcome. *Nothing else is possible.* We will consider all the cases as we did in the previous regime.

In the open-rule regime the committee once again has the first move. If it makes no motion, then x^0 persists. If it makes a motion, then that motion is open to amendment (hence the term *open* rule). We assume here that alternative amendments continue to be offered until no legislator wishes to offer another. So the committee proposal is initially pitted against the first amendment, the winner of that against the next amendment, and so on until all the amendments have been taken up; the survivor of that sequence is then pitted against the status quo (this last vote is often called the "vote on final passage").[40]

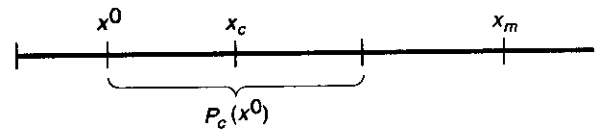
This procedure looks very much like the pure majority rule regime, *except that the committee has the first move.* Once it opens the gates, we're in the world of pure majority rule. This means that once a proposal is made, it will be amended and amended again, successful amendments converging the process toward x_m. Indeed, it doesn't even matter what the initial committee proposal is. The reality is:

$$\begin{aligned} &\text{open the gates} && \Rightarrow x_m \\ &\text{keep gates closed} && \Rightarrow x^0 \end{aligned}$$

The committee decision is really pretty simple. If the committee median voter, Ms. c, prefers x_m to x^0, then she makes a

[40] An alternative procedure would be to allow, after a committee motion, an amendment which is directly voted on. The winner stays on the floor and is subject to another amendment. The process continues until no more amendments are forthcoming, after which there is a final vote between the alternative left standing on the floor and the status quo.

FIGURE 5.14



motion (any motion); if Ms. c prefers x^0 to x_m, then the committee keeps the gates closed. Thus, all we need to inspect is Ms. c's preferred-to-x^0 set, $P_c(x^0)$, to see whether x_m is in it or not.

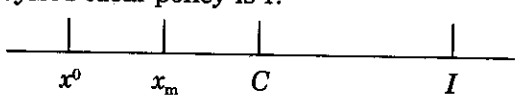
Recalling the three possible cases in the preceding section, for the parameter ordering of case 1 ($x^0 \leq x_m \leq x_c$), the committee clearly prefers x_m to x^0. For case 3 ($x_m \leq x^0 \leq x_c$), the committee clearly has the opposite preference. It is the case 2 ordering ($x^0 \leq x_c \leq x_m$) that is the interesting one. If c's ideal policy is less than halfway between x^0 and x_m, then she keeps the gates closed; if it is more than halfway, then she makes a motion.

The first of these case 2 situations is shown in Figure 5.14. What makes this especially interesting is that it represents a very frustrating situation. The committee, because it prefers the status quo to the median legislator's ideal, will keep the gates closed. *But both a committee majority and a chamber majority prefer every point in* $P_c(x^0)$ to x^0. That is, the open-rule environment, which at first blush appears to give a legislative majority potent authority, in fact penalizes both committee and legislative majorities. It gives the chamber *too much* authority—the right to amend whenever it wants. Its strength is its weakness, because it cannot promise *not* to use its authority; yet, it would be better off if it could credibly promise not to amend some proposal in $P_c(x^0)$ made by the committee (for, if it could precommit in this fashion, then the committee would be prepared to open the gates).

Case 5.2
The Importance of Compromise and Strategic Thinking

Our discussion of the closed-rule and open-rule regimes addresses the general question of how politicians and interested others think about legislative possibilities. In one-dimensional situations, as we have seen, politicians locate a proposal on the policy dimension relative to their own preferences and the status quo that will otherwise prevail if the proposed change is rejected. When faced with a choice between the proposal and the status quo, the politician votes for the alternative closer to his or her ideal. We demonstrated that, with the open-rule regime, once a proposal is made the dynamic of amendment activity leads inexorably to a unique outcome—the ideal point of the median legislator. With the closed rule, a proposal wins only if it is closer than the status quo to the median voter's ideal.

Failure to recognize these dynamics can lead to disappointment for principled, i.e., stubborn, lobbyists. Advocates of proposed legislation must take into consideration the preferences of the decision maker(s), the rules of procedure in effect, and the relative location of the status quo. An unwillingness to compromise in light of these strategic realities can keep the status quo in place, even though the possibility exists to defeat it with results satisfactory to the lobbyist. This is illustrated in the figure below, where the legislative median is x_m, the status quo is x^0, and a powerful lobbyist's ideal policy is I.



Without going into any of the specifics concerning how powerful lobbyists exercise their power, suppose that lobbyist I is in a position to undermine any change in the status quo if it finds the change not to its liking (perhaps by "bribing" influential legislators—that is, contributing to their campaign committees). Under the closed rule, the best I could hope for is the compromise point, C—a policy just a little bit closer than x^0 to x_m. If the lobbyist stubbornly refuses to accept C by seeking something more extreme, it loses. Under the open rule, it must be prepared to accept x_m, for this is where the process of amendment will drive the final result. In either of these cases, the lobbyist must be able to anticipate the best deal it can cut and settle for it. In particular, it must be especially sensitive to the fact that "the best deal it can cut" depends upon the procedural rules for amendments. Even though it is powerful enough to *undermine* proposed changes in x^0, it cannot *impose* its own will. It needs a little help from its (legislative) friends.

Some observers have cited the absence of such strategic thinking as a reason for the failure of the Equal Rights Amendment* and Proposition 174 in California, which would have implemented school choice as a voucher system.† Unwilling to compromise, lobbyists unwittingly kept their proposals further from the status quo than the compromise point required by the strategic realities. Politicians or voters voted against their proposals when more moderate versions very probably would have passed. The importance of such strategic thinking will be covered in greater depth in the next chapter.

* Jane Mansbridge, *Why We Lost the ERA* (Chicago: University of Chicago Press, 1986).

† Jack Anderson and Michael Binstein, " 'School Choice' Pitfalls in California," *Washington Post*, November 1, 1993.

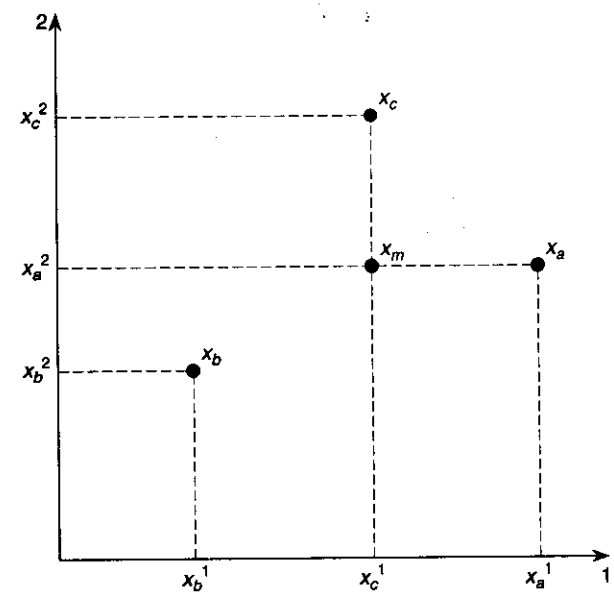
MULTIDIMENSIONAL EXTENSIONS Once we move into multiple dimensions, matters get a bit more dicey. In a pure majority rule regime, the results of the McKelvey Chaos Theorem loom large. Putting to one side the highly unlikely circumstance that legislator preferences are distributed in a radial symmetric manner, we know that $W(x) \neq \varnothing$ for any x in the policy space. Anything can be beaten. In particular, any status quo, x^0 has a nonempty winset, $W(x^0)$. So long as a motion is made from that set, the status quo will be replaced. But then $x^1 \in W(x^0)$, in turn, has a nonempty winset of its own, $W(x^1)$. A motion $x^2 \in W(x^1)$ will replace x^1. Under the assumptions we have made about legislative voting,[41] an existing status quo is continually replaced.

