Rethinking Federalism

Robert P. Inman and Daniel L. Rubinfeld

Federalism is a founding political principle of the U.S. Constitution and one of our country's recent intellectual exports. In Europe, the former Soviet Union, South Africa, and elsewhere, the view that effective government will involve a well-chosen mix of local and central governmental decision-making is now accepted. Federalism questions—how many local and state governments there should be, how they will be represented in the central government, and how policy responsibilities should be allocated between the central government and the lower tiers—are once again a central research concern of constitutional lawyers, political scientists, and economists alike.¹

America’s federalism debates were initially resolved by the U.S. Constitution. The resolution of the tension over which levels of government should do what has evolved during the past two centuries: from a period of “dualism” (1790–1860) in which states and the central government had comparable responsibilities; through an early period of “centralizing federalism” (1860–1933) in which the still modest federal responsibilities grew; through a later time of “cooperative federalism” (1933–1964), which marked a substantial growth in social programs arising out of the Depression; and finally to a period of “creative federalism” since 1964 in which

¹ For a presentation of our views on the current debates in federalism, see Inman and Rubinfeld (1996b, 1997). For other perspectives on current research on federalism, see Bednar and Eskridge (1994) on constitutional issues; Bermann (1994) on subsidiarity and the European Union; Bird, Ebel, and Wallich (1995) on federalist issues in the former Soviet Union; Ahmad and McLure (1996) on federalism in the new South Africa; and Rivlin (1992) on a federalism approach to leading U.S. policy issues.
the federal government has taken a direct and active role in the problems of state and local governments (Scheiber, 1969).

This most recent period of creative federalism was spurred in part by Walter Heller and Joseph Pechman's call in the 1960s for general revenue sharing from the federal government to the states. Heller and Pechman feared that progressive federal taxation would lead to growing federal budget surpluses and a "fiscal drag" on the economy; their solution was to share these surpluses with the more fiscally needy state and local sector (Heller, 1966; Perloff and Nathan, 1968). What proved particularly "creative" about the period of creative federalism, however, were the arguments for additional federal grants-in-aid offered by state and local government officials and their elected Washington colleagues (Beer, 1972). From 1962 to 1976, the number of separate federal grants programs to state and local governments increased from 160 in 1962 to 412 by 1976 (ACIR, 1978, pp. 25, 32), and the amount of money allocated by these programs over this time rose from $42 billion to $169 billion (in 1996 dollars). From 1976 to 1996, total federal grants to state and local governments has risen an additional $46 billion to $215 billion (in 1996 dollars), a real growth rate for that time of about 1.2 percent per annum (Council of Economic Advisers, Economic Report of the President, 1997, Table B-83). This expansion of centrally financed grants-in-aid drives a significant fiscal wedge between the costs and benefits of financing state and local governments.2

There are currently a number of efforts to check the drift toward centralization in the U.S. fiscal structure and reallocate funding responsibilities for redistributive services from Washington to state capitols. Of course, the current initiatives are not new; many of them date back to the "new federalism" espoused during Reagan's first term. It is too early to tell, but it is possible that these current reforms will mark the beginning of a new period in U.S. federalism. If so, this newest federalism period is likely to be built on an intellectual foundation that reflects recent work in public economics, law and economics, and political economy. In what follows, we sketch out our view of the principles that could form that foundation.

Three Principles of Federalism

Three alternative principles, or models, of federalism can be identified in contemporary debates. In considering the implications of these principles, it is helpful to bear in mind that those who value a federal system typically do so for some mix of three reasons: it encourages an efficient allocation of national resources; it fosters political participation and a sense of the democratic community; and it helps to protect basic liberties and freedoms. The means for implementing these three objectives

---

2 Quigley and Rubinfeld (1996) offer an empirical overview of the changing system of intergovernmental grants. Inman (1988) has estimated that this fiscal wedge between those paying the costs of public services and those receiving the benefits created by federal grants has resulted in a "Harberger triangle" of allocative inefficiency equal to about $.17 per dollar of federal aid distributed.
involves decisions about the institutions of federalism: the number of lower-tier governments, their representation in the central government, and the assignment of policy responsibilities between the vertical tiers of government.

**Economic Federalism**

The principle of economic federalism prefers the most decentralized structure of government capable of internalizing all economic externalities, subject to the constitutional constraint that all central government policies be decided by an elected or appointed "central planner."

