

PART III

---

**Public choice in a representative democracy**

## Federalism

Everyone knows that a great proportion of the errors committed by the State legislatures proceeds from the disposition of the members to sacrifice the comprehensive and permanent interest of the State, to the particular and separate views of the counties or districts in which they reside. And if they do not sufficiently enlarge their policy to embrace the collective welfare of their particular state, how can it be imagined that they will make the aggregate prosperity of the Union, and the dignity and respectability of its government, the objects of their affections and considerations? For the same reason that the members of the State legislatures will be unlikely to attach themselves sufficiently to national objects, the members of the federal legislature will be likely to attach themselves too much to local objects. The States will be to the latter what towns and counties are to the former. Measures will be too often decided according to their probable effect, not on the national prosperity and happiness, but on the prejudices, interests, and pursuits of the governments and people of the individual States.

The Federalist (James Madison)

In Part III we examine the properties of the different institutions of representative government that have been devised to supplement or replace direct democracy as a means of representing individual preferences. We begin with the United States's contribution to the evolution of representative government – *federalism* – because it is in some ways related to the theory of clubs reviewed in the preceding chapter.

### 10.1 The logic of federalism

#### 10.1.1 *The assignment problem*

Imagine a polity of nine persons divided into three local communities with three persons each. There are two public goods to be provided,  $G_L$  and  $G_F$ , where  $G_F$  is a public good like national defense, which when supplied to one community benefits all, and  $G_L$  is a public good with localized *spillovers*, like police protection. Let  $G_L$  and  $G_F$  be single-dimensional public goods, and the nine members of the polity have single-peaked preferences with ideal points as depicted in Figure 10.1. All

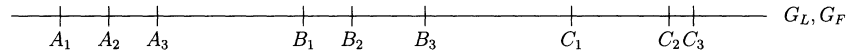


Figure 10.1. Ideal points in a federalist polity.

nine consume  $G_F$  in equal quantities. Individuals  $A_1$ ,  $A_2$ , and  $A_3$  belong to local community  $A$ , and they consume only the amount  $G_L$  supplied to their community. The same is true for the three individuals belonging to  $B$  and the three belonging to  $C$ .

Suppose now that the amounts of both  $G_L$  and  $G_F$  are decided by the larger community of nine using the simple majority rule, and that the quantity of  $G_L$  chosen by the larger polity is provided to each of the three smaller ones. With single-peaked preferences and single-dimensional issue spaces, the logic of the median voter theorem applies, and the quantities of both  $G_L$  and  $G_F$  provided correspond to the ideal point of the median voter in the polity of nine,  $B_2$ . Because  $G_F$  has the properties of a pure public good for the larger community of nine, any quantity of  $G_L$  and  $G_F$  chosen *must* be consumed by all nine citizens and, given that the simple majority is used to select this quantity, the amount  $B_2$  can be regarded as the optimum. But the public good properties of  $G_L$  allow different quantities of it to be provided to each of the three local communities. It is obvious from the location of the ideal points in Figure 10.1 and the assumption of single-peaked preferences that the members of community  $A$  can all be made better off if a smaller quantity of  $G_L$  than  $B_2$  is supplied, and the members of  $C$  will be better off with a larger quantity. Such quantities will be chosen if each local community can choose its own quantity of  $G_L$  using the simple majority rule. Thus, a superior institutional arrangement to having the quantities of both  $G_L$  and  $G_F$  decided by the larger community is to *assign* the authority to decide  $G_F$  to the larger community, and the authority to decide  $G_L$  to the three smaller ones. Having done so, one has created a *federalist* state.

A federalist state has two salient properties: (1) separate and overlapping levels of government exist and (2) different responsibilities are attached to the different levels of government. The polar case of a federalist system would have specific authorities for different activities assigned to each level of government, with each level able to determine both the expenditure levels for the activities assigned to it, and the taxes to cover these expenditures. No federalist country fits this polar case, however. In the United States, for example, primary responsibility for law enforcement lies with state and local levels of government, but Congress has passed laws governing certain criminal offenses, and federal police – like the FBI – often duplicate or assist the activities of state and local police. In many countries commonly thought of as federalist states, like the Federal Republic of Germany, regional and local levels of government have very limited authority to levy their own taxes, and thus are limited to allocating funds raised by and transferred to them by the central government. Nevertheless, all countries that are generally regarded as federalist in structure exhibit these two salient features to some degree.

A normative justification for the state is to provide public goods and resolve market failures and social dilemmas. The extension of this logic to a political community composed of states within states would determine the bounds of each governmental component on the basis of the extent of the spillovers from the public goods it was assigned, or the size of the community that was caught in a prisoners' dilemma. If the dimensions of the spillovers from two public goods were identical, both could be assigned to the same level of government. If, however, no two public goods had exactly the same geographic spillovers, the optimal federalist structure would see each public good provided by a different government, each law determined and enforced by a different government. An individual could easily be a citizen in thousands of different, overlapping governments.

Quite obviously such a situation could *not* be optimal. We have run into our old friend "transaction costs," and just as the existence of transaction costs ultimately explains the existence of the state, the existence of transaction costs explains why a federalist state is optimally composed of a few levels of government with multiple functions assigned to each level rather than thousands of levels of government with one task each. Even in a direct democracy, each citizen must incur the costs of participating in the meetings that decide what actions are warranted. He then must monitor those who carry out the tasks he has authorized. Replacing direct democracy with representative democracy lifts the burden of actually deciding budgets and taxes from the citizen to his representatives, but adds the burdens of having to participate in the process that selects the representatives, and extends the citizens monitoring duties to both the representatives he has chosen and the bureaucrats who execute the collective decisions. If citizens are mobile across communities, we must add in the costs of their having to decide which community to live in, and then of moving to it. When those who design a federalist system address the assignment problem, they must balance all of these transaction costs of having multiple levels of government against the informational efficiencies that arise from having the dimensions of a governmental unit match the benefits this government can provide to its citizens.

Why then not have a single level of government that decides all issues? The nonoptimality of this assignment in our example is contingent upon our constraining the upper level of government to choose the *same* levels of  $x$  for each community. But there is no reason to impose such a constraint. With freedom to select different  $G_L$ s, a proposal to supply  $A_2$  of  $G_L$  to community  $A$ ,  $B_2$  to  $B$ , and  $C_2$  to  $C$  defeats a proposal to provide  $B_2$  of  $G_L$  to all three communities. The three citizens of  $B$  are indifferent between these two proposals, while the six voters from  $A$  and  $C$  favor separate levels of  $G_L$ . Thus, a single assembly of all citizens assigned responsibility for determining the levels of both local and national public goods could, in principle, duplicate the outcomes arising when the public goods are assigned to local communities. The explanation for why the levels of *all* public goods are not decided in a single assembly of all citizens (or of their representatives) is again a matter of transaction costs. Once we expand the list of public goods to include all local, regional, and national goods, the task that a single assembly would face

deciding the levels of each bundle of goods for each community would become mind-boggling.<sup>1</sup>

### 10.1.2 *Federalism with geographic representation*

The discussion so far has assumed that direct democracy is employed at both levels of government. Let us now assume that representative government is employed at the higher level of government. Each local community elects one representative to an assembly of the higher level of government. (The remainder of Part III deals with representative government, and so we shall not go into the details of how it functions here, but rather consider simplified versions of it.)