Suppose we alter the setup ever so slightly. We retain the condition, from pure majority rule, that anyone is free to make a motion to change the status quo. But we assume that decision making takes place *one dimension at a time*, in some preset order. The first person recognized to make a motion on the initially designated dimension states his or her amendment to x^0; this amendment can only change x^0 on the dimension currently under consideration. The group continues to focus on amending the status quo on this dimension until no more amendments are offered. Once it completes its task, the group turns its attention to the next dimension. It continues in this manner until there is no dimension left on which any legislator wishes to alter the status quo level.

It is easy to see, as we show in Figure 5.15, that this multidimensional version of pure majority rule mimics the result of the one-dimensional setting. There are three legislators with ideal points $x_a=(x_a^1, x_a^2)$, $x_b=(x_b^1, x_b^2)$, and $x_c=(x_c^1, x_c^2)$. For any status quo (not pictured), $x_0=(x_0^1, x_0^2)$, motions are entertained, first on dimension 1 and then on dimension 2. At the

[41] Namely, that everyone votes their preferences rather than voting strategically (which we take up in the next chapter).

FIGURE 5.15

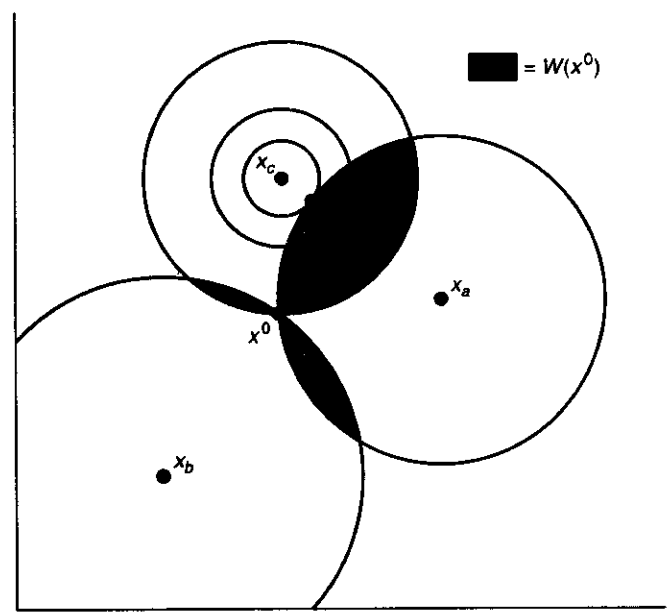


end of the day, when all motion-making and voting are said and done, the final outcome is the *multidimensional median*, $x_m=(x_c^1, x_a^2)$, since legislator c is median on the first dimension and legislator a is median on the second. This does not mean that $W(x_m)=\varnothing$. There are points, for instance, that both a and c prefer to x_m. Rather, it means that on any dimension—say the first—holding policy *fixed* on the other dimension, no movement away from x_c^1 would be supported by a majority. (For any point on the horizontal line through x_m—the points in which policy on the first dimension changes but remains fixed on the second—two of the three legislators always prefer x_m to it. The same holds for points on the vertical line through x_m. The only points preferred by a majority to x_m require changes on *both* dimensions at once.)

Thus, the multidimensional version of pure majority rule yields one of two possible conclusions, depending upon whether there is additional institutional structure or not. In the pure case, the status quo is continuously vulnerable to change. The group's choices are never very durable, since it is always in someone's interest to introduce a motion to change it, and it is always in some majority's interest to comply. In the case of institutional structure in the form of dimension-by-dimension decision making, the result is both predictable and centripetal. The median ideal point on each dimension prevails under the procedure described above (although it need not be the same median voter each time, of course).

Since we will take up the multidimensional versions of the open-rule and closed-rule regimes in the chapter on legislatures in Part IV, we will be especially brief on this subject now. Imagine, in Figure 5.16 (a reproduction of the spatial positions in Figure 5.13), that Ms. c is an agenda-setter and the status quo is x^0. If her proposals are subject to amendment by the parent legislature, then we are back to the wild-and-woolly open-rule majority system. Under a closed rule, however, she can make a take-it-or-leave-it proposal, one that is *not* subject to amendment but only to an up-or-down vote. The petal-shaped shaded regions comprise c's opportunity set—it is $W(x^0)$. The circles centered on x_c are various of c's indifference contours. Her objective is to move the final policy outcome onto the indifference contour of smallest radius (hence closest to her ideal point) that still lies in $W(x^0)$. The point in this figure (a big black dot!) at the tangency between one of the petals of $W(x^0)$ and the smallest indifference curve of Ms. *c* is the proposal she will make, which a majority (*a* and *c*) will then support. With the closed rule, then, a monopoly agenda setter has considerable power, though constrained by majority preferences.

FIGURE 5.16



CONCLUSION

The spatial model will be used time and time again in the analyses of the remainder of the book. We've put forward the basic ingredients in this chapter and briefly explored majority rule in electoral and small group settings. But we've assumed a good deal of naïveté on the part of our voter/legislators (apparently, only candidates and agenda setters are wily). We want to relax this unrealistic feature in Chapter 6. In Chapter 7 we want to move beyond majority rule and examine the multitude of ways human creativity has manifested itself in devising sometimes bizarre and intricate ways for groups to arrive at decisions.

What is exciting about the spatial model to an analytical political scientist is the opportunity it affords to capture many of the interesting details of political competition—whether between candidates in an election or between alternative motions in a legislative assembly—and, at the same time, to do so in a fairly clean and simple manner. We appreciate that the reader may not agree entirely with the last sentiment, since this chapter has required your undivided attention and careful reading. Nevertheless, political scientists over the past thirty or so years have found the model to serve as the principal building block for the analysis of political rivalries of all stripes.

