Political Economics

Explaining Economic Policy

Torsten Persson and Guido Tabellini

The MIT Press
Cambridge, Massachusetts
London, England

8 Electoral Rules and Electoral Competition

Our exploration of comparative politics begins in the world of preelection politics. In our discussion, electoral competition takes place between two opportunistic and rent-seeking candidates. In the electoral campaign these candidates make binding promises of how much they will tax voters and how they will spend the revenue. They can choose to please all voters, promising to levy low taxes or to supply a public good benefiting everyone. They can please some voters but not others, targeting redistribution to some specific groups. Or they can please no one and appropriate rents for themselves. How do politicians, competing to win the election choose among these alternatives? And—most importantly—how does the electoral rule shape their choices? These are the questions addressed in this chapter.

We can draw extensively on the models and results of several earlier chapters. As in chapter 7, we use a model with probabilistic voting to handle electoral equilibrium in cases where policy is inherently multidimensional. As in chapter 4, we allow for endogenous rents in addition to the traditional assumption of pure office motivation. The chapter focuses on the electoral rule, however, drawing on very recent work by Persson and Tabellini (1999b) and Lizzeri and Persico (1998).

In particular, we contrast multiple-district and single-district elections. Throughout the chapter, we shall loosely refer to single-district elections as proportional and to multiple-district elections as majoritarian. This is different from the more precise labeling in the literature on comparative politics in political science. There proportionality refers to another concept, namely to the electoral formula deciding on how votes translate to seats in individual electoral districts, more-proportional systems having a ratio of seat shares to vote shares closer to 1 for every party, compared to majoritarian systems. Given our perspective of policy selection from the viewpoint of society as a whole, however, we think it makes sense to refer to larger districts as being associated with more proportionality. Clearly, proportional representation requires that each district have more than one member. Empirically, one also finds a strong correlation between district magnitude and the electoral formula: one common form of elections combines single-member districts with plurality rule, whereas countries with proportional representation usually have relatively large districts. The reader is free, however, to disagree with our labeling, which actually conflates two different concepts. Naturally, electoral systems differ in many other respects: some have two (or more) tiers of electoral districts, some set minimum thresholds to obtain representation, and so forth.[1] Unfortunately, the simple model of two-party electoral competition that we will use in the chapter cannot capture these aspects well.

1. See Taagepera and Shugart 1989, Lijphart 1994, and Cox 1997 for extensive and illuminating treatments of the different dimensions of electoral systems as well as an account of the variation in real-world electoral systems across time and countries.

Under all electoral rules, the political equilibrium of this model implies that public goods are underprovided relative to the social optimum, that politicians earn positive rents, and that influential voters benefit from targeted redistribution. The electoral rule does make an important difference, however. Majoritarian elections induce politicians to pay most attention to voters in marginal electoral districts. Compared to proportional elections, this induces less public-good provision (since the benefits of the public good for voters in the other districts are disregarded) but more concentrated redistribution targeted toward the marginal districts. At the same time, however, majoritarian elections also reduce rents for politicians. Because voters in the marginal districts are more mobile, electoral competition is stiffer, and voters punish politicians more severely for wasteful spending. Thus from a positive point of view, the theory has clear predictions for how the electoral rule affects the composition of spending. The normative implications are ambiguous, however, as proportional elections induce more public-good provision but also more rents for politicians.

Clearly in this model we can capture only a few possible effects of different electoral rules. In particular, we hold the party structure fixed, ignoring convincing theoretical arguments (as in Cox 1987a, 1990a, 1990b and Myerson 1999) and pervasive empirical evidence (as in Lijphart 1994 and Taagepera and Shugart 1989) for a larger number of parties under proportional elections. Our excuse is pragmatic: we simply do not know how to analyze multidimensional policy consequences of electoral competition in a multiparty setting.

In the next two sections, we formulate the basic model and discuss the political trade-offs it entails. In sections 8.3 and 8.4, we study the equilibrium under proportional (single-district) and majoritarian (multidistrict) elections, respectively. Section 8.5 compares equilibria under proportional and majoritarian elections in a different economic model with local public goods and broad redistributive programs. This comparison provides an additional perspective on our basic comparative politics question, that is, which kind of policies we should expect which kind of electoral rules to favor.

8.1 The Economic Model

Consider a society with three distinct groups of voters, denoted $J = 1, 2, 3$. Each group has a continuum of voters with unit mass. Preferences over government policy are identical for every member of group J and given by the quasi-linear utility function

$$w^J = c^J + H(g) = 1 - \tau + f^J + H(g). \tag{8.1}$$

Here, c^J is the private consumption of the average individual in group J, τ is a common tax rate, f^J is a transfer targeted to individuals in group J, and g is the supply of a (Samuelsonian) public good, evaluated by the concave and monotonically increasing function $H(g)$. Thus we assume that income gross of taxes is equal to 1 for all individuals, that taxes are nondistorting, and that the tax rate is the same for every group.

The public policy vector $\mathbf{q}$ is defined by

$$\mathbf{q} = [\tau, g, r, \{f^J\}] \geq 0,$$

where all components are constrained to be nonnegative. Any feasible policy must satisfy the government balanced budget constraint

$$3\tau = \sum_J f^J + g + r. \tag{8.2}$$

The component r reflects (endogenous) rents to politicians and is a deliberate object of choice. As discussed in chapter 4, we can think of r as an outright diversion of resources, such as corruption or party financing, or more generally, as an allocation of resources benefiting the politicians' private agenda but appearing as an inefficiency for the voters. From the voters' viewpoint, these rents thus constitute a pure waste. As in chapter 4, rent extraction is associated with some transaction costs $(1 - \gamma)$, such that only γr benefits the politicians.

To make the public finance problem more interesting, we could extend the model with economic behavior affected by the policy, say, a labor supply choice distorted by taxation. Below we comment on how our results would change in this richer formulation. But even the simple model at hand entails a very rich micropolitical problem. It involves three conflicts of interest: that between different voters (over the allocation of redistributive transfers, $\{f^J\}$), that between voters and politicians (over the size of rents, r), and that between different politicians (over the distribution of these rents among themselves). As we shall see in this chapter (and the next), different electoral rules alter the intensity of and the interaction between these conflicts, basically by inducing more or less competition between politicians or voters.