This view elevates the goal of economic efficiency to the highest priority; only if two federal structures are equally efficient in the allocation of resources do other goals of federalism—political participation or the protection of individual rights—come into consideration. Oates's (1972) classic *Fiscal Federalism* still provides the most complete description of economic federalism; essentially, the central government is assigned responsibility for those public activities distinguished by significant externalities involving spatially dispersed populations, while local governments have responsibility for those public activities for which such spillovers are limited or absent. Decentralization to small jurisdictions is justified because, as Oates (1994, p. 130) put it more recently, "The tailoring of outputs to local circumstances will, in general, produce higher levels of well-being than a centralized decision to provide some uniform level of output across all jurisdictions . . .".

The appropriate number of local (or lower-tier) governments is specified so that all economies of scale in the provision of public services to households are just exhausted.\(^3\) When public services are pure public goods for which the marginal cost of adding another user will be zero (national defense, basic research), or when there are inefficiencies arising from externalities across jurisdictions, then under economic federalism the central government will be assigned responsibility for those services. However, for public services that become congested as more households use the service—that is, to accommodate additional households at current service levels, additional public facilities must be provided—then relatively small communities are more likely to provide the service efficiently. When the average cost per user of providing a given level of a "congestible" government service just equals the marginal cost of adding one more user, then the community has reached its efficient size. Important public services such as education, police and fire protection, sanitation, recreation, and even public health can be produced efficiently by relatively small communities, perhaps as small as 10,000 households.\(^4\)

\(^3\) This statement assumes that there are sufficient number of each "preference type" of household to achieve the efficient level of public goods provisions. If there are too few households of a particular type then one must balance allocative efficiency—satisfying demands—against technical efficiency—reaching the minimal efficient scale for the community.

\(^4\) On education, see Hanushek (1986); on police services, see Craig (1987); on fire services, see Duncombe and Yinger (1993); on parks and recreation, see Edwards (1990); on sanitation, see Gonzalez, Means, and Mehay (1993). For services where efficient production will require larger user populations, small communities can band together to form purchasing cooperatives, although writing contracts for such cooperatives is a subtle matter; see Williamson (1976). State governments often sanction such "contracts" when allowing local governments to form special districts.
Tiebout (1956) presented the first systematic argument as to how a decentralized federal structure could be used to achieve economic efficiency in the provision of public services; Bewley (1981) made Tiebout's insights precise. In the Tiebout economy, most public services are assumed to be congestible and efficiently provided by small communities. Thus, lower-tier governments are given significant policy responsibilities. Households are assumed to be freely mobile; they shop among local jurisdictions for that community which offers their preferred package of services, taxes, and regulations. In this institutional structure, if any jurisdiction were to provide public services inefficiently, households would move to another jurisdiction that was more efficient. It is this variety and the pressure it imposes on the unfavored communities and states which Justice Brandeis most likely had in mind when advocating local and state governments as "laboratories" for the design of public policies.  

However, when there are significant intercommunity interdependencies (like pure public goods or spillovers), Tiebout's competition among small governments may result in economically inefficient public policies. Potential examples of such inefficiencies include low income assistance (Gramlich, 1985), regulation (Oates and Schwab, 1986), and local income and business taxes (Inman and Rubinfeld, 1996a). The principle of economic federalism assigns the central government the task of correcting such misallocations.

The structure of central government decision-making under economic federalism is relatively simple. A single central planner is elected and charged with providing public goods and correcting intercommunity spillovers. The planner can rely on the voting mechanism to reveal voter (presumably, the median voter) preferences or perhaps apply more sophisticated revelation mechanisms such as auctions (Grossman and Helpman, 1994) or demand-revealing processes (Laffont, 1987).

The central government can provide public goods and correct spillovers in either of two ways: provide the good directly or mandate outcomes (a "quantity" control), or subsidize or tax the local governments to provide the efficient levels of the activity on their own (a "price" control). Central governments currently use both approaches. In the United States, national defense, old-age social security, and environmental protection are directly provided or mandated by the central government, while low-income assistance, interstate highways, and basic research are largely managed through central government price subsidies or matching grants to state and local agencies or nonprofit organizations.