A single chapter in a book, of course, can only portray the spatial model at its simplest. But even the simple formulation possesses nonobvious implications. In the context of one-dimensional pure majority rule with single-peaked preferences, for example, whether in two-party electoral competition or legislative policy choice, the magnetic attraction of the median participant's ideal point is powerful. Majoritarian politics is subjected to centripetal forces, producing outcomes that obervers describe with words such as "compromise," "moderate," or "centrist." At the very least, then, the simple spatial model provides a rationale or explanation for the inexorable movement of majority-rule competition toward the center of participant preferences. Its surprise value lies in the fact that this centripetal dynamic is *not* because "that's where the votes are." As we demonstrated earlier, movement toward the median voter's ideal is an equilibrium tendency in a pure majority rule arrangement *even if there are very few voters at the center of things.*

In the legislative realm, where rules of procedure and agenda-setting committees constrain the operation of pure majority rule, there were other surprises. At least in the world of one-dimensional politics, only a limited number of things are possible. A committee system operating under the open rule

can produce only one of two possible results. It can, by "closing the gates" and not permitting a motion to be proposed, keep policy at the status quo. Or, if it should make a motion, the sequence of amendments permitted under the open rule will drive the outcome to the median legislator's ideal. These are the only possibilities. Under a closed rule, there are three possibilities. If the gates are kept closed (possibly because the committee and legislature are at loggerheads, with the status quo between their respective median ideals), the status quo remains intact. If the committee median's ideal lies between the status quo and the median legislator's ideal, then the committee's ideal will be proposed and will pass. Finally, if the median legislator's ideal lies between the status quo and the committee's median ideal, then the outcome is the point closest to the committee's median that leaves the median legislator of the full legislature just indifferent between it and the status quo. The details are found in this chapter. The surprise, however, is found in the conclusions, first, that only a small number of items are possible under various legislative procedural regimes and, second, that these small number of things differ from regime to regime. Put differently, *institutional arrangments—the political ways of doing business—matter profoundly for the outcomes that emerge from the political process.*

The spatial model also allows us to begin to assemble explanations for why convergence to the center is not always complete. The centripetal tendency is always present, to be sure, but there may be countervailing tendencies as well. In the context of Downs's model of elections, for example, we noted that a politician may fear he will lose his extremist support (to abstention or to a third-party entrant) if he converges too much toward his opponent. In the legislative arena, powerful agenda setters may, through their control of motions and amendments, prevent the process from converging on the median legislator's ideal policy, either because the agenda setter

can propose and get passed something she likes better or because she chooses to keep the gates closed.

Thus, the great advantages of the spatial model are its (relative) simplicity, its analytical power, and the "surprises" it produces. Not only do we begin to understand things that we may have long appreciated in an intuitive fashion (like the tendency toward moderation in majority-rule systems); but also, we develop a sophisticated understanding of new things. While surely not a perfect explanatory tool, it's a pretty good start.

One of the matters that we touched on only briefly and unsystematically was strategic thinking. Many applications, especially early in the history of spatial modeling, assume that voters and legislators are "honest" in their voting behavior. When confronted with two alternatives, they simply vote for their favorite, doing so without regard for subsequent consequences. Yet there are many circumstances in which a rational person will think things through in a more sophisticated fashion, sometimes coming to the conclusion that, in a particular voting opportunity, she should *not* vote for her favorite alternative. We call this *sophisticated* or *strategic* behavior, the subject of the next chapter.

6

Strategic Behavior: Sophistication, Misrepresentation, and Manipulation

In our models of social choice and spatial decision making, voters vote their preferences. To some, this is the essence of rational behavior. To others, however, rationality is more subtle and nuanced. It entails doing the best one can with what one's got, and this sometimes requires making strategic maneuvers, investments, sacrifices, and retreats. It is embodied, for example, in the Protestant work ethic, which encourages deferred gratification in order to harvest later returns. Some might claim that Protestants, in some perverse sort of way, *like* deferred gratification, but we think their ethic (an ethic common among Asians, Jews, and others, too) reflects the strategic decision to maximize over the long haul by resisting enslavement to short-term preference satisfaction. Our models need to reflect this possibility, for a failure to take strategic capabilities into account may result in disaster (as Case 6.1 demonstrates). This chapter is devoted to elaborating on the multiplicity of ways strategy rears its head.

FROM SHEPSLE & BONCHEK. 1997. ANALYZING POLITICS. NEW YORK: NORTON

7

Voting Methods and Electoral Systems

In the last few chapters an implicit theme has emerged: It is nearly impossible to arrange for the making of fair and coherent group choices. Preference cycles, agenda manipulation, strategic misrepresentation of preferences, heresthetical maneuvers, and so on frustrate our best attempts. The *coup de grâce*, developed in this chapter, is that "popular sovereignty"—by which we mean any method for allowing individuals in a group to affect their own fates through voting—is not unambiguous either. There are lots of different ways to cast and count votes or "do" majority rule, for instance. If all these methods differed only in the details but not in the final result, then we could relegate the matter of details to politics junkies to chat about. Alas, the devil is in the details. In this chapter, therefore, we explore the procedural context of voting—the rules by which small committees and large electorates make choices.

We partition the discussion into two sections, according to what it is the group is choosing. The first part of the discussion focuses on how relatively small groups—a set of friends, a club, a committee—choose some alternative from a set of available alternatives. We call these arrangements *voting methods*. The second part of the discussion emphasizes how relatively large groups (called electorates) choose a specific thing (called a legislature). We call these arrangements *electoral systems*. In each part of this discussion, we are not so much interested in conveying specific details about the myriad ways of making group choices; rather we want to hammer home the simpler facts that *there are myriad ways of making group choices, that each has merit in some circumstances, but that the rationale for none seems decisive or compelling in all situations.*

VOTING METHODS

The Problem with Methods of Voting—They Matter!

Getting down to business, suppose we have a group of 55 individuals, choosing among five alternataives, $\{a, b, c, d, e\}$.[1] Of the 120 possible complete and transitive strict preference orderings an individual might adopt as his or her own,[2] there are only six distinct orderings, or "opinions." represented in this particular group. They are listed in Table 7.1, along with the number of group members holding each. (The underlining will be explained later.) For the sake of discussion we describe six different "reasonable" ways for the group to arrive at a choice among the five alternatives. The reader may well be able to devise others and should rest assured that, in human history, a multitude of alternative methods have been devised.[3]

Consider, as a first method, the simplest of them all, ***sim-***

[1] This absolutely evil example was invented by Joseph Malkevitch and is displayed in Figure 2 of his article, "Mathematical Theory of Elections," *Annals of the New York Academy of Sciences* 607 (1990): 89-97.

[2] Recall that there are five possible first-preference alternatives, four remaining possibilities for second preferences, and so on—or, 5 x 4 x 3 x 2 x 1 = 120 ways to order five alternatives.

[3] A considerably more systematic treatment of alternative voting methods is found in William H. Riker, *Liberalism Against Populism* (San Francisco: Freeman, 1982), chapter 4. Somewhat more technical, yet quite insightful, is Peter Fishburn, "A Comparative Analysis of Group Decision Methods," *Behavioral Science* 16 (1971): 38-44.

TABLE 7.1
AN EVIL EXAMPLE

I (18)	II (12)	III (10)	IV (9)	V (4)	VI (2)
a	b	c	d	e	e
d	e	b	c	b	c
e	d	e	e	d	d
c	c	d	b	c	b
b	a	a	a	a	a

SOURCE: Joseph Malkevitch, "Mathematical Theory of Elections," *Annals of the New York Academy of Sciences* 607 (1990): 89–97 (Figure 2)

ple plurality voting. Each voter casts a single vote for a single alternative, and the alternative with the most votes wins. This is one of the hallmarks of the Anglo-American system for electing legislators.[4]

A slight variation is the ***plurality runoff,*** in which each voter casts a single vote for a single alternative, and the two alternatives with the most votes move to a second stage in which the balloting is repeated between these two survivors according to simple plurality voting.

An even more intricate and general form of runoff is the ***sequential runoff.*** Here each voter casts a single vote for a single alternative, the alternative with the fewest votes is eliminated, and the balloting is repeated. This procedure continues until there is a single alternative left.