The socially optimal policy here entails $r^* = 0$ and g^* such that $H_g(g^*) = \frac{1}{3}$. Redistributive transfers are indeterminate, because of our assumptions that taxes are lump sum and the marginal utility of consumption is constant. But with identical voters, concavity of the utility of private consumption would imply equal transfers f^{J*} to all J in the social optimum (see problem 1 of this chapter). Moreover, a tiny amount of tax distortion would imply no transfers whatsoever: $f^{J*} = 0$ for all J.

8.2 The Politics of Electoral Competition

Before the elections, two parties or candidates (A and B) commit to policy platforms, $\mathbf{q}_A$ and $\mathbf{q}_B$. They act simultaneously and do not cooperate. The winning party's platform is implemented. As we emphasize below, the precise conditions for winning depend on the electoral rule. Both parties are rent-seeking. Thus as in chapter 4, when announcing its policy platform, party P maximizes the expected value of rents, namely

$$E(v_P) = p_P \cdot (R + \gamma r), \tag{8.3}$$

where R denotes the (exogenous) ego rents associated with winning the elections, and p_P denotes the (endogenous) probability that P wins the right to set policy, given $\mathbf{q}_A$ and $\mathbf{q}_B$.

We assume probabilistic voting. Thus as in earlier chapters, the election outcome is uncertain when platforms are chosen, and different voters evaluate the ideological or personal attributes of these parties in different ways. Specifically, let $W^J(\mathbf{q})$ denote the preferences of voters in group J over government policy. That is, $W^J(\mathbf{q})$ is the indirect utility obtained by substitution of (8.2) into (8.1). Then voter i in group J votes for party A if

$$W^J(\mathbf{q}_A) > W^J(\mathbf{q}_B) + (\delta + \sigma^{iJ}), \tag{8.4}$$

where the term $(\delta + \sigma^{iJ}) \lesseqgtr 0$ reflects voter i's ideological preference for party B. This term includes two components; δ is common to all voters, and σ^{iJ} is idiosyncratic.

As before, the random variable δ reflects the general popularity of party B and is uniformly distributed on $[-\frac{1}{2\psi}, \frac{1}{2\psi}]$. Thus the density of this distribution is given by ψ and the expected value of δ is zero. Since δ is realized between the announcement of the party platforms and the election, the parties announce their platforms under uncertainty about the election outcome.

The distribution of individual ideology σ^{iJ} differs across groups J, and it is uniform on

$$\left[-\frac{1}{2\phi^J} + \overline{\sigma}^J, \frac{1}{2\phi^J} + \overline{\sigma}^J\right], \qquad J = 1, 2, 3.$$

Thus two parameters, $\overline{\sigma}^J$ and ϕ^J, fully characterize this distribution, and groups differ over both. In other words, groups differ in their average ideology, captured by the group-specific means $\overline{\sigma}^J$. But they also differ in their ideological homogeneity, a higher density ϕ^J being associated with a narrower distribution of σ^{iJ}. We make specific assumptions about differences in these distributions. Suppose we label the three groups according to their average ideology $\overline{\sigma}^J$: $\bar{\sigma}^1 < \bar{\sigma}^2 < \bar{\sigma}^3$. Then we assume that group 2 also has the highest density: $\phi^2 > \phi^1, \phi^3$. This is the substantial assumption. For convenience, we also assume that $\overline{\sigma}^2 = 0$ and that $\overline{\sigma}^1\phi^1 + \overline{\sigma}^3\phi^3 = 0$.[2]

Figure 8.1, in which we have drawn the distributions for σ^{iJ} in the three groups, illustrates these assumptions. Each of the three groups has an ideologically neutral voter with $\sigma^{iJ} = 0$,

2. We assume that the parties know these group-specific distributions when they announce their policy and that the electoral uncertainty derives entirely from uncertainty about the common component, δ. Alternatively, we could have generated electoral uncertainty by assuming the group means $\overline{\sigma}^i$ to be random.

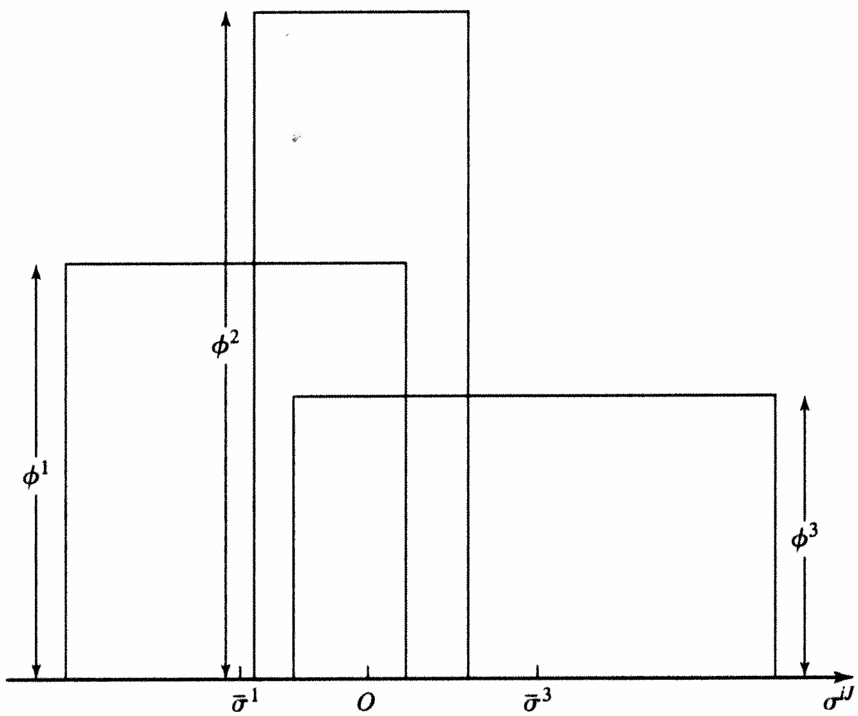

Figure 8.1

and the farther to the right we go in the figure, the more likely we are to find a voter voting for party B. As the figure illustrates, our assumptions imply that the group which on average is ideologically neutral also has the largest number of ideologically neutral voters.[3] (We sometimes find it more convenient to talk about the number of voters instead of the density, even though we are formally assuming that the distributions are continuous.) It is natural to think of this group as consisting of middle-class voters.