Assume first that the authority to decide the level of  $G_F$  has been assigned to the representative assembly of the higher level of government, and that the levels of  $G_L$  continue to be decided using direct democracy and the simple majority rule. Under these assumptions, the quantities of  $G_L$  chosen in the three communities will continue to correspond to the ideal points of the median voter in each community, and can be regarded as optimal, given the constraint of using the simple majority rule.<sup>2</sup>

Let us assume that representatives are elected using the plurality or first-past-the-post system. The candidate receiving the most votes wins. If the only issue that the representative assembly decides is the quantity of  $G_F$ , then the candidates can be expected to compete for votes by promising to vote for certain levels of  $G_F$ , if elected to the assembly. The issue to be decided is which position along the  $G_F$  line the winning representative occupies. The median voter theorem again applies, and the three elected representatives favor the quantities of  $G_F$  corresponding to points  $A_2$ ,  $B_2$ , and  $C_2$ . If the representative assembly decides the quantity of  $G_F$  using the simple majority rule,  $B_2$  is chosen and this system of federalism and representative government selects the same outcomes as would be chosen under direct democracy at each level.

Assume now that the representative assembly is also authorized to decide the quantities of  $G_L$ . If we again assume that the same quantity of  $G_L$  must be supplied to each local community, then the outcome under this form of geographic representation is the same as under direct democracy. Representative  $A_2$  favors  $B_2$  over any point to its right, representative  $C_2$  favors  $B_2$  to any point to its left, and representative  $B_2$  favors  $B_2$  over all other points, so  $B_2$  wins. Geographic representation in this case produces the same outcome as direct democracy would, and the same is also true if we allow the quantities of  $G_L$  supplied to each community to vary.

<sup>1</sup> The classic economic studies linking the characteristics of public goods and transaction costs to the assignment of tasks are by Tullock (1969), Oates (1972), and Breton and Scott (1978). This literature is reviewed by Inman and Rubinfeld (1997).

<sup>2</sup> The chosen quantities will not in general maximize the sum of the utilities of the communities' members or the sum of their consumer surpluses. Thus, by these normative criteria, the chosen outcomes are inferior to, say, those that would be chosen using the demand-revelation process. The proposition that the outcome corresponding to the ideal point of the median voter does not maximize the sum of utilities is demonstrated in Chapter 20. The different normative criteria that one might apply to the collective decision process are the subject matter of Part V.

The situation changes, however, once we expand the dimensionality of the issue set and introduce taxes.

## 10.2 Why the size of government may be “too large” under federalism

### 10.2.1 Logrolling

Assume that we continue to have a single-dimensional issue space as depicted in Figure 10.1 with the same nine voters as before. A representative assembly has been elected consisting of three representatives who favor the positions  $A_2$ ,  $B_2$ , and  $C_2$ . The assembly is now free, however, to determine both the spending on the local public good and the taxes to pay for it. One possible outcome would be to supply the amounts  $A_2$ ,  $B_2$ , and  $C_2$  of  $G_L$  to the three communities with tax rates chosen so that each community’s tax payments just covered its own consumption of  $G_L$ . But this outcome would be inferior from the point of the  $B$ s and  $C$ s, say, to one in which a tax was levied on the  $A$ s but no  $G_L$  was supplied to them, and this revenue was used to pay for  $G_L$  in  $B$  and  $C$ . Such an outcome would lower the effective tax rates that the  $B$ s and  $C$ s would have to pay for  $G_L$  and  $G_F$ , and thus would shift the ideal point of each member of community  $B$  and  $C$  to the right in Figure 10.1. Thus, their representatives would favor higher levels of  $G_L$  in both communities. A coalition between the representatives of  $B$  and  $C$  would favor this outcome, therefore, to one in which each community chooses its own level of  $G_L$  and pays for it out of its own taxes. If a coalition between the representatives of  $B$  and  $C$  could form, it could impose this outcome, and there would be “too much”  $G_L$  provided in these communities relative to the levels that would arise if the provision of  $G_L$  were the responsibility of each local government.

This example resembles Tullock’s (1959) example of the overprovision of roads among a community of 100 farmers, where each farmer is served by one road. Tullock does not assume the existence of representative government, and the overprovision outcome might well occur in a direct democracy. The individual citizens in  $B$  and  $C$  have just as great of an incentive to discriminate against  $A$  as do their representatives. This sort of discrimination and potential inefficiency is not per se a product of their being a federalist system and geographic representation; it is due solely to the use of the simple majority rule. What federalism and geographic representation are likely to affect is the *form* that discrimination and redistribution take, *not* their existence.

To see this, consider what might occur under the polar alternative to this form of geographic representation – *at-large* representation. All voters no matter where they live must choose from the same list of candidates. Of course, nations consist of more than nine persons, and their national legislatures have hundreds of seats. Citizens do not choose among *individuals* to fill these seats, but among *parties*. In an at-large representation system, citizens choose from a list of parties, and several parties can be expected to win seats in the legislature.<sup>3</sup> Let us, therefore, think of

<sup>3</sup> The characteristics of multiparty systems are the subject of Chapter 13.

the citizens as being represented by parties, but continue for simplicity to assume that there are only nine citizens in the entire polity with preferences for  $G_L$  and  $G_F$  as in Figure 10.1. If we continue to assume that these citizens are separated geographically into three smaller communities containing the three  $A$ s and so on, then it would be reasonable to expect an at-large election to produce three parties – the  $A$ ,  $B$ , and  $C$  parties – with equal numbers of seats in the national assembly. This allocation of seats can be expected to produce *exactly* the same outcomes as under a geographic system of representation. With a geographic distribution of preferences over the set of collective decisions to be decided, as just described, there is no reason to expect great differences in the outcomes under at-large representation.

Consider now, however, a different geographic distribution of preferences. Instead of communities  $A$ ,  $B$ , and  $C$ , we have three communities 1, 2, and 3, with citizens  $A_1$ ,  $B_1$ , and  $C_1$  in community 1;  $A_2$ ,  $B_2$ , and  $C_2$  in community 2; and so on. The high, medium, and low demanders are dispersed evenly across the country. With this geographic distribution of preferences over the set of collective decisions to be decided, geographic representation will lead to the three communities being represented by individuals who hold the positions  $B_1$ ,  $B_2$ , and  $B_3$ .  $B_2$  will again be the quantity of  $G_F$  chosen. If the quantities of  $G_L$  are selected in the higher level of government's assembly, a coalition between the representatives of two of the local communities can again be expected to discriminate against the third by, say, taxing it for  $G_L$  but not providing it, and instead providing more  $G_L$  to themselves.