The fourth method we examine allows voters to express preeferences about *all* the alternatives. According to the ***Borda count,*** a scoring system much like that used in international track competitions, each voter expresses personal

[4] When we discuss it in the latter half of this chapter as an electoral system, we will call it *first past the post.*

preferences over the five alternatives by awarding four points to his or her first choice, three to the second choice, two to the third, one to the fourth, and none to the fifth. These points are totaled and the alternative with the most points wins.

Fifth, the ***Condorcet procedure*** seeks to determine whether there is some specific alternative that can secure a majority against each of the others in a pairwise round-robin tournament. If so, that is the winner. If not, then we will need to provide some alternative procedure (perhaps one of the others).

Finally, consider the method of ***approval voting,*** invented by the political scientist, Steven Brams, and the operations research scholar, Peter Fishburn. It puts no limit on the number of votes an individual can cast. Each individual casts votes for all those alternatives he "approves of." This means she may cast votes for all the alternatives if she wishes, none of them, or any number in between. The winner is the alternative that receives the most approval votes.[5] All of the methods are listed in Display 7.1.

With the data of Table 7.1 we can determine how the various forms of popular sovereignty listed in Display 7.1 perform. Voters are assumed to vote sincerely. Display 7.2 provides the results, with the winning alternative given in bold.

The simple plurality method produces a victory for *a*, since it has the most first-preference supporters, though a peculiar victory it is, given that all but the first group *hate* this alternative. This is made all too apparent when we look at the plurality runoff procedure in which *b* triumphs; indeed, *any* alternative that made it to the "finals" against *a* would have

[5] An individual who votes for all the alternatives states that all are "above threshold," or acceptable. The impact on the final outcome is exactly the same as the voter who abstains (or, equivalently, states that she approves of no alternative). What matters in approval voting is that an alternative do well *relative to its competitors*; its absolute vote total is less important.

DISPLAY 7.1
SOME VOTING METHODS

Method	Description
Simple plurality voting	Alternative with most votes (plurality) wins.
Plurality runoff	Top two vote getters move to a second round; new balloting determines second-round winner by simple plurality voting.
Sequential runoff	Alternative with fewest votes is eliminated and balloting repeated; elimination procedure continues until one alternative remains.
Borda count	Alternatives assigned points in accord with voter rank orders. The alternative with the largest sum of points wins.
Condorcet procedure	Pairwise round-robin tournament determines if one alternative defeats each of its rivals.
Approval voting	Each voter casts votes for any alternative he or she approves of. The alternative with the most votes wins.

beaten it. The sequential runoff procedure produces *c* as the final outcome. The Borda count gives the victory to *d*. And the Condorcet procedure shows that alternative *e* receives a majority (28 or more of 55 votes) against every other alternative. In short, each of the first five preference-based methods of group choice yielded a *different* winner. The sixth, approval voting, yielded a tie between *d* and *e*.

So we must conclude that the rules of preference aggregation matter, and sometimes (as in this example) they matter a lot. It is evident in this example that whoever chooses the

DISPLAY 7.2
ELECTORAL SYSTEM RESULTS

Simple plurality:	**a=18**	b=12	c=10	d=9	e=6
Plurality runoff:					
round 1	**a=18**	**b=12**	c=10	d=9	e=6
round 2	a=18	**b=37**			
Sequential runoff:					
round 1	**a=18**	**b=12**	**c=10**	**d=9**	e=6
round 2	**a=18**	**b=16**	**c=12**	d=9	
round 3	**a=18**	b=16	**c=21**		
round 4	a=18		**c=37**		

Borda count:

a=4(18) + 0(12) + 0(10) + 0(9) + 0(4) + 0(2)= 72
b=0(18) + 4(12) + 3(10) + 1(9) + 3(4) + 1(2)=101
c=1(18) + 1(12) + 4(10) + 3(9) + 1(4) + 3(2)=107
d=3(18) + 2(12) + 1(10) + 4(9) + 2(4) + 2(2)=136
c=2(18) + 3(12) + 2(10) + 2(9) + 4(4) + 4(2)=134

*Condorcet:**

	a	b	c	d	e
a	—	18	18	18	18
b	37	—	16	26	22
c	37	39	—	12	19
d	37	29	43	—	27
e	**37**	**33**	**36**	**28**	—

Approval:†

a	18+0+0+0+0+0	=18
b	0+12+10+0+4+0	=26
c	0+0+10+9+0+2	=21
d	**18+12+10+9+4+2**	**=55**
e	**18+12+10+9+4+2**	**=55**

* Reading across each row in the matrix gives the number of votes that each *row* alternative gets when paired against each of the *column* alternatives. Thus, reading across the second row, alternative *b* gets 37, 16, 26, and 22 votes out of 55 against *a*, *c*, *d*, and *e*, respectively.
† Each group is assumed to cast votes for all the alternatives above the line in Table 7.1. Thus, every group but the third votes for three of the five alternatives; the third group approves of four out of the five.

method of counting noses determines, finally and decisively, the final outcome. Put somewhat differently, whatever procedure is in place for choosing the method determines the final outcome. But putting procedures in place is the business of constitutions; they say either directly what method will apply or indicate who (or what body) gets to decide. No wonder constitutional politics are such struggles! So much is at stake. And while the example above provides something of a worst-case scenario (what with each method producing a uniquely different outcome), the reader should understand that it is not an altogether extreme instance. As the section title says, the problem with voting methods is that they matter!

Thinking about Voting Methods

At present there is no generally accepted way to think about voting methods. There are so many different ways to vote, and so many potentially useful criteria to bring to bear on alternative systems, that it is easy to become quickly confused. Here we want only to suggest a couple of directions for thought. Generally speaking, a voting method may be thought of in terms of (1) the inputs required, (2) what the procedure does to those inputs, and (3) the output or outcome produced. That is, the final *outcome* is a *function* of the *inputs* (written: outcome = F(inputs)), and we can think about each of the three italicized components separately.

1. In terms of inputs, plurality voting makes the simplest demands on voters (and, perhaps, for this reason, it is a commonly used method of group choice); each person must simply name an alternative (his or her first preference if a sincere voter; something else otherwise). On the other hand, the demand on voters in a runoff plurality election depends upon how it is administered. If voters are expected to show up for two separate rounds, then less information at each round is required (but the cost of showing up both times is greater). In

round one the same data as in a simple plurality contest is required. In round two, relative preference between the two highest vote-getters in round one is needed. If voters wish to economize on trips to the polls, showing up only once, then they must provide more information on that one occasion. Specifically, they must provide information on every possible pairwise comparison (if there are five alternatives, then there are ten comparisons), since it is not known in advance *which* pair will advance to the second round. However, all this comparison data is contained in a voter's preference ordering, so that's all a voter need provide at the outset. This is precisely the same data required for a sequential runoff election, a Borda count procedure, and a Condorcet procedure, too. Approval voting requires as much information about preferences as each voter wants to reveal.[6]

Allowing these six methods to stand for the many hundreds of voting methods that have sprung from human creativity, the point we wish to make here is that methods differ as to what they require—a single alternative, a subset (of whatever size each voter wishes), or an entire preference ordering. There may be grounds for preferring one method over another, quite apart from the particular result each may yield, based on the ease of administering it or on the desire to economize on the burden of the voter. On the other hand, the necessary inputs may depend on what it is you want to get out of the group choice—something we discuss below. Wherever one stands on these or a host of other criteria for thinking about inputs, it is patently clear, we hope, on the basis of required inputs alone, that democratic voting, broadly understood, takes on a multiplicity of forms.