Which of the two approaches—quantity controls or price subsidies—is to be
preferred by the central government depends upon the particular economic circumstances of the public good or intercommunity spillover. Direct provision of the public good or mandates by the central government will be preferred when the social marginal benefit curve of the good or corrected spillover is relatively inelastic and the social marginal cost curve is relatively elastic (Weitzman, 1974; Inman, 1982). Lump-sum grants targeted to a particular service can also be used in this case; if tightly monitored, targeted lump-sum grants are functionally equivalent to direct provision.

Untargeted lump-sum grants may also have a role to play in the efficient federal economy. For example, to ensure the efficient location of private sector workers across fiscally competitive jurisdictions, lump-sum transfers from the residents of the fiscally favored community to the residents in the fiscally disadvantaged locality may be needed (Boadway and Flatters, 1982; Myers, 1990). Further, if local tax administration is inadequate, it may make sense for central government to collect tax revenues for, and then transfer grants revenues to, state and local governments. To avoid the moral hazard of having local governments view such transfers as "blank checks" from the central government, the amount of such grants should be firmly tied to a publicly reviewed and locally decided tax rate. There are even circumstances where efficiency requires a grant from local government to the central government. Boadway and Keen (1996) present an analysis with tax interdependencies, in which tax decisions by local government increase the marginal cost of raising central government revenues. In this case, a grant from local governments to the center internalizes the costs that those governments impose on the central government. Finally, intergovernmental transfers can be used to improve the equity performance of the local public sector. For example, a state may decide that school districts should receive a certain amount of money depending directly on the rate at which they tax themselves—not on the tax base. This "tax base equalization aid" would involve transfers from districts with high tax bases to those with lower bases (LeGrand, 1975). The case for such transfers is strengthened if certain local public services such as education are considered "merit wants" (Musgrave, 1987); categorical matching grants can be designed to increase their provision (Inman and Rubinfeld, 1979).

For most economists, the principle of economic federalism, with its recommended institutions of competitive decentralized local governments and a strong central government to provide pure public goods and control intercommunity externalities, essentially defines what federalism is about. However, the principle has had only mixed success as a guide to economic policy. Its strength has been to articulate how fiscal competition among decentralized local governments can ensure the efficient provision of congestible public services; several recent studies offer empirical support for the proposition that competitive local governments do provide citizens the public services they want at the lowest cost (Brueckner, 1982; Bergstrom, Roberts, Rubinfeld, and Shapiro, 1988; Gramlich and Rubinfeld, 1982; Rubinfeld, 1987). The primary weakness of the principle of economic federalism has been to advocate the central government as the only institution best able to provide
pure public goods and correct interjurisdictional externalities. With our growing understanding of how central government policies are decided, the deference of economic federalism to a strong central government may be excessive. For example, there often appears to be little connection between actual interjurisdictional spillovers and the size or structure of federal grants received (Oates, 1994; Inman, 1988). Alternative principles of federalism, ones which explicitly recognize the potential failings of central government policy-making, should be considered too.

Cooperative Federalism

The principle of cooperative federalism is to prefer the most decentralized structure of government capable of internalizing all economic externalities, subject to the constitutional constraint that all central government policies are agreed to unanimously by the elected representatives from each of the lower-tier governments.

The insights of cooperative federalism spring from the law and economics literature. Like economic federalism, cooperative federalism embraces the goal of economic efficiency as its central objective and advocates the use of lower-tier governments to provide congestible public services. However, cooperative federalism is much less optimistic as to the ability of a central government alone to resolve the intercommunity inefficiencies which might arise. Thus, the principle of cooperative federalism requires all central government policies to be unanimously approved by the elected representatives from each of the lower-tier governments. Since central government political agreements will be achieved through bargaining between all affected parties, any central government policies which are unanimously approved will likely be Pareto-improving. These agreements can take place directly within a central legislative body (Wittman, 1989) or through intergovernmental agreements between subsets of local governments which are then approved by the central government or by some agreed-upon neutral party, like an appointed court (Ellickson, 1979).

Agreements between lower-tier governments will require those who are harmed by the fiscal policy of some other jurisdiction to "compensate" the residents inside that jurisdiction for removing the offending policy. As Coase (1960) and others have pointed out, when there are sufficient benefits to the outside residents from removing a harmful policy, then compensation can be paid to inside residents so that all residents—outside and inside—are better off. In practice, such compensation would be paid through an intercommunity agreement in which jurisdictions raise taxes and pay compensation to their neighboring governments.