Ideological here has nothing to do w/ preferences over policy.

Recalling our discussion in chapter 4, we can use this figure to illustrate how the parties evaluate the announcement of different policies. Suppose party A contemplates a deviation from a common policy announcement, $\mathbf{q}_A = \mathbf{q}_B$. Such a deviation alters the number of votes party A can expect by changing the identity of the swing voters. For example, a lower tax rate τ or more public goods g benefit voters in all groups symmetrically. Taken separately, such measures thus push the identity of the swing voter in all groups to the right by the same distance, say to σ', and party A can expect to capture the voters between 0 and σ' in all groups (as the expected value of δ is equal to zero). Similarly, more transfers to group 1, financed by fewer transfers to group 3, shift the swing voter in group 1 to the right and the

3. With more general distributions, this association between the group's average ideological position and the number of ideologically neutral voters would be a natural property of all unimodal distributions.

swing voter in group 3 to the left by the same distance (recall that we assume the groups to be of the same size). This redistribution implies a net gain in votes, as there are more swing voters in group 1 than in group 3, that is, $\phi^1 > \phi^3$. Finally, higher rents r mean losing votes in all three groups and a lower probability of winning. As the announced policies must respect the budget constraint, the two parties effectively trade off votes for votes, or rents for votes, when designing their platforms.

As a final preliminary, we define $\pi_{A,J}$, that is, the vote share of party A in group J. Given our assumptions about the group-specific distributions, $\pi_{A,J}$ can be expressed as

$$\pi_{A,J} = \phi^J[W^J(\mathbf{q}_A) - W^J(\mathbf{q}_B) - \delta - \overline{\sigma}^J] + \frac{1}{2}, \tag{8.5}$$

where the expression within square brackets is a formal definition of the swing voter in group J. Clearly, party B's vote share in group J is given by $1 - \pi_{A,J}$. From both candidates' point of view, $\pi_{A,J}$ is a random variable, since it is a transformation of the random variable δ capturing party B's average popularity.

8.3 Single-District (Proportional) Elections

Consider first the equilibrium policy under an electoral rule in which it is equally important to win votes in all groups. Specifically, we study a very stylized case in which (as in the Netherlands or Israel) there is only one voting district, comprising all voters in the population. We assume that there is perfect proportional representation, in the sense that the parties obtain a seat share in perfect proportion to their vote share in the entire population. We then add the specific winning rule that the party obtaining more than 50% of the seats earns the right to set policy according to its electoral platform. Under this electoral rule, p_A is clearly given by

$$p_A = \text{Prob}\left[\frac{1}{3}\sum_J \pi_{A,J} \geq \frac{1}{2}\right], \tag{8.6}$$

where the probability refers to the random variable δ. By (8.5) and our previous assumption that δ has a uniform distribution, we have

$$p_A = \frac{\psi}{3\phi}\sum_J [\phi^J(W^J(\mathbf{q}_A) - W^J(\mathbf{q}_B))] + \frac{1}{2}, \tag{8.7}$$

where $\phi \equiv \sum_J \phi^J/3$ is the average density across groups. By symmetry, party B's probability of winning is $(1 - p_A)$.

Given our distributional assumptions and the concavity of $H(g)$, a unique equilibrium exists, in which both A and B choose the same policy. Formally, they face the same maximization problems, since $p_B = (1 - p_A)$ and since $\mathbf{q}_A$ and $\mathbf{q}_B$ enter (8.7) symmetrically but with opposite signs. To characterize the equilibrium policy, we maximize party A's objective function (8.3) with regard to $\mathbf{q}_A$, taking $\mathbf{q}_B$ as given. Exploiting (8.1), (8.2), and (8.7), and evaluating the resulting first-order conditions at the point $\mathbf{q}_A = \mathbf{q}_B$, we obtain the conditions that must hold at an equilibrium.

The equilibrium involves positive redistribution to group 2 only; that is, $f^2 > 0$, and $f^1 = f^3 = 0$ (we restrict ourselves to parameters such that some component of f is positive). This stark result follows from there being more swing voters in group 2, by our assumption that $\phi^2 > \phi^1, \phi^3$, and from the marginal utility of private consumption being constant. Thus both parties target their redistribution programs toward the middle class, since this group contains the most-responsive voters.

The equilibrium supply of public goods follows from the optimal trade-off between g and f^2. The corresponding condition is

$$\phi^2 \cdot 1 = \sum_J \phi^J \cdot H_g(g), \tag{8.8}$$

where 1 refers to the marginal utility of private consumption and superscripts refer to groups. Intuitively, cutting the supply of the public good to all voters by one unit can yield an additional unit of redistribution for the middle-class group. This means an expected gain of votes proportional to $\phi^2 \cdot 1$ in group 2 (captured by the left-hand side) but an expected loss of votes in every group J proportional to $\phi^J \cdot H_g(g)$ (the right-hand side). It is optimal for the two parties to equate the marginal gain of votes to the marginal loss of votes. Recalling our definition of $\phi \equiv \sum_J \phi^J/3$, (8.8) can be rewritten as

$$H_g(g) = \frac{\phi^2}{3\phi} > 1/3.$$

Thus compared to the socially optimal policy, the public good is underprovided in political equilibrium: $g < g^*$. Politically, targeting transfers to the mobile middle-class voters pays more than trying to please all voters with public goods.

A similar trade-off between τ and f^2 pins down the optimal tax rate. Raising the tax rate by one third for all voters can also yield an additional unit of redistribution to the middle-class group. This leads to the complementary slackness condition:

$$\phi^2 \cdot 1 \geqslant \sum_J \phi^J \cdot \frac{1}{3} = \phi \qquad [\tau \leqslant 1].$$

Here the gain of votes in group 2 always exceeds the loss, as $\phi^2 = \max_J[\phi^J]$. Since taxes are not distortionary, the optimum is a corner solution with $\tau = 1$.