2. We won't spend much time on the procedures themselves

[6] That is, the voter can either submit a subset of the full set of alternatives (the "approved of" alternatives) or, as in Table 7.1, he or she can hand in a preference ordering with a line drawn below the "approved of" alternatives.

and what they do, since we described a few of them already in Display 7.1. We do want to point out, however, that these procedures (and any others you can think up) have their peculiarities. Plurality rule, for instance, is especially odd. Alternative *a* was the plurality winner in the example above, yet *it loses to every other alternative in pairwise comparison.* Additionally, the Condorcet winner, *e*, which many would take as a strong normative candidate for the group choice, actually got the *fewest* votes in the plurality contest. Runoffs, whether simple or sequential, have the perverse possibility of eliminating an alternative that can beat every other in a pairwise contest (*e* never made it very far in these runoffs). The Borda count method (indeed, this is true of all the methods) is very vulnerable to strategic behavior. Notice that the twelve voters of group II or the ten voters of group III in Table 7.1, who prefer *e* to *d*, can actually give *e* a victory by misrepresenting their preference ranking for *d* (pretending it is lower in their ranking).[7] The Condorcet procedure does not always produce a winner—and then what do you do?

Finally, there is an issue that applies to each of the voting methods we have described, but we will discuss it in terms of approval voting, since its proponents seem so unperturbed by it.[8] In the example above there are five alternatives. Those alternatives might be various motions (say, what movie the fraternity house should rent this evening) or candidates (say, which of the sorority sisters should be the representatives on the Greek Council). However, which motions are moved, or which candidacies are activated, depend intimately on the voting method, *F*. You might figure, for example, that *Forrest Gump* would get a lot of second-choice votes from your frat brothers, and thus have a good chance of winning if the voting method were the Borda count. But you also believe it wouldn't get many first-preference votes, so you probably wouldn't even bother proposing it if the decision rule were plurality voting. To look at a *fixed* set of alternatives, and compare the outcome under plurality against that of approval or the outcome under Borda against that of Condorcet, for example, misses this point. In the jargon of the field, the set of alternatives or candidates is *endogenous*—that is, highly dependent on the method of counting heads. In this regard, our hunch is that approval voting encourages a larger number of alternatives to come forward (so to speak) than many other voting methods. Candidates know they do not have to be the top choice of a voter, but merely *among* the "approved of" set, and thus may find it easier to rationalize their prospects of victory. Likewise, a motion need not be the favorite of many voters, but only *among* the favorites, in order to prosper under approval voting, a fact that may give encouragement to potential motion-makers. So, the question comes down to whether it is better or worse for a group to have a rich set of items from which to choose or a more spartan set. Is more always better than less? This question will take on an interesting political significance in the second part of this chapter when we examine proportional representation versus other systems for electing legislatures.

3. It may seem odd that we even need to discuss the output of a voting method, since it is no more than the thing that is chosen. But exactly what is that "thing"? We have somewhat abstractly described alternatives by letters, suggesting that the thing the group must choose is some unitary entity, some element of the set $\{a, b, c, d, e\}$. But we can quickly complicate matters quite a bit. Shortly, for instance, we will talk about ways of choosing members of a legislature. The "thing" here

[7] M. Borda, a member of the French Academy of Sciences in the late 18th century, was informed by one of its other members, M. Condorcet, that the Borda count could be manipulated. Borda is reported to have sniffed, "My method is only for honorable men."

[8] See, especially, the various writings of Brams and Fishburn on the merits of approval voting, including "Approval Voting," *American Political Science Review* 72 (1978): 831–847, and their book *Approval Voting* (Boston: Birkhauser, 1983).

could be a single legislator, a group of legislators from a multimember district, or the legislature in its entirety. Or, to give an example with a different emphasis, imagine that what we, as a group, are doing is choosing *instructions* to give our agent. Imagine that our agent must choose among the five letters for us, but she will not know in advance (nor will we) whether all five are available or only some subset thereof. She will need to know more than the group's favorite, since that alternative may turn out to be unavailable. Thus the group, in this instance, needs to choose a *collective preference ordering* by which it wishes its agent to be guided. The point here is that one set of criteria to evaluate a voting method is relevant when all the group needs to do is choose a letter, but an entirely different set may come into play when the group needs to come up with a full preference ordering. The nature of the output, then, should affect the way we think about voting methods.

This entire discussion produces a serious philosophical puzzle. If the "wish of the group," or the "collective will," or the "public interest"—whatever you want to call the output of group deliberation—is to be ascertained from the inputs individual group members bring to the voting method, and those inputs vary from method to method, then how are we to give meaning to "wish of the group," "collective will," or "public interest"? Using one method may yield one conclusion about what the group wants, while using a different method yields a different conclusion because it operates on some new set of alternatives. This is crazy! But it is even worse. Suppose, as in the example associated with Table 7.1, the alternatives under consideration remain constant, and that voter preferences have also remained fixed. Then, with alternatives and preferences fixed, it seems only natural to presume the wish of the group is well defined—it is whatever it is. The only thing that might change is the way in which those wishes are revealed or ascertained by the voting method. Yet we have seen that the outcome *does* change (most evilly displayed in the example

at the beginning of this chapter). But surely it seems perverse to conclude that the group wish has changed just because the method we have used to ascertain it—and *only* the method—has changed. Are we driven to this conclusion? We ask the reader to think hard about this question, for it has motivated much of the discussion of the last five chapters. We know of no definitive answer to it, although we shall examine some in the summary that follows this chapter.

We leave this part of our discussion in a woefully incomplete state. Our principal purpose, however, has been less to provide a broad summary of the myriad methods of conducting group voting than to convey by illustration a sense of fragility in group life. The decisions a group reaches, as the last few chapters have suggested, depend not only upon the options made available, not only on the order in which some agenda setter presents them, not only on the degree to which group members reveal or misreveal their preferences, but also on the way they conduct the actual decision making. And all those other things, likewise, are influenced by the voting method we adopt. A group decision surely reflects member preferences. *But it also reflects much more*, a theme to which we will return in our summary.

Electoral Systems

Just as there are many voting methods, there also is an incredible variety of electoral systems. We will restrict ourselves to systems for electing legislatures, using these institutions to represent a broad class of elected governance arrangements.[9] We claim here that electoral systems may be thought of in terms of the degree to which their "core value" is *representa-*

[9] In a sense, a legislature *is* a generic elected institution. For example, a president, governor, or mayor may be thought of as a one-person legislature, an elected court as a one- (or more) person legislature, and so on.

tion or *governance*. By the former we mean an electoral arrangement that places priority on the degree to which the elected reflect (or represent) the beliefs and preferences of the electors. By the latter we mean an arrangement yielding elected representatives capable of acting decisively, of governing. Obviously, both of these purposes are noble. Yet, they often operate at cross-purposes because an arrangement that emphasizes representativeness may make governance more difficult, and vice versa.[10]

Associated with the end of the spectrum giving priority to representation is the broad family of electoral methods known as *proportional representation (PR)*. At the other end of the spectrum is the family of *plurality voting* methods (to which we referred earlier in this chapter). We begin our discussion with plurality methods that are common in the United States and Great Britain. This will be followed by a discussion of the more exotic PR methods found in continental Europe. We conclude by suggesting why these broad classes of electoral system are seen as either representation-oriented or governance-oriented.