---

7 This specification of the interstate bargaining assigns the "property rights" to public policy to the government passing the policy. The alternative is to assign property rights to the affected governments; in this case, governments would have veto power over the actions of others. Each assignment has its problems when information is less than perfect, raising the possibility of extortion. Most constitutions assign the property rights to policy to the governments passing the laws.
which in turn return those funds to groups initially favored by the inefficient policy.\textsuperscript{8} Thus, cooperative federalism views the primary function of the central government as encouraging and enforcing interjurisdictional contracts to provide pure public goods and to correct the failings of lower-tier fiscal competition.

There a number of reasons, however, that cause us to be skeptical that interjurisdictional Coasian bargains can be effective. Arguably the most important source of bargaining failure is the inability of the parties to agree how the economic surplus generated by the bargaining process should be divided (Cooter, 1982), since this may well involve irreconcilable ideas of fairness (Hoffman and Spitzer, 1985; Sutton, 1987). Furthermore, Coasian bargainers may make poor estimates of each other’s threat point or miscalculate the chances that the other party will accept a compromise offer, thus taking a hard line that prevents agreement. Unless costs and benefits are common knowledge, all sides are likely to seek strategic advantage by concealing information (Myerson and Satterthwaite, 1983). Jurisdictions being asked to change may overstate the compensation they require for changing. Finally, if many governments are adversely affected by one state’s public policy, the affected jurisdictions may have a difficult time determining how much each is willing to pay, since each individual jurisdiction will face a free-rider incentive to understate how it is affected and what it would pay, hoping that the other jurisdictions will bear the costs of the change. Strategic interplay becomes ever more complicated as the number of bargaining jurisdictions increases beyond two or the bargain is one of many (Mailath and Postlewaite, 1990).

Enforceability of agreements can also be a problem. In principle, intergovernmental agreements are legally enforceable (Ellickson, 1979), but in practice, when the violating party is a state or local government, the central government’s only recourse may be military action. When the ultimate enforcement mechanism becomes so costly, all sides face incentives to renege on prior agreements.\textsuperscript{9} When jurisdictions are tempted to renege, there is less incentive to reach agreements in the first place.

How well has cooperative federalism done in providing public goods or controlling intercommunity externalities? The overall record has not been impressive. Even agreements among few jurisdictions often fail to achieve fully efficient outcomes (Coates and Munger, 1995), and U.S. states often engage in non-cooperative behavior when significant benefits might arise from cooperation (Kolstad and Wolak, 1983). In fact, the limitations of cooperative federalism were evident from our nation’s beginning. The U.S. Constitution was largely a response to the inability of

\textsuperscript{8} One important application of this Coasian approach to intercommunity externalities is found in the work of Myers (1990) and Krelove (1992), who argue central government grants-in-aid are not necessary to correct misallocations arising from excessive, fiscally-induced relocations. Rather, the “over-populated” community has an incentive to pay the residents of the “under-populated” community not to relocate, and a Coasian agreement to this effect can be written between the two communities to correct the market failure. No central government aid is needed.

\textsuperscript{9} This point was well-appreciated by Hamilton in his critique of the Articles of Confederation in \textit{Federalist} 15: “The consequences of this (The Articles of Confederation) is, that though in theory their resolutions concerning those objects are laws, constitutionally binding on the members of the Union, yet in practice they are mere recommendations, which the States observe or disregard at their option.”
the Articles of Confederation to achieve agreement among the states for financing the defense of the newly independent colonies (Rakove, 1989). As a consequence, Article I, Section 8 of our Constitution explicitly allocates the task of providing that defense to the new national Congress. Nationally provided social insurance, the other major expenditure activity of the U.S. central government, arose too from a failure of U.S. states to cooperatively respond to growing unemployment. During the Great Depression, rather than working together to provide a common level of social insurance, the states chose instead to act alone to keep as many of their current jobs as possible through low taxes and low unemployment assistance (Patterson, 1969, ch. 2 and 3), a problem which still exists today (Helms, 1985; Feldstein and Vaillant, 1994). Under Roosevelt's New Deal, the national government decided to fund some insurance systems directly (the Social Security Act of 1935) and to use matching grants to encourage states to take on the task of providing income assistance for the indigent elderly, the blind, and mothers with dependent children (Patterson, 1969, p. 88; Wallis, 1984).