Clearly, the more responsive is the middle-class group (the higher is ϕ^2), the higher is the opportunity cost of public goods. Thus the two parties find it optimal to announce a lower supply of public goods and to increase transfers to this powerful group. We therefore have a comparative statics result that will prove useful when comparing electoral rules: the larger is the density of ideologically neutral middle-class voters (the higher is ϕ^2), the smaller is equilibrium public-good provision, and the larger are the transfers to middle-class voters. With distortionary taxes and an interior solution for τ, the equilibrium tax rate also increases in ϕ.

Finally, the optimal rents implied by the two candidates' platforms are obtained as in chapter 4. Consider the trade-off between r and f^2. The complementary slackness condition corresponding to this margin is

$$p\gamma \leqslant -[R+\gamma r]\cdot\frac{\partial p}{\partial r} = [R+\gamma r]\cdot\frac{\phi^2\psi}{3\phi} \qquad [r \geqslant 0]. \tag{8.9}$$

The left-most expression reflects the marginal benefits of additional rents, whereas the remaining expressions reflect the inframarginal rents times the greater probability of losing the election. As is evident from the condition, equilibrium rents r might well be positive. Because p is equal to $\frac{1}{2}$ in equilibrium, this is more likely when R (the exogenous rents) are low. As in chapter 4, electoral competition does not eliminate rent seeking. The parties perceive electoral uncertainty because they are not perfect substitutes in the eyes of any voters except for swing voters. This implies that $\frac{\partial p}{\partial r}$ is negative but finite. The more imperfect substitutes are the two parties, the larger are equilibrium rents. Hence we have a second comparative statics result: a larger number of ideologically neutral (swing) voters (a higher ϕ^2) reduces equilibrium rents in an interior optimum. Finally, rents are larger the higher is the variance in electoral outcomes (the lower is ψ). Higher variance implies that the expected vote share is not very sensitive to policy; given this, candidates find it optimal to take a greater risk by insisting on larger rents.

This is why rents are high when there are 'ethnic' party systems.

8.4 Multiple-District (Majoritarian) Elections

What if elections are instead conducted under plurality (first-past-the-post) rule in multiple one-seat electoral districts? Specifically, assume that there are three electoral districts, each with one seat. We then add the following winning rule: earning the right to set policy requires winning at least two seats out of three. This setting can be interpreted as a parliamentary election in which two competing parties field candidates in all three districts running on the same platform. The party winning in a majority of the districts has a majority in the assembly and can thus implement its preannounced policy. Alternatively, the setting can be interpreted as a presidential election (as in the United States), where (because of the structure of the electoral college) a candidate, again, needs only a majority of the votes in a majority of the districts, rather than in a majority of the population, to be elected president.

We start with a simplifying assumption: the three electoral districts coincide with the three groups in the population. Later we show that all comparative politics results generalize if groups and districts do not completely overlap. In the present setting, existence of equilibrium is not guaranteed without further assumptions. Essentially, we must assume that the ideological bias toward party A in group 1 and toward party B in group 3 are large enough; that is, the group-specific means $\overline{\sigma}^1$ and $\overline{\sigma}^3$ are sufficiently distant from zero. The conditions for the existence of equilibrium are included as (part of) problem 2 of this chapter. When these conditions are fulfilled, we have an equilibrium in which A and B announce equal policies and the entire competition takes place in the "marginal district" made up of the middle-class (group 2) voters. Party A wins district 1 with large enough a probability and loses district 3 with large enough a probability so that neither party finds it optimal to seek voters outside the marginal district; recall that only two districts are required for winning the election.

In the present model, the electoral uncertainty derives from an aggregate popularity shock. In its equilibrium, district 2 is the pivotal voting district with probability 1, whereas districts 1 and 3 both are pivotal with probability 0. In a more general setting in which we also added (independent) shocks to the group-specific means $\overline{\sigma}^J$, we would instead have a probability of district 2 being pivotal lower than unity but higher than the positive probabilities of districts 1 and 3 being pivotal. This would lead to the same kind of qualitative comparative politics results and ensure existence of an equilibrium without too stringent assumptions.[4]

Under our assumptions, the relevant expression for party A's probability of winning the election is just the probability that party A wins district 2. By the same argument as in the previous section, this can be written as

$$p_A = \text{Prob}\left[\pi_{A,2} \geq \frac{1}{2}\right] = \psi \cdot [W^2(\mathbf{q}_A) - W^2(\mathbf{q}_B)] + \frac{1}{2}. \tag{8.10}$$

Compared to (8.6), the expression in (8.10) thus depends only on what takes place in the marginal district. We may now follow the same steps as in the previous section to characterize the policies in a convergent electoral equilibrium. Obviously, only the middle class—the sole group in the marginal district—gets the entire transfer budget. Furthermore, it is optimal for both candidates to propose more redistribution than under proportional elections. Intuitively, such redistribution has the same benefit to the parties as under proportional elections, namely

4. Strömberg (1999) shows formally how to derive equilibria for this more general case in a probabilistic voting model similar to the one of this chapter but applied to purely redistributive policy within the U.S. electoral college.

the marginal votes gained from the middle-class voters, but the costs are smaller, as the parties do not now internalize the votes lost in the nonmarginal districts. As a result, it is still optimal to set the tax rate at its maximum: $\tau = 1$. With distortionary taxes, however, the lower costs of taxation would have led to a higher tax rate.

The sharper incentives to redistribute also shape the optimal supply of public goods, because the optimal trade-off between f^2 and g now fulfills

$$\phi^2 = \phi^2 \cdot H_g(g). \tag{8.11}$$

By (8.11), $H_g(g) = 1$, whereas by (8.8), $H_g(g) < 1$ under proportional elections. Clearly, the supply of public goods is smaller under majoritarian elections.