General Remarks

One of the leading contemporary students of the theory of electoral systems, Gary Cox, has defined an electoral system in terms of five bits of information.[11] For Cox, as for us, the critical separation is between plurality and proportional systems, but five bits of information can be used to characterize each.

[10] For an elaboration on this theme, with special reference to the U.S. Congress, see Kenneth A. Shepsle, "Representation and Governance: The Great Legislative Tradeoff," *Political Science Quarterly* 103 (1988): 461–484.

[11] Two especially important papers by Cox are mathematically advanced, but well worth examining, if only to get a feel for the kinds of analysis scholars like Cox are able to do. See his "Electoral Equilibrium under Alternative Voting Institutions," *American Journal of Political Science* 31 (1987): 82–109; and "Centripetal and Centrifugal Incentives in Voting Systems," *American Journal of Political Science* 34 (1990): 903–936.

These five describe the resources given to voters, what the voters can do with their resources, and finally, how the electoral formula produces a final outcome—in effect, the inputs, procedures, and outputs we discussed earlier. Generally speaking, Cox maintains the following distinction between plurality and proportional arrangements.[12]

> By a plurality formula, I mean one in which voters cast votes for individuals (rather than party lists) and the top . . . vote-getters win seats. . . . Proportional formulas, on the other hand, are those in which voters vote for parties, and seats are allocated in proportion to the vote polled by each party. . . .

The five bits of information are v (number of *votes* per voter); p (if $v > 1$, whether voters must cast all v votes, or may *partially abstain*); c (if $v > 1$, whether voters may *cumulate* their votes, or must distribute them); k (the number of legislators to be elected per district, known as the *district magnitude*); and f (the *electoral formula*). Electoral systems can be represented by this information, and Display 7.3 lists some common plurality types.

Plurality Systems: First (or More) Past the Post

The most famous of the plurality systems is single-member districts and first-past-the-post (FPP). As Display 7.3 describes, each voter gets one vote, may cast it for any candidate he or she pleases, and the single candidate with the most votes (not necessarily a majority) is elected. The legislature thus consists of legislators elected from separate districts in this manner. This is the electoral system found in Great Britain and many of her former dependencies (including, of course, the United States).

The key feature, it seems to us, is that each district, or constituency, gets but a single representative. This may

[12] "Centripetal and Centrifugal Incentives in Voting Systems," 905–906.

DISPLAY 7.3
ALTERNATIVE ELECTORAL SYSTEMS

	v	p	c	k	f
First Past the Post (FPP)	1	no	no	1	Plurality
Single-Nontransferable Vote (SNTV)	1	no	no	$k>1$	Plurality
Limited Vote (LV)	$<k$	yes	no	k	Plurality
Cumulative Vote (CV)	$\leq k$	yes	yes	$k>1$	Plurality

mean—and this is in fact a common complaint with the system—that the winner is not particularly representative of the district in which he or she is elected. The district may be 60 percent male and 40 percent female; whoever is elected will not represent (in the sense of "reflect") a rather sizable chunk of the electorate. Another (melting pot) district may be 25 percent Roman Catholic, 23 percent Greek Orthodox, 16 percent Jewish, 15 percent Baptist, 11 percent Episcopal, and 10 percent agnostic. Its representative will not share religious and cultural traditions with at least three-fourths of the constituency. If the district magnitude were more generous—if k were larger—then it would be possible to represent more of a district's heterogeneity. But it would also mean a larger legislature, for one thing, and one almost certainly with a greater heterogeneity of views. This might make it more difficult for representatives to *govern*—to debate, deliberate, form coalitions and compromises, and ultimately come to some conclusion on public policy issues facing the society. This is the "great tradeoff" between representativeness and governance.

Oddly enough, an ethnically homogenous society, Japan, until very recently employed a plurality system with a larger district magnitude. It is identified in Display 7.3 as the method of single nontransferable vote (SNTV). In this arrangement

each voter is still endowed with but *one* vote, but now the k highest vote-getters are elected from the district, where k is given in advance as the district's magnitude.

Just as SNTV is but a small alteration of FPP (namely, a change in district magnitude from 1 to some $k > 1$), the method of limited vote (LV) is only a slight alteration of SNTV. Specifically, LV endows each voter with more than one vote (but fewer than k), still allowing for k winners. If, under FPP, *one* representative is to be elected from a district and each voter casts *one* vote, and, under SNTV, *more than one* representative is to be elected from a district and each voter (still) casts *one* vote, then, under LV, *more than one* representative is elected from a district and each voter may cast *multiple* votes. Thus, a district may elect four legislators under LV by giving each voter, say, two votes. This method, however, offers an additional strategic maneuver to voters—"plumping" (as the English called it in the eighteenth century), or voting only for your favorite candidate.[13] Consider the district with $k = 4$ and $v = 2$ as just mentioned. Suppose one of us is considering casting his votes for his two favorites in the field of candidates. The latest public opinion poll showed his second-choice candidate running in fourth place and his favorite candidate in fifth. If he proceeds to vote for both candidates, his second choice may win, but just possibly at the expense of his first choice, who will finish just out of the running. He might be better advised to cast only a single vote for his first choice, thus foregoing the support he had planned to give to his second choice, but just possibly helping to elevate his favorite into

[13] In *Safire's Political Dictionary* (New York: Random House, 1978) the following entry is found under *plump*:

"One of the English election phrases for which there is no equivalent in the United States," wrote the *New York Tribune* in 1880, "is 'plumping.' Whenever [an English] constituency returns two members, each voter can give one vote each to any two candidates but he cannot give his two votes to any one candidate. If he chooses he can give one vote to only one candidate, and this is termed 'plumping'."

fourth place. The LV method permits this sort of *partial abstention.*

Cumulative voting (CV) goes one step further, permitting voters not only partially to abstain, but also to *cumulate* their votes.[14] In the example of the previous paragraph, where the voter has two votes and there are four candidates to be elected, but the method is now CV, he could cast *both* of his votes for his first-preference candidate. The state of Illinois, until 1980, elected state legislators from multimember districts via CV. Similarly, nearly half of the states permit and more than one-fourth actually *require* CV in elections for boards of directors of publicly traded corporations. The idea behind CV is that well organized minorities, by cumulating their votes, can assure themselves a modicum of representation. Cumulative voting in multimember districts, for example, has become one of the "methods of choice" for a number of civil rights activists for electing county commissioners, school boards, and other local officials, primarily in the south, in order to assure minority representation.[15]

Equilibrium in Plurality Systems

Scientists and engineers can usually make intelligent remarks about and comparisons among alternative engines. The same is true here regarding the electoral engines we have just described; thus far, however, man has been much more creative in *inventing* electoral engines than in *understanding* their operating characteristics.

[14] Safire, in the entry referred to in the previous footnote, goes on to indicate that the American usage of the term "plump" differs from the English. In the United States, "to plump" means to cumulate your votes for a particular candidate.