Finally, the macro-management of the economy, perhaps government's single most important regulatory activity, can also be viewed as the response of the central government to a failure of Coasian bargaining among the states. Individual states do have some power to use fiscal policy to stimulate their economies (Gramlich, 1987; Inman and Rubinfeld, 1994), but the fact that most states are small, open economies severely limits their ability to implement effective aggregate demand fiscal policies. Only an agreement among the many states to coordinate their fiscal policies will work. The struggle in Europe to form a viable monetary union illustrates how difficult such agreements on appropriate macroeconomic policies can be.

To be completely clear, our point here is not that the levels of central government defense spending, social insurance, or macroeconomic stability have always been optimal; rather, it is that if the nation had waited for states to agree unanimously on such policies, economic outcomes would almost surely have been far worse. Our reading of the historical and contemporary evidence does not provide much support for the claim that lower-tier governments can solve their important collective action problems on their own through unanimous Coasian agreements. If economic federalism seems too biased in favor of centralization, cooperative federalism seems to bias the fiscal constitution too far in the other direction.

**Democratic (Majority-Rule) Federalism**

The principle of democratic (or majority-rule) federalism is to prefer the most decentralized structure of government capable of internalizing all economic externalities, subject to the constitutional constraint that all central government policies are agreed to by a simple (51 percent) majority of elected representatives from lower-tier governments.

Like the principles of economic and cooperative federalism, democratic federalism also embraces the use of lower-tier governments to provide congestible public services, and again, the number of lower-tier governments is determined by the technology of public services. With regard to its views on the economic perfor-
mance of the central government, however, democratic federalism stands between economic federalism and cooperative federalism. Unlike economic federalism, it does not implicitly assume that the central government will provide public goods and regulate interjurisdictional spillovers efficiently. In contrast to cooperative federalism, only majority-rule—not unanimity—is required to make a decision. Of course, there is no guarantee that policies chosen by a majority-rule legislature will be efficient either. Democratic federalism seeks to balance the potential efficiency gains of greater centralization in a world of local spillovers and pure public goods against the inefficiencies which might arise when a democratic central legislature sets policies (Tullock, 1969; Inman and Rubinfeld, 1996b). Considering this trade-off requires a specification of the federal institutions of government. What is the extent of local representation in the central legislature? Should there be an independently elected executive with veto powers? How should policy responsibilities be assigned to the different tiers of government?

In thinking about how the institutions of a democratic central government might be specified, it is useful to contrast two commonly used approaches to legislative decision-making. The first assigns agenda-setting powers to a small subset of members, say the speaker of the house or a key legislative committee. Other members in the legislature then simply vote—up or down—on the items in the approved agenda. Most likely, policies will be approved by a bare majority—a minimal winning coalition—in this strong agenda-setter legislature (Baron and Ferejohn, 1989). A second strategy shares agenda-setting powers among all members, giving each legislator a right to select a most preferred policy in that policy area most germane to the legislator’s constituents. This second approach to legislative decision-making involves each legislator deferring to the preferred policies of all other legislators, provided the other legislators defer to the legislator’s own policy requests (Weingast, 1979; Niou and Ordeshook, 1985; Weingast and Marshall, 1988). The guiding norm here is one of deference—“You scratch my back, I’ll scratch yours”—and it typically results in legislative proposals which are approved nearly unanimously. For this reason such legislatures are often called “universalistic.”

There are reasons to believe that minimum winning coalition legislatures may be more economically efficient than universalistic legislatures (Inman and Fitts, 1990). But this alternative environment will not arise unless the representatives themselves prefer to do business in a minimum winning coalition environment. A single legislator choosing between the closed rules of a bare majority, minimal-winning-coalition legislature or the more open rules of a universalistic legislature managed by a norm of deference will typically favor the more open rules. With closed rules, you must belong to a winning coalition to have your constituent’s projects approved; without strong political parties or additional side-deals, the probability of being in that coalition is at best 50:50. In constituent-based politics, it may be better to have the sure slice of a smaller pie under universalism, than run the risk of no slice at all under minimal-winning-coalition politics.