Finally, equilibrium rents are also smaller. To see this, note that the complementary slackness condition for f^2 versus r now becomes

$$p\gamma \leqslant -[R + \gamma r] \cdot \frac{\partial p}{\partial r} = [R + \gamma r] \cdot \psi \qquad [r \geqslant 0]. \tag{8.12}$$

The condition is identical to (8.9), except that ψ replaces $\frac{\phi^2 \psi}{3\phi}$ in the expression for $-\frac{\partial p}{\partial r}$. Since $\phi^2 < 3\phi = \sum \phi^J$, higher rents make the candidates lose votes at a higher rate in majoritarian elections. Intuitively, the electoral competition is stiffer, because it is now focused on the district with the most responsive voters. Since the election outcome is more sensitive to policy, the two parties become more disciplined and forego some prospective (endogenous) rents.

Qualitatively, nothing happens to these comparative politics results if we relax the extreme assumption about a perfect overlap between groups and districts, provided that the middle-class group 2 is a *dominant group* in one of the districts.[5] Let the population share of group J in district k be denoted by $n^{J,k}$. Then group 2 is a dominant group in one of the districts if $n^{2,k} > \frac{1}{3}$ and $n^{1,k}, n^{3,k} < \frac{1}{3}$ in some k. If the middle class dominates district 2, in this sense, electoral competition will take place only in district 2. Furthermore, district 2 is an asymmetric replica of the whole population, in which group 2 receives more weight. As illustrated in figure 8.2, this asymmetry has the same effect as a higher relative density $\frac{\phi^2}{3\phi}$ of group 2 under proportional elections, the result of which was discussed in the previous section: more redistribution toward group 2, less public goods, and less rents. Problem 4 of this chapter deals with the formal derivation of this result.

This chapter's central comparative politics results can be succinctly summarized as follows. Majoritarian elections concentrate electoral competition in some key marginal districts, resulting in increased targeted redistribution towards a narrower constituency. This

5. The conditions for the existence of equilibrium become stricter, however, as we relax the assumption of perfect overlap.

// this should

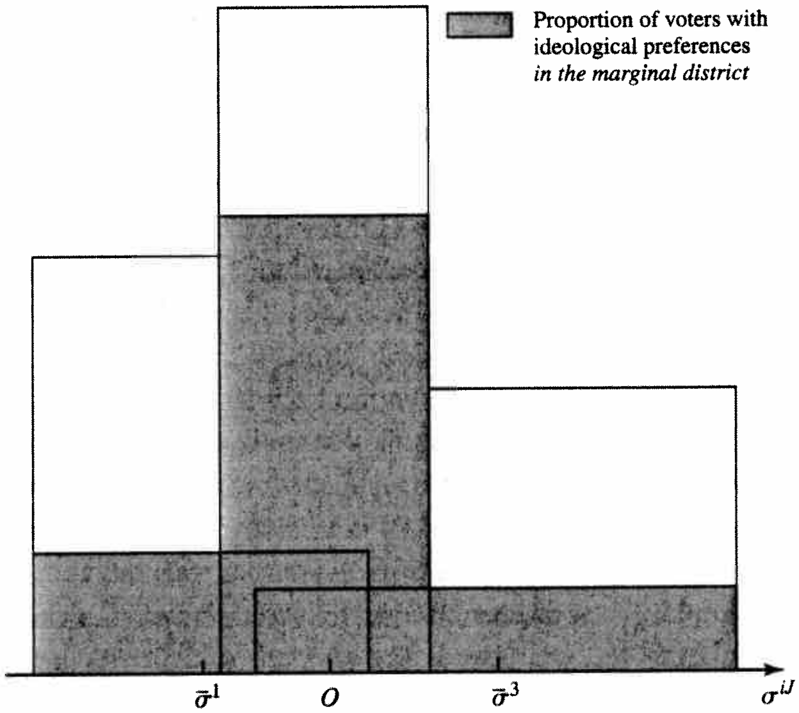

Figure 8.2

is associated with a smaller supply of public goods, as the benefits of the public good for voters in the other (nonmarginal) districts are disregarded. It is also plausible to assume, as we have done, that voters in these marginal districts are more mobile, in the sense of being more responsive to economic benefits. This makes electoral competition stiffer and reduces equilibrium rents. Extending the model with distortionary taxes, we also get the prediction that majoritarian elections should be associated with larger governments. Problem 3 of this chapter formulates the model including distortionary taxes.

8.5 Broad versus Targeted Redistribution

The theoretical concepts of spending in the previous model, public goods versus targeted redistribution, do not perfectly match with the available empirical measures of spending. For instance, transportation, health, and other classes of public spending on goods and services combine general public goods with local public goods and redistribution in kind; they may also entail some inefficiency (corresponding to r in the model). Furthermore, as we stressed in chapter 6, redistributive transfers are often designed to favor broad groups of citizens, such as the elderly or the unemployed, which are generally spread out over the electoral districts. The general results suggested by the model are that majoritarian elections induce politicians to target resources to specific groups in specific districts, reducing nontargeted spending that benefits the population at large. In many cases, narrow interests can be targeted by increasing local public-good provision and reducing general transfers. That is, the theory generates predictions for nontargeted versus targeted spending, rather than for public goods versus redistribution as such. We now illustrate this point by contrasting majoritarian and proportional elections in a different economic model that has both local public consumption, as in chapter 7, and broad redistribution, as in chapter 6. Its main result is that majoritarian elections induce more spending on local public goods and less spending on redistributive transfers than proportional elections.