The situation is likely to be quite different under at-large representation, however. Here one could again expect  $A$ ,  $B$ , and  $C$  parties to win seats by promising to represent the high, medium, and low demanders for  $G_L$  and  $G_F$ . A coalition between two of the parties would now be based on the levels of their demands for the public goods, and discrimination would likely be against either the high- or the low-demand groups, depending on which coalition formed. If the differences in the demands for the public goods were based on differences in incomes, with the  $C$ s having the highest incomes, then the discrimination and redistribution that would result under the simple majority rule would be related to an individual's income and not her geographic location. Note that under at-large representation, if the levels of  $G_L$  in each local community were decided at the national level, a coalition between the  $A$  and  $B$  parties would *ceteris paribus* favor a uniform provision of  $G_L$  in all local communities that was between  $A_3$  and  $B_1$ , and thus *less* than the outcome in two of the three communities, when they alone are responsible for this decision.<sup>4</sup> This is why the words "too large" are placed in quotation marks in the title of this section. In a federalist system, the discrimination and redistribution that result when higher levels of government provide local public goods are likely to result in greater quantities of local public goods being supplied to some communities than would occur if each local community chose its own quantity, and smaller quantities in others.

This latter conclusion depends on there being the kind of exploitation of the minority by the majority under the simple majority rule that Tullock described in

<sup>4</sup> Baron (1993) presents a model in which the provision of a local public good by the central government can have equally ambiguous results.

his farmers/roads example. Some have argued that this sort of discrimination does not in fact occur. We examine their arguments next.

### 10.2.2 *Universalism*

The results of the previous section presume that a winning coalition in the legislature of the central government exploits the potential inherent in the majority rule and provides the local public good to only its members and/or provides these goods using discriminatory tax formulas. This sort of “tyranny” by the majority coalition has obvious attractions for its members but, given the high probabilities of cycling, it also has its risks. The representative or party that finds itself in a winning coalition today may be on the losing side tomorrow. To avoid such risks, several authors have claimed that legislatures adopt a norm of *universalism*. Every local community is supplied any local public that the central government supplies.<sup>5</sup>

If the legislature of the higher level of government is to choose the levels of  $G_L$  for each local community using a norm of universalism, one might expect it to supply the amounts  $A_2$ ,  $B_2$ , and  $C_2$  if the citizens’ preferences were as depicted in Figure 10.1. The empirical evidence in support of universalism suggests that the high demanders have greater influence in the legislature, however (Weingast and Marshall, 1988; Hall and Grofman, 1990). Thus, instead of the set of outputs  $A_2$ ,  $B_2$ , and  $C_2$  being provided,  $A_3$ ,  $B_3$ , and  $C_3$  are.

Often the effect of geographic representation seems not to be that a particular local public good is provided by the central government to each local community, but that different local public goods are provided. Each representative in the federal legislature proposes a “pet project” that her constituents would like to see the federal government finance. The application of the norm of universalism results in all of their wishes being fulfilled.

Schwartz (1994) presents a model to explain why this comes about. Each representative is concerned only about being reelected, and her constituents are only concerned about their pet projects being provided. The constituents ignore the costs of these projects, which are spread across the entire federal polity, and the outcome is that bundles of local public goods get provided by the central government that individually would have been turned down by their respective local communities. *Too much* of each local public good is provided.

### 10.3 Intergovernmental grants under federalism

One important feature of federalist systems is that one level of government may not actually provide a public good for another level, but merely transfers money to it. These intergovernmental grants are usually from the higher levels to the lower ones, but not always. The European Union’s budget is made up of grants from the 15 national governments of the countries in it. In this section we explore the properties

<sup>5</sup> An explanation for universalism that is less dependent on the assumption of selfish utility maximization would be that members are moved out of a sense of fairness to apply legislation universally. See Weingast (1979), Niou and Ordeshook (1985), and Weingast and Marshall (1988).



of such grants. We first look at the normative argument for having such grants, and then at the empirical evidence relating to their effects.

### 10.3.1 Intergovernmental grants to achieve Pareto optimality

Intergovernmental grants can improve the allocation of resources when a locally provided public good has positive externalities. One example of such a public good is highways. In some federalist systems, like the United States, each regional government is responsible for building and maintaining the roads in its political jurisdiction. In addition to its own citizens using these roads, however, the citizens of other jurisdictions sometimes use them. To achieve Pareto optimality the demand for maintenance and construction of these roads should be measured by summing the demand schedules of all users, whether they are citizens of the jurisdiction or not. But the political system only records the demands for roads in a given jurisdiction of its own citizens. The demand for roads by all citizens in the country is underestimated, and the resulting quantity of road services provided is less than the Pareto-optimal quantity.

In this example, the provision of roads by one community leads to a *positive externality* with respect to other communities, and the problem can be investigated like that of any externality. The problem is exactly like that analyzed in Section 2.6 except that the externality is symmetric. In the case of two communities,  $A$  and  $B$ , the amount of the public good,  $G_A$ , that  $A$  consumes equals its own provision of roads,  $R_A$ , plus a fraction  $s_A$  of the amount of roads supplied by  $B$ ,  $0 < s_A < 1$ ; and the same is true for  $B$ :

$$G_A = R_A + s_A R_B, \quad G_B = R_B + s_B R_A. \quad (10.1)$$

If all citizens in  $A$  have the same incomes,  $Y_A$ , and utility functions,  $U_A(X_A, G_A)$ , then they will unanimously agree to construct the amount of roads that maximizes the following Lagrangian:

$$L_I = U_A(x_A, G_A) + \lambda(Y_A - P_x x_A - P_r R_A), \quad (10.2)$$

where  $P_x$  and  $P_r$  are the prices of the private good  $X$  and roads, and  $G_A$  satisfies (10.1). Maximization of (10.2) leads to the familiar first-order condition

$$\frac{\partial U_A / \partial G_A}{\partial U_A / \partial X_A} = \frac{P_r}{P_x}. \quad (10.3)$$

An analogous condition could be derived for the representative citizen from  $B$  (all  $B$ s also have identical utility functions).

To obtain the Pareto-optimal quantity of roads, we maximize the utility of a representative  $A$  with respect to the four decision variables  $X_A$ ,  $X_B$ ,  $R_A$ , and  $R_B$ , subject to the constraint that the utility of a representative  $B$  is held constant, and the aggregate budget constraint.

$$L_{PO} = U_A(X_A, G_A) + \lambda(\bar{U}_B - U_B(X_B, G_B)) + \gamma(Y_A + Y_B - P_x X_A - P_x X_B - P_r R_A - P_r R_B). \quad (10.4)$$

This yields the four first-order conditions:

$$\begin{aligned}
 \frac{\partial L_{PO}}{\partial X_A} &= \frac{\partial U_A}{\partial X_A} - \gamma P_x = 0 \\
 \frac{\partial L_{PO}}{\partial X_B} &= \frac{\partial U_B}{\partial X_B} - \gamma P_x = 0 \\
 \frac{\partial L_{PO}}{\partial R_A} &= \frac{\partial U_A}{\partial G_A} \frac{\partial G_A}{\partial R_A} - \lambda \frac{\partial U_B}{\partial G_B} \frac{\partial G_B}{\partial R_A} - \gamma P_r = 0 \\
 \frac{\partial L_{PO}}{\partial R_B} &= \frac{\partial U_A}{\partial G_A} \frac{\partial G_A}{\partial R_B} - \lambda \frac{\partial U_B}{\partial G_B} \frac{\partial G_B}{\partial R_B} - \gamma P_r = 0.
 \end{aligned}
 \tag{10.5}$$