[15] For various points of view, see Lani Guinier, "The Representation of Minority Interests," and Kenneth Benoit and Kenneth Shepsle, "Electoral Systems and Minority Representation," both in Paul Peterson, ed., *Classifying by Race* (Princeton, N.J.: Princeton University Press, 1995), pp. 21–49 and 50–84.

The key to understanding the equilibrium tendencies of the alternative plurality systems we have described is the *number of candidates in a district.* But before developing this thesis, we first must say what we mean by *equilibrium* in this context. We restrict analysis to one-dimensional spatial representations (like those described in the first part of Chapter 5). Each voter is assumed to have a unique ideal point and single-peaked preferences on the one dimension. Candidates compete for votes by identifying with specific locations on the dimension. Qualitatively, we consider two equilibrium tendencies. A *central tendency* is one in which the candidates tend to converge on the median voter's ideal point location. A *dispersed tendency* is one in which the candidates tend to distribute themselves along the dimension, adopting distinctive policy positions. The former is dominated by centripetal forces in which electoral competition drives candidates toward one another, while the latter is dominated by centrifugal forces which drive candidates away from one another in order to distinguish themselves from one another.

Cox shows how the number of competing candidates is a key parameter in determining whether centripetal or centrifugal incentives dominate. Cox cuts the cases according to whether the system permits cumulation of votes or not (see the column labeled c in Display 7.3). He then shows that when cumulation is not allowed (c = no), "if the number of candidates competing for election is small enough relative to the number of votes per voter [v in Display 7.3], then centripetal forces will dominate (in the sense that equilibria will be centrist); but if the number of candidates is large enough, then centrifugal forces become strong enough to create a certain amount of dispersal in equilibrium. When cumulation is allowed (c = yes), then "centrifugal forces will always dominate."[16]

[16] "Centripetal and Centrifugal Incentives of Electoral Systems," 912. Cox makes "small enough" and "large enough" precise.

In principle, the reader could take a copy of Cox's paper (in which more carefully and precisely stated propositions may be found) and a governmental handbook describing a specific electoral system and, on the basis of these two sources of information, make forecasts about how candidates will actually distribute themselves politically in election contests. Cox illustrates this for the case of Japan, a country whose electoral system (at the time Cox wrote) was described by SNTV:

> Most of the 511 members of the Japanese House of Representatives are elected in districts of magnitude 3, 4, or 5 in which each voter has a single nontransferable vote. . . . In all of Japan's multimember districts [the number of candidates is large relative to the number of votes per voter]. This means that there is never a median voter or central clustering result predicted for Japan. Instead, the dispersion result applies and predicts that, if there is an equilibrium, then candidates will not be bunched together anywhere along the left-right spectrum; and some candidates will adopt [extreme positions].[17]

Social scientists are not rocket scientists—that's already been conceded. But we are growing increasingly sophisticated about how various procedures of social choice actually work. In this section we have done no more than to illustrate a small fraction of the rich class of plurality-like electoral systems and how they may be analyzed. Electoral systems can be boiled down to a relatively small number of parameters and equilibrium analysis conducted on them. Depending upon the resources given a voter (particularly, the number of votes), how those resources may be deployed (particularly, whether they can be cumulated and whether they may be only partially used), and the nature of the task (particularly, the number of legislators to be elected), it is possible to determine whether candidates have incentives to cluster centrally or disperse

[17] "Centripetal and Centrifugal Incentives of Electoral Systems," 915.

themselves. This, in turn, tells us something about the kinds of legislatures these electoral systems produce.

We should remember, however, that not all electoral systems are horse races among individual candidates. A large class of systems operates at a more highly aggregated level—one in which parties, rather than individual candidates, are the strategic players. Legislative representation in these systems are determined by the proportion of the popular vote each party receives. We shall look at these systems before comparing them to the horse-race variety. Even before doing that, however, we will look at one last interesting plurality-like system that would seem to defy Cox's classification scheme.

A Most Unusual Plurality System: The Single Transferable Voter (STV)

The single-transferable-vote system, sometimes called the *Hare system* after one of its early students, differs from other multimember plurality systems in that each voter essentially reports his or her entire preference ordering over the candidates. Riker describes it as follows:

> The rule for the *single transferable vote method* is: For districts with S seats and m candidates ($m \geq S$), the voters, V in number, mark ballots for first choice, second choice, . . . , and m^{th} choice. A quota, q, is calculated thus:
>
> $$q = V/(S + 1) + 1$$
>
> and q is rounded down to the largest integer contained in it. If a candidate receives at least q first-place votes, he or she wins, and any surplus votes (i.e., the number of first-place votes in excess of q) are transferred to nonwinning candidates in proportion to the appearance of these candidates in next place on all ballots of the initial winner. Another candidate who then has q first-place [plus] reassigned votes wins, and his or her surplus is transferred to the next nonwinning candidate on his or her supporters' ballots (again in proportion to their appearance in next place) and so on

until all seats are filled. If at any point in the process (including the beginning) no candidate has q first-place and reassigned votes, the candidate with the fewest first-place and reassigned votes is eliminated and all the ballots for her or him are transferred to candidates in the second (or next) place on those ballots; and this is repeated until some candidate has q votes.[18]

To make all this concrete, suppose there were 100 voters ($V = 100$) in a district charged with electing 3 representatives ($S = 3$). The quota—known as the Droop quota (Mr. Droop was a friend of Mr. Hare)—is $q = 100/(3 + 1) + 1 = 26$. That is, any candidate receiving 26 votes can be assured that no more than two other candidates can get as many as she.[19] If, in fact, a candidate got in excess of 26 first-preference votes, and all the remaining candidates did not, then the preference orderings of the excess voters are consulted for the second preference listed and those votes are distributed to them. If this pushes some other candidate over the 26-vote quota, then he or she is deemed elected. This continues until all three candidates have been elected or until fewer than three have been elected and no remaining candidate has the quota. In this case, the process starts eliminating candidates, starting with the one with the fewest total votes. All of that candidate's votes are distributed to the candidates named second on each ballot. This continues until all seats are filled.

The STV method is used to elect the parliament in Ireland (called the *Dail*) and the city council of the authors' city of

[18] Riker, *Liberalism Against Populism*, p. 49. Duncan Black refers to STV as a system of *proportional representation*, because it tends to approximate the representativeness that many PR systems display. See his famous *Theory of Committees and Elections* (Cambridge, Eng.: Cambridge University Press, 1958). But it is decidedly a plurality system in that election is dependent on getting more (first-preference plus transfer) votes than other candidates.

[19] If *three* other candidates got at least 26 votes, then they would jointly have 78 votes, which would mean that, together with the candidate who already has 26 votes, 104 votes had been cast. But this cannot be, since only 100 votes can be cast. Therefore, it also cannot be that more than three candidates receive 26 or more votes each.