Again, there are three regions, $J = 1, 2, 3$. The economic model is different, however. Like the model in section 6.4, it has employed and unemployed individuals. Employed individuals consume $c = y(1-\tau)$, where y is income and τ a nondistorting tax. Unemployed individuals receive an unemployment subsidy, f. Individuals are risk averse and evaluate private consumption with a concave utility function $U(\cdot)$. Let us also assume that individuals differ in the probability of being employed, as in example 4 of chapter 2. Let n^k denote the probability that an individual of type k is employed. The average value of n^k in the population is n, which also denotes the fraction of employed individuals. There are K different types, $k = 1, 2, \ldots, K$, and each type forms a continuum. Individuals also draw utility from local public consumption, and g^J denotes local public consumption per capita in region J. At present, suppose that the utility from local public consumption is linear (below, we discuss the implications of relaxing this assumption). Thus an individual of type k residing in region J has preferences:

$$w^{kJ} = n^k U(c) + (1 - n^k)U(f) + g^J.$$

For simplicity, we abstract from rents r. Summing over risk types k, the government budget constraint can be written as

$$ny\tau = (1 - n)f + \frac{1}{3}\sum_J g^J.$$

As before, individuals trade off economic benefits and ideology when deciding how to vote. Let $W^{kJ}(q)$ be the indirect utility function of a voter of type k in region J as a function of the policy vector $\mathbf{q} = [\tau, f, \{g^I\}]$. Then the swing voter of type k in region J is defined as usual, namely as a voter with an ideological bias given by

$$\sigma^{kJ} = W^{kJ}(\mathbf{q}_A) - W^{kJ}(\mathbf{q}_B) - \delta.$$

Finally, suppose that the distribution of individual ideological preferences is uniform and specific to each region cum risk type, with density ϕ^{kJ} for type k in region J. As before, different regions also differ in the mean of the distribution, and these means are sufficiently different that the conditions for existence of a political equilibrium under majoritarian elections are satisfied (see problem 2 of this chapter). Under majoritarian elections, regions and voting districts coincide. Region 2 is the middle district, with zero mean and the highest density.

In this setup, we can easily describe the difference between political equilibria under proportional and majoritarian elections. Repeating the steps of the previous sections, proportional equilibria solve the problem of maximizing $\sum_k \sum_J \alpha^{kJ} \phi^{kJ} W^{kJ}$, where α^{kJ} is the share in the population as a whole made up of region J individuals of type k. Majoritarian equilibria instead solve the problem of maximizing $\sum_k \alpha^{k2} \phi^{k2} W^{k2}$. That is, as already discussed in the previous section, majoritarian elections concentrate the electoral competition in the middle region/voting district.

To illustrate these different equilibria, make the simplifying assumption that all types have the same distribution of ideological preferences: $\phi^{kJ} = \phi^J$ for all k. That is, only individuals from different regions/districts differ systematically in their distribution of ideological preferences. Furthermore, assume that $\alpha^{kJ} = \frac{1}{K}$ for all pairs (k, J) (recall that the population of each region has mass 1 and that there are K types). Under these simplifying assumptions, it is easily shown that both electoral systems imply full unemployment insurance: $c = f$. Intuitively, the political equilibrium sets f and τ in both cases so as to maximize the utility of the average voter, which calls for full insurance.

The two electoral systems differ with regard to local public-good provision, however. Given the linearity of preferences, both electoral systems have only $g^2 > 0$, with $g^J = 0$ for $J \neq 2$. Under majoritarian elections, optimal public-good provision (the derivative of the objective function with respect to g^2) satisfies the following condition:

$$\frac{1}{3} U_f(f) \phi^2 = \phi^2. \tag{8.13}$$

The right-hand side of (8.13) captures the marginal political benefit of more public consumption for region/district 2: the marginal economic benefit of g^2 is unity, and ϕ^2 gives the marginal number of swing voters in district 2. The left-hand side captures the marginal political cost: as all districts pay for the public good, district 2 bears only one-third of the cost, and the opportunity cost of less resources is less unemployment benefits for all K consumer types in district 2, each type having $\frac{1}{K}$ members.

Repeating the same steps under proportional elections, we instead get the following optimality condition:

$$\frac{1}{3} U_f(f) \sum_I \phi^I = \phi^2. \tag{8.14}$$

The marginal political benefit of public consumption in region 2 is still given by ϕ^2, but the expression differs from (8.13) on the left-hand side. The costs of less unemployment insurance in all other regions are now internalized, such that the cost for any single district, namely $\frac{1}{3}U_f(f)$, is multiplied by the term $\sum_I \phi^I$. It follows immediately from (8.13) and (8.14) that unemployment insurance is more generous under proportional elections. Moreover, since $c = y(1-\tau) = f$, taxes must be lower under proportional elections. Because proportional elections induce lower taxes and more unemployment insurance, they must also be associated with less total public consumption. The results are thus analogous to those in the previous section: compared to proportional elections, majoritarian elections entail more targeted spending (local public consumption), less nontargeted spending (unemployment insurance), and a larger size of government (higher taxes).

8.6 Discussion

How general are these results? First, do the results in section 8.5 generalize if we assume concave preferences for local public consumption, as in chapter 7? Some of the results become ambiguous. Concave preferences for g^J introduce an additional difference between policy under the two electoral rules. Under proportional elections, all groups J receive some local public consumption. Under majoritarian elections, only group 2 receive a positive amount (as the other groups/districts have zero weight in the politicians' objective function). It thus becomes ambiguous which electoral rule implies higher overall public consumption. Even though citizens in the marginal district experience higher public consumption under majoritarian elections than under proportional elections, this may be offset by the lower spending on other groups. Problem 5 of this chapter provides a formal discussion of these points in an example with concave preferences for private and public consumption.

A potential fragility of some theoretical results to the specification of preferences arises also in the model of sections 8.1–8.4, in which we postulated linear preferences for private consumption. Concave preferences might lead to some ambiguities in the comparison of electoral systems. Therefore, the general and robust insight is that by concentrating competition in the marginal districts, majoritarian elections lead to more narrowly targeted redistribution in those districts. Whether this is accompanied by lower spending on nontargeted items and higher overall spending may depend on the model.[6]

As stated earlier in the chapter, we have not taken into account that different electoral rules may generate differences in the number of parties. A possible counterargument to our results is that allowing for a larger number of parties under proportional elections could allow parties that are more narrowly specialized to cater more effectively to the preferences

6. Problem 1 of this chapter considers proportional elections with concave preferences for private and public goods.

of narrower interest groups. This is a possibility, but it is hard to develop this idea into a specific hypothesis without the aid of a formal model. Unfortunately, multiparty competition in a multidimensional policy context goes beyond the current state of the art in the explicit modeling of political equilibria.

Myerson (1993a) studies voters' ability to control corruption in electoral competition under different electoral rules, assuming that some political parties are (exogenously) corrupt, whereas others are not. In that case, proportional elections are better for the voters than majoritarian elections, because they allow voters not to vote for corrupt parties without compromising their partisan preferences.