From (10.1) we have

$$\frac{\partial G_A}{\partial R_A} = 1, \quad \frac{\partial G_A}{\partial R_B} = s_B, \quad \frac{\partial G_B}{\partial R_B} = 1, \quad \frac{\partial G_B}{\partial R_A} = s_A.
 \tag{10.6}$$

Substituting from (10.6) into (10.5) and rearranging yields (10.7):<sup>6</sup>

$$\begin{aligned}
 \frac{\partial U_A / \partial G_A}{\partial U_A / \partial X_A} &= \frac{P_r}{P_x} - s_B \frac{\partial U_B / \partial G_B}{\partial U_B / \partial X_B} \\
 \frac{\partial U_B / \partial G_B}{\partial U_B / \partial X_B} &= \frac{P_r}{P_x} - s_A \frac{\partial U_A / \partial G_A}{\partial U_A / \partial X_A}.
 \end{aligned}
 \tag{10.7}$$

Equation (10.3) states the condition that is fulfilled when the representative citizen from *A* maximizes her utility ignoring the consequences of this decision for *B*. An analogous condition holds for *B*. Substituting these into (10.7) gives us

$$\begin{aligned}
 \frac{\partial U_A / \partial G_A}{\partial U_A / \partial X_A} &= (1 - s_B) \frac{P_r}{P_x} \\
 \frac{\partial U_B / \partial G_B}{\partial U_B / \partial X_B} &= (1 - s_A) \frac{P_r}{P_x}.
 \end{aligned}
 \tag{10.8}$$

To achieve the Pareto-optimal supply of roads in the two communities a Pigouvian subsidy must be offered to a community per unit of roads purchased that equals the proportionate spillovers from its roads onto the other community.

One way to obtain this outcome is for the higher level of government to levy lump sum taxes on both communities and then offer each of them subsidies in the form of *matching grants*. The effects of a matching grant on a local community's purchases are illustrated in Figure 10.2. In the absence of any grant the community faces the budget constraint *BB* and purchases  $X_0$  of the private good and  $G_0$  of the public good. A 50 percent matching grant results in the federal government's purchasing one unit of *G* for every unit of *G* purchased by the local community and is equivalent

<sup>6</sup> Compare the analogous derivation for externalities in equations (2.34) through (2.41).

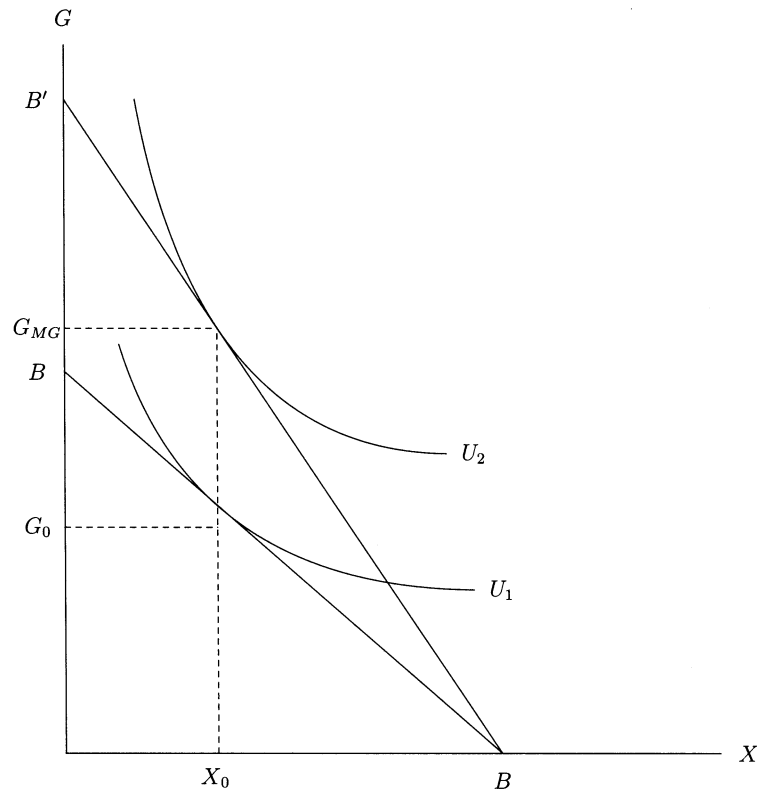


Figure 10.2. The effects of matching grants.

to a 50 percent reduction in the price of  $G$  to the local community. This shifts the community's budget constraint line out to  $BB'$ , and it now purchases  $G_{MG}$  of the public good. If  $G$  is a normal good, both the substitution and the income effects of the matching grant cause the amount of  $G$  purchased to increase. These two effects work in opposite directions with respect to its consumption of the private good  $X$ , however, and the net effect could be a reduction in the amount of  $X$  purchased. Should this occur, the increase in the amount of  $G$  brought about by the matching grant would be *greater* than the amount of money actually transferred from the central to the local government. Thus, matching grants are a potentially powerful way to influence the patterns of spending by local communities.

A second form of intergovernmental grant often employed is an *unrestricted* or *general* grant. As its name implies such grants are unconditional and allow the local government the freedom to spend the money any way it chooses, including as a *tax expenditure*, that is by cutting its local taxes to some or all of its citizens and thereby allowing them to use some of the grant to increase their private consumption. The effects of an unrestricted grant are illustrated in Figure 10.3. In the absence of the grant the community's budget constraint is  $BB$  and it purchases  $X_0$  and  $G_0$  of the two goods. The unrestricted grant allows the community to increase its consumption of the private good by  $B' - B$  if it offsets all of the grant by a tax cut, or to increase its consumption of public goods by this amount. The community's budget constraint

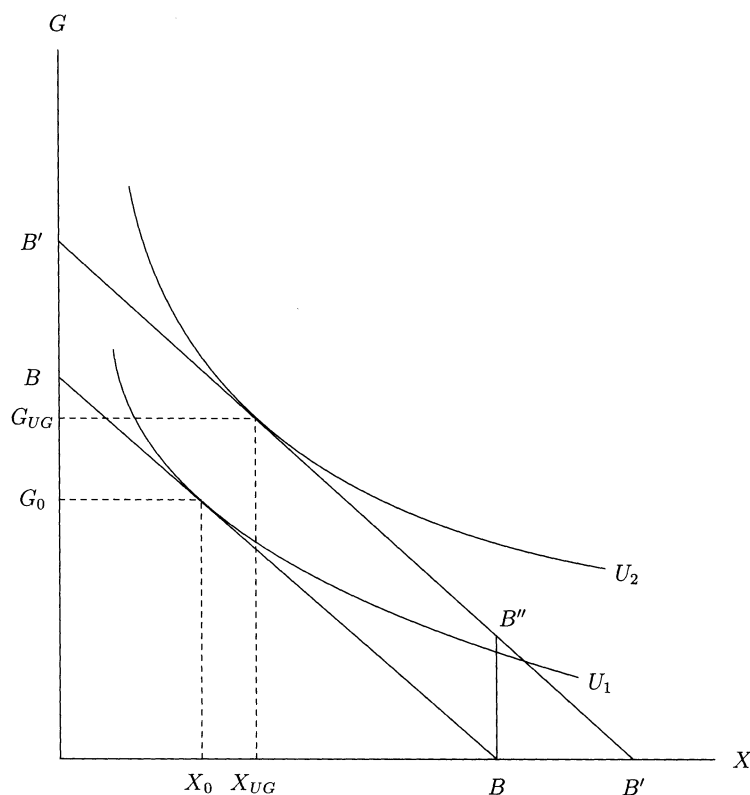


Figure 10.3. The effects of unconditional and earmarked lump-sum grants.