Universalistic legislatures operating under a norm of deference run a signifi-
cant risk that their chosen policies will be economically inefficient. The essential problem is that each legislator chooses a program that will disproportionately benefit their own constituency, with the costs paid by residents of all jurisdictions. Because of this cross-subsidy, each legislator has an incentive to ask for too much of their own preferred good or regulation (Weingast, Shepsle, and Johnsen, 1981; Persson and Tabellini, 1994). The subsidy becomes larger, and the potential inefficiencies greater, the greater is legislative representation, or equivalently, the smaller are the legislative jurisdictions relative to the nation as a whole. The legislative norm of deference allows these inefficiencies to stand, not just for one jurisdiction but for all jurisdictions represented in the legislature. Inman (1988) and Inman and Fitts (1990) offer some tentative evidence on the magnitude of the allocative inefficiencies created by such a legislature.

What can be done to strike a more appropriate balance between the gains of centralized assignment and the costs of this assignment when the legislature is inefficient? One set of options is to reform the central government's legislative process in ways that would discourage an inefficient universalistic legislature, perhaps by strengthening the hand of political parties over members' decisions (Aldrich, 1995) or by increasing executive powers (Fitts and Inman, 1992).

Alternatively, one might adjust the institutions of federalism. For example, if given the legislative process and size of the legislature, the assignment of policy responsibility to the central government is less efficient than retaining those responsibilities at the local level, even with associated spillovers, then constitutional assignment can reallocate the activity to the lower tiers of government. In effect, this is what President Reagan sought to achieve through the informal influence of his presidency with his 1982 budget and his "new federalism" reforms. However, since the Reagan reforms asked the central legislature to surrender influence over spending, it is perhaps not surprising these efforts did not survive his presidency.10

Another institution of federalism that can be adjusted is the extent of representation of local governments to the national legislature. Does every community have direct representation or does the constitution combine communities, with groups of local jurisdictions electing one representative? If the central government's legislature operates under a norm of deference, economic inefficiencies are likely to be greater the larger the legislature and the smaller the unit of representation.

---

10 Formal constitutional assignment of functions to lower-tier governments might also be tried, but such constraints, when effective, are often quite "heavy-handed." Either the assignment excludes everything from the central level or, if exceptions are written, potentially nothing. Recent efforts to write an effective, but flexible, federal balanced budget amendment is a case in point. The current Rehnquist Supreme Court is struggling to find a more nuanced interpretation of assignment in our Constitution. Since Garcia v. San Antonio Metropolitan Transit Authority (469 U.S. 528 [1985]), the Court has abandoned trying to define assignment by governmental function. In United States v. Lopez (115 S. Ct. 1624 [1995]) the Court embraced a process approach to evaluating congressional actions which affects states. The Court now requires national laws to explicitly state a national interest (for example, interstate commerce) being rationally served by the legislation. This does not seem a particularly high hurdle. Inman and Rubinfeld (1997) suggest another approach to raising the bar using a Federalism Impact Statement (or "FIST").
tation; so for efficiency, smaller legislatures and larger units of representation may be preferred (Inman and Fitts, 1990; Gilligan and Matsusaka, 1995). The efficient legislature is unlikely to be very small, however. Legislatures serve as bargaining halls and help to reveal preferences. A small legislature puts all the burden for preference revelation on the candidate election, losing the potential efficiency gains which come with face-to-face deal-making. Setting the size of the efficient legislature requires us to balance the gains from having more voices heard against the risk that too many bargainers means only inefficient or unstable deals are done (Buchanan and Tullock, 1962, ch. 7).

Which Principle of Federalism Should One Choose?

Constitutions establish the rules for collective decision-making. The unique contribution of a federal constitution is to allow for multiple tiers of governments, each with a domain of policy responsibilities. Federal constitutions must specify the number of lower-tier governments, the representation of those governments to the central government, and the assignment of policy responsibilities between the upper and lower tiers.