What about the empirical evidence? It is so far very scant but still indicates some systematic differences for different electoral rules. Motivated by the model in sections 8.1–8.4, Persson and Tabellini (1999b) consider a sample of more than fifty democracies around 1990. Controlling for other economic and social variables, they find that spending on public goods (such as education, order and safety, transportation, or health) as a percentage of GDP is indeed lower in countries with majoritarian elections (countries with plurality rule in single-member districts). The empirical results are fragile, however, as the specification of the model and the sample of countries affects them. Persson and Tabellini (1999b) also find weak evidence of larger total government expenditures in countries with majoritarian elections.

Preliminary results in the same cross section of countries, again controlling for economic and social variables, also suggest that governments in majoritarian countries spend more (as a fraction of GDP) on government consumption and less on transfers, compared to those in proportional countries. This is in line with the model of section 8.5, as many redistributive transfers are related to broad welfare and pension programs and hence cannot easily be targeted to specific districts. Government consumption, on the other hand, reflects public employment and purchases of goods and services from the private sector. This kind of spending may allow politicians more discretion in targeting the benefits to certain electoral districts.

Milesi-Ferretti, Perotti, and Rostagno (1999) study a panel data set for government spending in the OECD countries from the 1960s onward. They relate the size and composition of government spending to different measures of the degree of proportionality, obtaining results that are broadly consistent with the predictions in this chapter. Thus they find that transfer payments by general government—measured both as a share of government expenditures and as a share of GDP—are strongly positively related to the degree of proportionality. The share of total expenditure in GDP is also positively related to the same measures.

In the paper mentioned at the end of chapter 7, Baqir (1999) studies the effect of different electoral arrangements in U.S. cities and finds a negative result: the size of city government is not significantly affected if city council members are elected in different districts (ward systems) rather than in the whole city (at-large systems).

More empirical work is certainly needed on these important issues. More theoretical work on how the electoral rule shapes the incentives or electoral competition would also be welcome. It is reassuring, however, that Lizzeri and Persico (1998) arrive at the same conclusions as this chapter on the trade-off between redistribution and public goods, relying on a different model (an extension of Myerson 1993b) of preelection politics than the probabilistic voting model employed in this chapter. As we shall see in the next chapter, the electoral rule may also shape the postelection policy choices.

8.7 Notes on the Literature

Research on the questions addressed in this chapter is still in its infancy. The modeling in sections 8.1–8.4 draws on Persson and Tabellini (1999b). Lizzeri and Persico (1998) ask a very similar question in a different model of electoral competition, namely that of Myerson (1993a). They obtain a very similar result: that majoritarian electoral competition discourages public-good provision in favor of targeted redistribution. With regard to rents and wasteful spending, our model draws on Polo (1998) and J. Svensson (1997), as well as other works already mentioned in chapter 4.

Empirical work on the effects of the electoral rule on the composition and size of public spending can be found in Persson and Tabellini 1999b, which relies on recent cross-sectional data (five-year averages) for 54 countries. Milesi-Ferretti, Perotti, and Rostagno (1999) use panel data from 1960 for twenty OECD countries. Whereas the former study mostly employs dummy variables to characterize the electoral rule, the latter uses alternative continuous measures of the degree of proportionality. Baqir (1999) deals with the determinants of the size of U.S. city governments, including the rules for electing members to city councils.

Political scientists have studied electoral rules extensively. Some classics are Bingham Powell 1982, Lijphart 1984, Taagepera and Shugart 1989, and Cox 1997. Their analysis has been confined, however, to political phenomena and to the party structure. Myerson (1995) surveys some of this literature. A much emphasized result is that proportional elections tend to increase the number of parties represented in the legislature. Naturally, this can also profoundly affect policy formation. Roubini and Sachs (1989), Grilli, Masciandaro, and Tabellini (1991), Edwards and Tabellini (1994), Alesina and Perotti (1995b), and Hallerberg and von Hagen (1999) discuss this point informally with regard to budget deficits, pointing out how proportional elections are associated with coalition governments in the data, and these governments, in turn, with debt accumulation.

Cox (1987a, 1990a, 1990b) characterizes equilibria in the one-dimensional spatial model under a variety of electoral rules. Rivière (1998) contrasts electoral systems with more than two parties. Morelli (1999) takes an important step toward making the party structure endogenous under different electoral rules. Rigorous theoretical studies of multidimensional policy formation with an endogenous party structure under alternative electoral rules have not yet been attempted, as the problem is very difficult.

8.8 Problems

1. Proportional elections and concave utility

Consider a society with three distinct groups of voters, denoted $J = 1, 2, 3$. Each group has a continuum of voters with unit mass. Preferences over government policy are identical for every member of group J and given by the utility function

$$W^J = u(c^J) + H(g) = u(1 - \tau + f^J) + H(g).$$

Here, c^J is the private consumption of the average individual in group J, τ is a common tax rate, f^J is a transfer targeted to individuals in group J, and g is the supply of a (Samuelsonian) public good, evaluated by the concave and monotonically increasing function $H(g)$. Assume $u(\cdot)$ also to be concave.

The public policy vector $\mathbf{q}$ is defined by

$$\mathbf{q} = [\tau, g, r, \{f^J\}] \geq 0,$$

where all components are constrained to be nonnegative. Any feasible policy must satisfy the government balanced budget constraint

$$3\tau = \sum_J f^J + g + r.$$

The component r reflects (endogenous) rents to politicians and is a deliberate object of choice. Rent extraction is associated with some transaction costs $(1 - \gamma)$, such that only γr benefit the politician.

Consider proportional elections:

Before the elections, two parties or candidates (A and B) commit to policy platforms $\mathbf{q}_A$ and $\mathbf{q}_B$. They act simultaneously and do not cooperate. The winning party's platform is implemented. Party P maximizes the expected value of rents, namely

$$E(v_P) = p_P \cdot (R + \gamma r),$$

where R denotes the (exogenous) ego rents associated with winning the elections, and p_P denotes the (endogenous) probability that P wins the right to set policy, given $\mathbf{q}_A$ and $\mathbf{q}_B$.