line shifts out to  $B'B'$ , and it now purchases  $G_{UG}$  of the public good and  $X_{UG}$  of the private good. If both  $G$  and  $X$  are normal goods, the amounts of each purchased increase. An unrestricted grant's only impact on the quantity of local public goods purchased comes through its income effect, and thus it can be expected to lead to a smaller increase in the local community's spending on public goods than a matching grant.

A third form of intergovernmental grant is an *earmarked* or *specific* grant. Earmarked grants can only be used to finance the programs for which they are earmarked, but they resemble unconditional grants in that they do not constrain the local government to spend any specific amounts of its own money on these programs. Thus, specific grants can also provide local governments with the freedom to reduce their taxes. An earmarked grant does not allow the community to reduce its consumption of the public good for which it is earmarked below the level of the grant. Thus, an earmarked grant of the same magnitude as the unrestricted grant just discussed would shift the community's budget constraint line out by  $B' - B$  from point  $B$  on the  $X$  axis (see Figure 10.3). The new budget constraint becomes the kinked line  $BB''B'$ . If the quantity of the public good that the community would have purchased in the absence of the specific grant exceeds the amount of the grant, which is the case in Figure 10.3, then an earmarked grant's only impact on the quantities of the two goods purchased also comes entirely through its income effect,

and the outcome is exactly the same as if there had been no conditions attached to the grant. If the quantity of the public good that the community would have purchased in the absence of the specific grant is less than the amount of the grant, then an earmarked grant increases the quantity of the public good purchased up to the level of the grant. Of the three grants discussed, clearly matching grants give the central government “the most bang for its buck” in affecting the direction of local government spending, and thus are most compatible with a “spillovers” rationale for intergovernmental grants.

An alternative justification for intergovernmental grants is to offset differences in fiscal capacities across communities. Consider again the example using Figure 10.1, and assume that the reason community *A* wishes to purchase less of the two public goods is that its citizens have lower incomes than in *B* and *C*. Following the argument for Pareto-optimal redistribution presented in Chapter 3, citizens in *B* and *C* may get utility out of transferring money to community *A*. A proportional or progressive federal income tax combined with a federal grant to *A* could thus be a form of Pareto-optimal redistribution.

If citizens in *A* prefer smaller quantities of public goods than those in *B* and *C* because they are poorer than the citizens of *B* and *C*, then consumption of private goods in *A* also will be lower than in the other two communities. If now citizens in *B* and *C* merely wish to raise the welfare of citizens in *A* through intergovernmental grants, then they will do so by voting for an *unrestricted* grant. Such a grant allows citizens in *A* to use the funds any way they choose, and thus to allocate these funds between public and private goods to maximize their utilities. It is the form of intergovernmental grant which is most compatible with citizen/consumer sovereignty. Thus, the logic underlying the optimal form of intergovernmental grant is completely reversed when the goal is to achieve Pareto optimality – where matching grants are to be preferred to correct inefficiencies arising from intergovernmental spillovers, unconditional grants are optimal to eliminate the “interpersonal externalities” that arise when the residents of wealthy communities contemplate the situation of people in poorer ones.

Sometimes it is argued that intergovernmental grants to poorer communities are needed, not simply to allow the citizens in these communities to increase their welfare as best they see fit by expanding their consumption of both private and public goods, but rather to allow (induce) them to increase their purchases of specific *governmentally provided* goods. Citizens in wealthy communities only get utility out of the additional consumption of certain public goods by a poor community. One example of such a good might be education. If *A* must provide education to its citizens out of its own tax revenues, the median voter favors  $A_2$  of education. But citizens in *B* and *C* believe that no child should receive less of an education than is implied by, say,  $B_2$  of education. If such were the case, matching or earmarked federal grants to local communities for education might be needed to achieve Pareto optimality.

There are other factors that affect the nature and size of intergovernmental grants, but many of these are hypotheses about why these grants *actually* exist, rather than hypotheses about why they *ought* to exist. We take them up in the next subsection, therefore.

### 10.3.2 *The empirical evidence on intergovernmental grants*

The analysis of intergovernmental grants leads to some very clear predictions as to their effects on local governments' spending. If a local government's budget is, say, 5 percent of the income of its residents, and the income elasticity of demand for local public goods is one, then 5 percent of any block grant to a local government should wind up as an increase in local government spending, and the rest be allocated to the private consumption and savings of its residents, since for such an unconditional grant, only its income effect is at work. This simple prediction has, however, been consistently disconfirmed in the empirical literature. Local government spending has been found to increase by anywhere from 25 percent of the size of the grant to over 200 percent, with the average estimate exceeding 50 percent.<sup>7</sup> Money from the central government transferred to a local government largely "sticks where it lands" – in the local government's budget. So consistent is this result that it has acquired its own name: the *flypaper effect*.

Such a consistent and dramatic refutation of the predictions of the simple governmental grants model has led to a huge literature, which has either reworked the model to try and get it to fit the data, or reworked the data to try and get them to fit the model.<sup>8</sup> The literature is too vast for us to wade very deeply into it. We content ourselves here, therefore, with an examination of two explanations for the flypaper effect that rely on public choice reasoning, and a brief look at the econometric criticisms.

One explanation for the flypaper effect is that it is due to *fiscal illusion* (Courant, Gramlich, and Rubinfeld, 1979; Oates, 1979). Tanzi (1980) has traced the concept of fiscal illusion back to John Stuart Mill and also cites Pareto as a source. But it is to the Italian economist Puviani (1897, 1903) that credit must go for emphasizing the importance of fiscal illusion to a positive theory of government (see also the discussion in Buchanan, 1967, pp. 126–43). The general idea of fiscal illusion is that there are certain revenue sources of the government that are unobserved or not fully observed by the citizens. If money from these sources is spent, some or all citizens benefit from these expenditures, and support for the government increases. Because the citizens are unaware of the source of these expenditures, they do not perceive the pain of having either paid higher taxes or foregone a tax cut to bring about this increase in expenditures. Thus, spending revenue from sources that are hidden from the citizens' view by fiscal illusion should increase the popularity of the government and thus those in government who seek reelection have an obvious incentive to spend any revenue that is subject to fiscal illusion, and seek revenue that has this characteristic. With respect to intergovernmental grants the fiscal illusion argument is that voters do not perceive that these grants are implicitly to *them* and not to those in their local government, and thus that all of the money could be

<sup>7</sup> The 25 percent figure comes from Gramlich and Galper. Kurnow (1963), the earliest study that Gramlich (1977) lists in his survey of the empirical literature, estimated increases in local expenditures that were 245 percent of the size of the grant.