Cambridge, Massachusetts. In Ireland there are 41 districts from which 166 members of parliament are elected; typical district magnitudes are 3, 4, and 5. The City of Cambridge elects a city council (nine members) from one at-large district. In both locations, there is much local lore about strategic behavior, as candidates campaign not only among their own supporters, but also among opponents' supporters. The purpose for a candidate is both to energize his own supporters as well as to try to get listed high up in the preference orderings of voters supporting other candidates, the latter in order to benefit from potential "reassigned" votes. Cambridge has refused to computerize its operations so that it takes nearly two weeks following an election to allow tens of thousands of paper ballots to be counted and recounted; many a local political junkie socializes at election headquarters during this period, cheering as one candidate or another surpasses the Droop quota and claims a council seat.[20]

We conclude by noting that a version of this method is used to elect members of the parliament in Australia. In that country, however, districts have a magnitude of only one, so the method is given a different name just to confuse everyone (naturally); it is known there as the *alternate vote*. With $S = 1$, the quota formula above becomes $q = V/2 + 1$. With 100 voters a candidate needs 51 first-preference plus reassigned votes to be declared the winner.[21]

[20] For those interested in learning more about this system, he or she should consult Gideon Doron, "The Hare Voting System is Inconsistent," *Political Studies* 27 (1979): 283-286; and Gideon Doron and Richard Kronick, "Single Transferable Vote: An Example of a Perverse Social Choice Function," *American Journal of Political Science* 21 (1977): 303-311.

[21] STV is also used to elect members of the Faculty Council of Harvard University. (It is something of an historical irony that it was recommended to the university by Kenneth Arrow while he was a member of the faculty!) It runs smoothly, virtually without controversy, and certainly with no strategic behavior, because so few faculty members offer themselves as candidates.

Proportional Representation

We will be analytically less precise about PR systems because, quite frankly, not nearly so much work has been done on them. The purpose of a proportional system is to produce a legislature that mirrors, in some fashion, the larger society. If, for example, the main cleavages in society are ethnic (as in many of the emerging democracies in Eastern Europe), or religious (Northern Ireland), or linguistic (Belgium), then a PR system will tend to reproduce them inside the elected legislature. Under most such systems, no stratum, unless especially small or poorly organized (or just plain stupid) is highly underrepresented.

One would think that the design of a PR system is straightforward. Let each citizen cast a single vote for his or her favorite party (or any other, for that matter). Add up the votes for each party. Give each party a proportion of legislative seats exactly equal to its proportion of the total popular vote. Voila!

Not so fast! First of all, the legislature's size is typically fixed in advance. For most "reasonably" sized legislatures, it is typically not possible to translate electoral proportions evenly into seat proportions. Suppose the Beer Lovers Party (an actual party in the 1992 Polish elections) captured 1 percent of the popular vote. How many seats should it receive in the 450-seat *Sejm* (the Polish House of Representatives)? It cannot be give the 4.5 seats to which it is entitled according to its electoral proportion. Most PR schemes—and there are actually quite a large number of them—differ primarily on how they handle the problem of allocating these "fractional seats."

A second issue involves exactly who should get elected to the legislature in the first place. If the Polish Beer Lovers Party receives one percent of the popular vote, as in the previous paragraph, should it get *any representation at all*? If the answer is in the affirmative, then at what point does the answer change to negative—0.5 percent? 0.25 percent? 0.10 percent? Where do you draw the line? PR systems differ quite dramatically on this matter. In practice most require that a party receive at least a specific minimum popular vote proportion before it is entitled to *any* parliamentary representation. This parameter, known as the *threshold*, varies dramatically from country to country. Poland (in its first few democratic elections) and Israel (throughout its democratic history) have had very low thresholds. For the 450-seat *Sejm*, a Polish party must obtain 1/450th, or 0.22 percent, of the popular vote to be awarded a seat. For the 120-seat *Knesset*, an Israeli party (until recently) needed 1/120th, or 0.83 percent, to win a seat. Germany, on the other hand, has a very high threshold; a party must obtain 5 percent of the vote before it qualifies for seats in the *Bundestag*. Thus, in the 1987 elections, the Green Party won 8.7 percent of the popular vote and was awarded 42 seats in the 497-seat *Bundestag* (8.45 percent of the seats); in the 1991 election their popular vote percentage fell just a hair below 5 percent *and they lost their entire legislative representation*. Clearly, high thresholds make for more disproportionality.[22]

Representation versus Governance: PR v. Plurality

There are few results to report on equilibrium properties of various PR systems, so we turn in this concluding section to a brief comparison of the two broad families of electoral system we have been discussing. In all our discussion we have been

[22] For a full discussion of this and related issues, the reader may wish to consult Rein Taagepera and Matthew Soberg Shugart, *Seats and Votes: The Effects and Determinants of Electoral Systems* (New Haven, Conn.: Yale University Press, 1989). For an interesting recent discussion on PR, consult the debate between Cox and Gallagher in the pages of the leading journal on electoral systems: Michael Gallagher, "Proportionality, Disproportionality, and Electoral Systems," *Electoral Studies* 10 (1991): 33-51; and Gary W. Cox, "Comment on Gallagher's 'Proportionality, Disproportionality, and Electoral Systems'," *Electoral Studies* 10 (1991): 348-352.

vague about the actual number of candidates that compete for legislative seats. Indeed, we have not ventured a conjecture about whether many or few candidates compete and, more significantly, whether or not the electoral system has anything to do with this.

In fact, there is a long literature on this very subject, the most famous proposition of which is known as *Duverger's Law*, named after the famous French political scientist who dared to call it an empirical regularity. Duverger's Law comes in two parts. The first, for which there is both strong argument and evidence, states that first-past-the-post, single-member-district systems are strongly associated with two-party (or two-candidate) competition. The idea here is that third parties and third candidates (or both) will ordinarily be loathe to enter the race because they have so little chance of winning; in turn, they have so little chance of winning because neither voters, nor campaign consultants, nor campaign contributors are likely to waste their votes, time, and money, respectively, on hopeless candidacies.[23] The second part of Duverger's Law, for which there is ample empirical support but less compelling analytical argument, states that PR systems are associated with multiparty competition.

As an analytical claim, it seems to us that the kernel of truth here is that districts in which there are, by design, a very small number of winners—only one in first-past-the-post; exactly k in k-past-the-post—discourage independent political entry and encourage cooperation, coordination, coalition, and merger-like political activity *before* elections. In districts where there are many possible winners, as in most PR systems (especially those with a very low threshold) and even in those k-past-the-post systems where k is quite large, independent political entry is encouraged and various forms of cooperation, coordination, coalition and merger-like political activity are ei-

[23] See Case 6.5.

ther discouraged altogether or deferred until after elections. First-past-the-post systems typically, and other "small" k-past-the-post systems often, resolve many conflicts *before* legislative politics commences. PR and "large" k-past-the-post systems, on the other hand, defer this kind of conflict resolution until the legislature convenes. Thus, parliamentary political conflict tends to be more muted and centrist in legislatures elected by FPP; indeed, there is typically a single majority party that can get on with the business of implementing its agenda. Legislatures elected by PR *reflect* rather than *resolve* political conflict in advance, depending upon post-election parliamentary politics—coalition government, for example—to discover the means for resolution.

It should not be surprising, then, that a number of democratic regimes seek to obtain the best of both worlds by implementing a "mixed" method. In Germany, for example, half of the *Bundestag's* members are elected from single-member districts according to first-past-the-post plurality. The remaining half is elected by party-list proportional representation so that the entire legislature is approximately proportional (subject, of course, to that high 5 percent threshold). In 1993, both Italy and New Zealand underwent changes in electoral law. Like ships passing (but not quite passing) in the night, New Zealand deserted plurality for an approximation of the German mixed method, while Italy deserted PR for the mixed method. The mixed method appears to enhance governance by keeping the number of parties relatively small, on the one hand, while maintaining a modicum of representativeness, on the other.