Assume probabilistic voting. Thus the election outcome is uncertain when platforms are chosen, and different voters evaluate these candidates' ideological or personal attributes in different ways. Specifically, let $W^J(\mathbf{q})$ denote the preferences of voters in group J over government policy. Then voter i in group J votes for party A if

$$W^J(\mathbf{q}_A) > W^J(\mathbf{q}_B) + (\delta + \sigma^{iJ}),$$

where the term $(\delta + \sigma^{iJ}) \lesseqgtr 0$ reflects voter i's ideological preference for party B. δ is common to all voters and σ^{iJ} is idiosyncratic. δ is uniformly distributed on $[-\frac{1}{2\psi}, \frac{1}{2\psi}]$. σ^{iJ} differs across groups J and is uniform on

$$\left[-\frac{1}{2\phi^J} + \bar{\sigma}^J, \frac{1}{2\phi^J} + \bar{\sigma}^J\right], \quad J = 1, 2, 3.$$

Assume further that $\bar{\sigma}^1 < \bar{\sigma}^2 < \bar{\sigma}^3$, $\phi^2 > \phi^1$, and $\bar{\sigma}^1\phi^1 + \bar{\sigma}^3\phi^3 = 0$.

a. Compute the social planner's policy choice.

b. How do transfers in the proportional election model compare with the social optimum?

c. Write down the condition for the choice of rents, r, in equilibrium.

d. How does the provision of the public good differ from that of the model with linear utility of voters?

2. Existence of majoritarian equilibria

For the same model as in problem 1, now assume multiple-district (majoritarian) elections: Assume that the three groups of voters, $j = 1, 2, 3$, correspond to three electoral districts, each with one seat. Earning the right to set policy now requires winning at least two seats out of three. Assume further that voters' utility for consumption is linear ($u(c) = c$).

Let g^* be the solution to

$$H_g(g^*) = 1.$$

a. Show that the level of transfer provided to group 2 voters in an equilibrium in which both parties compete only for district two is given by

$$f^{2*} = 3 - g^* - \frac{1}{2\psi} + \frac{R}{\gamma}.$$

b. Show that such an equilibrium exists if and only if

$$\bar{\sigma}^3 \geq 3 - g^* - \frac{1}{2\psi} + \frac{R}{\gamma} \quad \text{and} \quad \bar{\sigma}^1 \leq -\left(3 - g^* - \frac{1}{2\psi} + \frac{R}{\gamma}\right).$$

3. Distortionary taxes and electoral competition

Examine the two models, those of proportional and majoritarian elections (problems 1 and 2), with distortionary taxes: Voters must now choose e (effort). The amount of effort chosen determines the amount of income earned according to the production function $y = e$. Effort is a costly activity with a cost function $c(e) = \frac{e^2}{2}$. The tax rate is proportional, so that a voter with income y pays τy. Assume further that utility for consumption of voters is linear ($u(c) = c$).

a. Characterize the voters' choice of effort, their income, and the resulting government revenue as functions of the tax rate τ.

b. Show that taxes in the majoritarian model are higher and public-good provision lower than in the proportional model.

4. Districts and voter groups that do not coincide

In the same model as in problem 3, assume that districts and groups do not overlap and that $n^{2,district\,2} > \frac{1}{3}$ and $n^{1,district\,2} < \frac{1}{3}$, $n^{3,district\,2} < \frac{1}{3}$, where $n^{j,district\,i}$ denotes the share of voters of group j in district i. Assume that parameters are such that an equilibrium exists in which parties compete on district 2 alone.

a. Show that the majoritarian equilibrium corresponds to a proportional equilibrium with different weights for the groups. Compute these weights.

b. Compare the majoritarian model and the proportional model in terms of the weights given to the different groups.

5. Broad versus targeted policy and a concave utility for local public goods

Consider the following modification of the model in problems 1–4. Once more, there are three regions, $J = 1, 2, 3$. The economic model is different, however. Voters are employed or unemployed. Employed individuals consume $c = y(1 - \tau)$, where y is income and τ a nondistorting tax. Unemployed individuals receive an unemployment subsidy, f. Individuals are risk-averse and evaluate private consumption with a concave utility function $U(\cdot)$. Let us also assume that individuals differ in the probability of being employed. Let n^k denote the probability that an individual of type k is employed. The average value of n^k in the population is n, which also denotes the fraction of employed individuals. There are K different types, $k = 1, 2, \ldots, K$, and each type forms a continuum. Individuals also draw utility from local public consumption, and g^J denotes local public consumption per capita in region J. Suppose that the utility from the local public consumption is linear. Thus an individual of type k residing in region J has preferences

$$w^{kJ} = n^k U(c) + (1 - n^k) U(f) + H(g^J),$$

where $H(g^j)$ is a concave utility function for the local good. Summing over risk types k, the government budget constraint can be written as

$$ny\tau = (1 - n)f + \frac{1}{3}\sum_J g^J.$$

Individuals trade off economic benefits and ideology when deciding how to vote. Let $W^{kJ}(q)$ be the indirect utility function of the type k voter in region J as a function of the policy vector $\mathbf{q} = [\tau, f, \{g^I\}]$. Then the swing voter of type k in region J is defined as usual, namely as a voter with an ideological bias given by

$$\sigma^{kJ} = W^{kJ}(\mathbf{q}_A) - W^{kJ}(\mathbf{q}_B) - \delta.$$

Finally, suppose that the distribution of individual ideological preferences is uniform and specific to each region cum risk type, with density ϕ^{kJ} for type k in region J. Different regions also differ in the mean of the distribution, and these means are sufficiently different that the conditions for the existence of a political equilibrium under majoritarian elections are satisfied (see problem 2). Under majoritarian elections, regions and voting districts coincide. Region 2 is the middle district, with mean 0 and the highest density. Compare proportional and majoritarian elections in this model. Assume $\phi^{k,j} = \phi^j$.

a. Show that group 2 enjoys more public-good provision under majoritarian elections than under proportional elections.

b. Assume that $U(c) = \ln c$ and $H(g) = \ln g$. Show that the amount of total spending on public goods and the amount of spending on unemployment insurance are equal in the two election models.