<sup>8</sup> For surveys of this literature, see Gramlich (1977), Inman (1979), Fisher (1982), Heyndels and Smolders (1994, 1995), Hines and Thaler (1995), Becker (1996), and Bailey and Connolly (1998).

given to the voters if those in government chose to do so. The citizens' ignorance of the economics of intergovernmental grants leads to fiscal illusion regarding these grants. This fiscal illusion *allows* those in the local government to spend a higher fraction of the grant money than would maximize the voters' utilities. The local government exploits this opportunity and the "paradox" of the flypaper effect is explained.

The preceding explanation of the flypaper effect relies on the motivation of elected officers in local government to increase their popularity. A second explanation of the flypaper effect emphasizes the motivation of the unelected officers in local government who reside in its bureaucracies. Niskanen (1971) has hypothesized that government bureaucrats seek to maximize the size of their budgets. His theory also relies on information asymmetries, but now it is the elected members of government who lack information and are exploited by the appointed bureaucrats who have it.<sup>9</sup> An explanation for the flypaper effect based on this theory could run as follows: the central government provides a grant earmarked for education to a local government. The grant is less than the local government's current education budget, and thus is equivalent to an unconditional grant and should result in only a modest increase in local education expenditures. Members of the local education bureaucracy, however, are keen on spending the money and take advantage of the elected politicians' ignorance of the costs and benefits of education to convince them that this money "really is needed to improve the quality of local schooling." A large fraction of the grant winds up as an *addition* to the local education budget.<sup>10</sup>

The prediction of a modest budgetary impact of intergovernmental grants applies only to unconditional and (most) earmarked grants. The budgetary impact of matching grants could be large. It is not always easy to determine the nature of the grants made, however, and thus some matching grants have been included in the empirical studies that find a flypaper effect. This is one possible empirical explanation for it.

A related explanation is that an earmarked grant may, *implicitly*, be a matching grant (Chernick, 1979). When the central government decides to provide a local community with a grant earmarked for education, it presumably does so because it wants the local government to spend more on education. If the local government chooses to use most of the grant to cut taxes and not increase education outlays, the central government's objective has not been met. This outcome could significantly reduce the probability of a similar grant from the central government in the future. If members of the local government realize this – and if they do not, members of the central government are likely to make them aware of this danger – they will treat the earmarked grant as a matching grant and expand their education budget by more than the amount warranted from the income effect of the grant alone. It is better for the elected local politicians to obtain money from the central government, even if they must spend it on education, than not to obtain it at all.

<sup>9</sup> Niskanen's theory is discussed in Chapter 16 along with the related theory of Brennan and Buchanan (1980).

<sup>10</sup> Wilde's (1968, 1971) explanation for the flypaper effect anticipates Niskanen's model to some extent. Schneider and Ji (1987) provide empirical support for a bureaucratic power explanation by showing that the extent of competition between governments, which presumably reduces a bureaucracy's monopsony power, reduces the magnitude of the flypaper effect.

Several criticisms have been made of the econometrics employed in flypaper-effect estimations. But improving the econometrics alone does not seem capable of eliminating it.<sup>11</sup> Perhaps the simplest, and yet potentially most devastating attack on the empirical evidence for a flypaper effect is by Becker (1996). She is able to eliminate the flypaper effect simply by substituting a logarithmic functional form for the linear one commonly used. In much econometric work, such substitutions have only modest effects on the conclusions drawn. That it should have such a dramatic effect in this literature is noteworthy. Pending confirmation of Becker’s findings with other data sets, however, one still must conclude that a significant fraction of federal grant money seems to stick where it lands at the local governmental level.<sup>12</sup>

#### 10.4 Why the size of government may be “too large” and “too small” under federalism

Much of the public choice literature, as with that on the flypaper effect, argues in one way or another that government grows to be *too large*, too large in the sense of being larger than the size that would maximize the median voter’s utility, or would maximize some welfare function defined over the utilities of members of the community. There are some reasons to believe that at least some parts of the government sector may be too small in a democracy, however. This danger is particularly likely in a federalist state with geographic representation.<sup>13</sup>

To illustrate how government expenditures can be too large and too small at the same time we assume a two-level federalist state. Instead of assuming that the preferences of the median or representative voter are decisive, as in the two earlier models presented in this chapter, we assume that the preferences of those in the government are decisive. The main goal of elected government officials is assumed to be reelection. The more the government spends, holding taxes constant, the happier voters are and the higher the probability of incumbent politicians being reelected. Assuming that this probability increases at a diminishing rate, we can depict the marginal valuation of expenditures by elected officers in the local and federal governments as  $MV_L$  and  $MV_F$  in Figures 10.4a and 10.4b. (If we were to assume that it is the preferences of those in the bureaucracy that are decisive, and that they are budget maximizers, then these curves represent the marginal utilities of bureaucrats in the local and federal governments.)

While spending money wins votes, *ceteris paribus*, raising taxes loses them.  $MC_L$  and  $MC_F$  in Figures 10.4a and b depict the marginal costs in reduced popularity from raising the revenue to pay for the expenditures at the two levels of government. If the constitution assigns local public goods to the local government, and national public goods to the federal level, then the local and federal governments choose to supply the quantities  $G_L^0$  and  $G_F^0$ , where the marginal gain in the probability of

<sup>11</sup> See Wyckoff’s (1991) review and tests.

<sup>12</sup> Worthington and Dellery (1999) have confirmed Becker’s finding using grants data in Australia.

<sup>13</sup> Downs (1961) was one of the first to argue from a public choice perspective that government may be too small in a democracy.



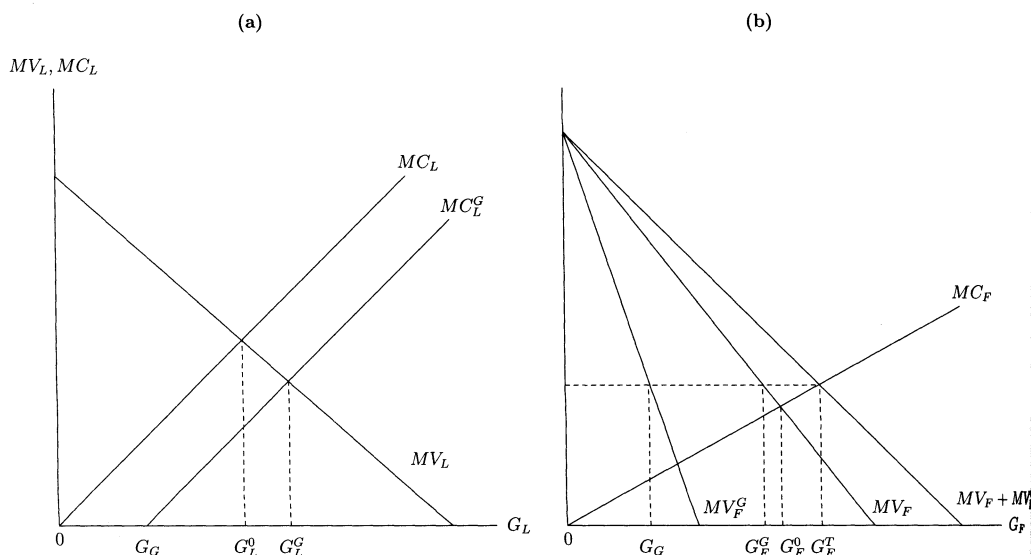


Figure 10.4. Effects of grants on government expenditures in a federalist system.

being reelected from an increase in expenditures just equals the reduction in the probability of being reelected from an increase in taxes.

Now assume that representatives to the national legislature are elected geographically, and that this legislature is free to provide local public goods directly, or to provide grants to the local government. These representatives can then increase their chances of being elected by both spending more on the national public goods and by spending more on the local ones. Let  $MV_F^G$  in Figure 10.4b represent the marginal valuation of those in the national legislature of increases in local expenditures for this reason. The marginal valuation of expenditures at the federal level on both national and local public goods is then  $MV_F + MV_F^G$ . The new level of total federal expenditures including grants to the local government or direct purchases of local public goods is  $G_F^T$ , which is made up of  $G_F^G$  in expenditures by the central government on national public goods, and  $G_G$  in grants or direct local expenditures ( $G_G = G_F^T - G_F^G$ ). The central government's total outlays increase; its expenditures on national public goods decline from  $G_F^0$  to  $G_F^G$ . Financing local public goods out of federal revenues has crowded some national public goods out of the federal budget. If  $G_F^0$  were the optimal level of expenditures on national public goods, then the shift of funding of some of the local government's budget to the national level would have resulted in *too small* of a federal budget on national public goods.<sup>14</sup>

Turning to the local level of government in Figure 10.4b, we see that a grant of  $G_G$  shifts the local government's marginal cost schedule over to  $MC_L^G$ . The new level of local expenditures including the grant is  $G_L^G$ . Even though we have assumed

<sup>14</sup> A model in which political competition leads vote-maximizing politicians to choose Pareto-optimal quantities of public goods is discussed in Chapter 12. This model would predict a decline in welfare from allowing the central government to finance local public goods, and the resulting outcome as depicted in Figure 10.4b.

that the grant was a block grant rather than a matching grant, by allowing the local government to cut taxes somewhat, the grant lowers its marginal costs of purchasing local public goods and results in an increase in expenditures on local public goods of  $G_L^G - G_L^0$ , an amount that exceeds that which we would expect from the income effect alone.<sup>15</sup>

A comparison of Figures 10.4a and 10.4b reveals that the *net effect* of the intergovernmental transfers on the total size of the government sector is *positive*. The decline in spending on the national public goods,  $G_F^0 - G_F^G$ , is less than the increase in expenditures on local public goods,  $G_L^G - G_L^0$ . A federalist form of government with geographic representation and intergovernmental grants can result in less than the optimal expenditures at the national level and more than the optimal level at the local level.

- Grossman (1989a) has tested the prediction that intergovernmental grants lead to a larger total government sector using both cross-sectional and time series data for the United States.<sup>16</sup> His cross-section estimates were for 1976–7 using data from the 48 continental states. One set of estimates is given in (10.9):

$$G = .036^{**} + 6 \times E^{-5^{**}} TR + 4 \times E^{-6^*} Y + 1 \times E^{-4^{**}} FTR + 3 \times E^{-4} MFG$$

$$n = 48, \quad \bar{R}^2 = .583 \quad (10.9)$$

\*\* = significant at 5 percent level, \* = significant at 10 percent level.

In this equation,  $G$  is state and local taxes as a fraction of personal income;  $TR$  is state transfers to local communities divided by state population,  $FTR$  is federal transfers to the state divided by state population, and  $MFG$  is state population divided by the number of multiple function governments in the state (basically cities and counties). The third variable was insignificant, but the other two were significant. The positive coefficient on  $TR$  indicates an increase in the size of the total government sector in a state in proportion to the amount of funds passed from the state-level government to local ones. The positive coefficient on  $FTR$  is evidence of the flypaper effect. Grossman’s estimate implies an elasticity of state expenditures out of federal grants of 31 percent. His time series estimates using federal, state, and local government expenditure data also confirm the hypothesis that intergovernmental transfers in a federalist system lead to an expansion of the total size of the government sector.<sup>17</sup>

A somewhat different form of governmental waste occurs when two governments compete to supply the same service. Here a form of “common pool” problem arises with both governments over exploiting the pool of tax payer resources.<sup>18</sup>

<sup>15</sup> The difference in the results from this model and the simple model of grants used in the previous section arises because that model implicitly assumed that the marginal costs of purchasing more public goods were constant in the absence of matching grants, where here we assume that the marginal costs to the politicians are increasing.

<sup>16</sup> The above model is a simplified version of Grossman’s. For a much more elaborate model of intergovernmental transfers in a federalist system, see Renaud and van Winden (1991).

<sup>17</sup> Although Grossman did not test explicitly for the presence of crowding out, the fact that one of the predictions of his model was supported leads one to expect that its other implications are also likely to be present in the data.

<sup>18</sup> See models and evidence provided by Flowers (1988), Migué (1997), and Wrede (1999).

Table 10.1. *Distribution of European Union expenditures by budget category, 1985 and 1995 (percentages)*

		1985	1995
Redistribution	Agriculture and fisheries	72.9	53.6
	Regional policy	5.9	13.6
	Social policy	5.7	11.9
Allocative efficiency	Research, energy, transport	2.6	5.6
	External policies	.	6.2
	Administrative costs	4.6	5.1
	Miscellaneous	4.4	4.5

Source: Goodman, 1996, pp. 101, 105–6.

The possibility that under geographic representation local interests shift local projects to the national budget and thereby crowd out national interests has largely been discussed in the context of the United States.<sup>19</sup> However, the European Union (EU) in many ways provides a cleaner and more dramatic example of this phenomenon. The most important decision-making body of the EU is the *Council*, which meets in Brussels. Each member country is represented on the Council by delegates appointed by each country's government. Thus, in the EU's most important decision-making body, representation is geographically based, as in the United States.

The Council faces a very severe budget constraint. Its funds come from contributions from the member countries, which are already pressing up against the upperbound of the tax revenue that they can raise (see Chapter 22). The entire budget of the EU amounts to less than 3 percent of the EU's GDP. Thus, if any local – in this case *national* – public programs work their way into the EU's budget, the potential for crowding out European-wide public goods is large.

On the other hand, until 1991 the Council made decisions using the unanimity rule. Given our discussion of this rule in Chapter 4, one might have expected that its use would prevent local public goods and involuntary redistribution from entering the EU's budget, but this has not been the case. Instead, the Council seems to have practiced the same kind of universalism that many see in the U.S. Congress. Table 10.1 breaks the EU budgets in 1985 and 1995 down into various large categories. Outlays that were purely or largely redistributive made up almost 85 percent of the EU budget in 1985, and almost 80 percent in 1995. Activities that could be fairly clearly identified as having salient public good properties accounted for only 2.6 percent of the EU budget in 1985. Even if we categorize the EU's aid to non-EU countries (External Policies) as a "Pareto-optimal redistribution," and thus a form of allocative efficiency activity, outlays to improve allocative efficiency accounted for only 11.2 percent of the EU budget in 1995.

Today the biggest single item in the EU's budget is, as it has always been, subsidies to farmers. One might argue, at the national level, that these could constitute a

<sup>19</sup> See, for example, Ferejohn (1974) and Fiorina (1977a).

form of Pareto-optimal redistribution. French citizens get utility from seeing French farmers better off, and therefore are willing to pay higher prices for food and higher taxes to subsidize their farmers. It is hard to press this argument at the level of the EU, however. To do so one would have to argue that the average citizen in, say, Portugal, gets utility from seeing French farmers better off – *even though the average French farmer is richer than the average Portuguese citizen*. What explains the predominance of agricultural redistribution in the EU's budget is the same kind of pork-barrel politics that has led to large farm subsidies in America. Each European country's farmers can impose sufficient political costs on its government to induce it to lobby hard for high subsidies.

Given the scale of redistribution in the EU, and given the size of its budget, there is nothing left over to finance those European-wide public goods that ought to justify its existence – like foreign policy and defense. Assuming that there *are* some European-wide public goods, then the EU's redistribution policies, fostered by its confederalist political structure with geographic representation, has resulted in too small of government outlays in the one area that should justify the EU's existence – the provision of these public goods.

### 10.5 The problem of centralization under federalism

The 1949 Constitution of the Federal Republic of Germany did indeed create a *federal* republic. The constitution assigned specific sources of tax revenue, like the personal and corporate income taxes, and wealth and death taxes, to the Länder, the regional governments in Germany. In 1950 roughly 40 percent of all tax revenue in Germany was raised by the regional and local governments. By 1995 this figure had fallen to a mere 7 percent, as the federal government had taken over all major sources of tax revenue (Blankart, 2000).

In 1929 expenditures by the federal government in the United States were less than half of state and local governments' expenditures. Today they are more than 50 percent *greater* than state and local expenditures.<sup>20</sup>

This process of centralization of governmental finances has been repeated again and again in many countries. So common is it that some Europeans refer to it as *Popitz's law* in reference to the German scholar who discussed "the power of attraction of the central government" more than 70 years ago.<sup>21</sup>

In Blankart's (2000) account of the workings of Popitz's law in Germany over the second half of the twentieth century, elected members of the Länder were willing accomplices in the process which stripped their governments of their tax authority. They did so to free themselves of the necessity of having to compete with one another in setting tax rates. The central government effectively helped organize a cartel among the Länder governments to eliminate tax competition.

Grossman and West's (1994) description of the process of centralization in Canada over the same period is very similar to that of Blankart for Germany. A cartel among

<sup>20</sup> See Table 21.1.

<sup>21</sup> See discussion in Vaubel (1994) and Blankart (2000).

the Canadian provinces in conjunction with the central government has significantly reduced differences in tax rates across the provinces. To reduce the competitive pressure Tiebout migration places on the provinces, *equalization* grants from the federal to the provincial governments have been instituted.

Cartels among lower levels of government to eliminate tax competition and migration do not merely help centralize governmental activity; they increase its scale. The mechanics of this process are much the same as those described earlier with respect to the effects of intergovernmental grants on government size. Grossman and West provide econometric evidence that links the centralization of government activity in Canada to the growth in size of its total government sector. Blankart provides more indirect evidence for Germany in the form of a comparison between Germany and Switzerland. Whereas governmental revenue sources have become dramatically more centralized in Germany since World War II, in Switzerland they have become more *decentralized*. During the same time period, Germany's governmental sector has grown 20 percent faster than Switzerland's.<sup>22</sup>

Switzerland's example shows that Popitz's law can sometimes be repealed. Several features of Switzerland's political institutions help account for this achievement. Swiss citizens are able to petition for a referendum to reconsider any major action by their elected government. These referenda have often been used to repeal increases in expenditures and taxes. Some local communities continue to employ direct democracy, thereby eliminating the possibility of those in government substituting their preferences regarding government programs for those of the citizens. Most importantly, Swiss citizens have consistently resisted attempts to weaken their direct control over government as, for example, in their repeated rejection of entry into the EU.<sup>23</sup>

Potentially, the constitution can also help preserve a decentralized federalist structure by clearly assigning different functions and revenue sources to the different levels of government. Such an assignment was present in the 1949 German Constitution, however, and it was simply amended to accommodate the centralization process, as was Canada's (Blankart, 2000; Grossman and West, 1994, p. 22). Section 8 of the U.S. Constitution assigns a short and rather specific list of functions to the federal government, except for the first in the list – to “provide for the common Defense and general Welfare of the United States.” This constitutional assignment of functions succeeded in preventing the central government from encroaching upon the activities of state and local governments for a century and a half, until the Constitution was “amended” by judicial reinterpretation in the 1930s. The lesson one draws from these examples is that a constitutional assignment of functions must be accompanied by procedures that make amendment of the constitution difficult, and that the judiciary must be steadfast in its interpretation

<sup>22</sup> See Table 21.2.

<sup>23</sup> For further discussion of the Swiss case, see Frey (1994). Vaubel (1996) also finds in a cross-national comparison that referenda on federal tax increases deter centralization. He identifies several other factors that deter centralization including, most importantly, the age of the constitutional court.

of the constitution for it to be effective in preventing the erosion of a federalist structure.<sup>24</sup>

*Bibliographical notes*

Classic studies of federalism in the economics and political science literatures, outside of *The Federalist* itself, include Riker (1964), Elazar (1966), Friedrich (1968), Oates (1972), and Breton and Scott (1978).

For a recent survey of the economics/public choice literature, see Inman and Rubinfeld (1997).

Gillette (1997) offers a general discussion of the assignment problem and the conflicts between governments that arise in a federalist system.

Strumpf and Oberholzer-Gee (2000) present evidence that the presence of geographically concentrated groups with intense preferences can influence the assignment of responsibility in a federalist system.

Filippov, Ordeshook, and Shvestova (2001) analyze the problem of avoiding political instability in federalist systems.

<sup>24</sup> Ackerman (1998) recounts the events of the 1930s that removed the constitutional barriers to the federal government's expansion of activities. Aranson (1992a,b) describes how the Supreme Court valiantly, but in the end unsuccessfully, tried to protect the country's federalism from attacks upon it by the Congress. See also Niskanen (1992).

Filippov, Ordeshook, and Shvestova (2001) are highly skeptical about the potential of jurisdictional assignments of governmental functions in the constitution to protect federalism.