For economic efficiency, all three principles of federalism embrace the logic of the Tiebout model and the use of lower-tier governments to provide congestible public services. It is in the specification of central government representation and in the assignment of pure public goods and spillovers to the central government that the three principles of federalism may disagree. Economic federalism has all local governments select one central government representative—a "president-planner"—and then assigns all pure public goods and spillovers as central responsibilities. It is the job of the president-planner to set and administer central government policies efficiently, presumably guided by principles of efficient (second-best) public finance. Cooperative federalism gives each local government one representative to the central government and then allows those representatives to fashion Coasian agreements to improve the welfare of their citizens. Agreements may involve all local governments—for example, an agreement to enforce agreements—or only a subset of governments. Cooperative federalism assigns all public goods and spillovers locally, unless local governments voluntarily agree to centralize. Democratic federalism jointly decides representation and assignment as part of an effort to balance the efficiency gains of central government provision

---

11 Although Buchanan and Tullock's (1962) classic Calculus of Consent does not discuss federalism explicitly, they do advocate an approach to constitution-writing much like that suggested by democratic federalism (p. 112): "As we have suggested, the costs of reaching agreement, of bargaining, are, from a 'social' point of view, wasteful. One means of reducing these costs is to organize collective activity in the smallest units consistent with the extent of the externality that the collectivization is designed to eliminate."

12 The president-planner's objective may, or may not, correspond to an ethically appealing social welfare function, but whatever the objective, the president-planner can be assumed to pursue it with the most efficient use of the policy instruments available. For a discussion of how such a president-planner might be chosen and set policies in a democracy, see Besley and Coate (1997).
against the inefficiencies which might arise when central legislatures decide what that level of provision should be (Inman and Rubinfeld, 1996b). Large legislatures will do a good job representing preferences of all citizens but may foster inefficiently large ("universalistic") budgets. Small legislatures are less representative but the budgeting may be less prone to excessive spending.

However one evaluates the economic efficiency performance of federal constitutions, it must be recognized that the federal institutions chosen will have important implications for political participation and the protection of individual rights and liberties, two other constitutional values central to past and current federalism debates. James Madison's arguments in Federalist 10 for a strong but highly representative central government as the best protector of individual rights helped to define the representation and assignment outcomes in our Constitution. Current concerns in the European Union over political participation and the Union's "democratic deficit" may have a similar effect by elevating the European Parliament to greater institutional importance in the new EU's constitution (Garrett and Tsebelis, 1996).

Finally, there are good reasons to think that efficiency, participation, and the protection of individual rights may at times conflict, and that setting the institutions of the federal constitution will require hard choices. For example, a strong central government built on the principles of economic federalism or democratic federalism with a small legislature is likely to be the relatively more efficient federal structure for the provision of public goods and spillovers. However, such a structure may shortchange the valued goal of political participation which is typically best served by giving small local governments stronger central government representation and more policy responsibilities (Frug, 1980; Inman and Rubinfeld, 1997). Individual rights might also be threatened by a strongly centralized federal constitution, Madison's arguments notwithstanding. Legal scholars concerned with the protection of individual political rights (McConnell, 1987; Rapaczynski, 1986) have strongly criticized from a rights perspective the Supreme Court's recent Garcia decision giving the central government carte blanche for setting public policies, while Easterbrook (1983) and Weingast (1995) both argue forcefully for decentralized policy assignments as the best means for protecting individual property rights. Those charged with selecting a principle of federalism must understand and then balance these potentially important tradeoffs between economic efficiency, political participation, and individual liberties.

13 After years of frustration in trying to establish a workable assignment principle based upon the Tenth Amendment, the Supreme Court in Garcia v. San Antonio Metropolitan Transit Authority (469 U.S. 528 [1985]) gave the task of deciding the allocation of federal and state policy responsibilities to the central government, arguing that state representation in the U.S. Senate would adequately protect state policy interests. Two more recent Supreme Court decisions have placed some modest limits of what Washington can do. In United States v. Lopez (115 S. Ct. 1624 [1995]) the Court required Congress to make explicit the connection between national regulations of state activities and a national objective, while in New York v. United States (505 U.S. 144 [1992]) and more recently in Printz v. U.S. (1997 U.S. Lexis 4044 [June 27, 1997]), the Court overturned the use of unfunded mandates on the states.
Welfare Reform: Budget Cutting or A New Federalism?

The U.S. Constitution is a broadly representative but centralized federal constitution. In response to Madison's concerns that the new democracy be representative of all the people, the Constitution requires representation of the populace in the House of Representatives and equal voices for all the states in the Senate. To avoid the collective action problems inherent in the Articles of Confederation, the Constitution through the Tenth Amendment assigns all policy powers to the majority-rule central government. Ultimately, this representative central government will decide which tier of government will set America's economic policies.

We may be at the start of a quiet revolution in Washington's view of how to assign policies. In summer 1996, President Clinton signed the Personal Responsibility and Work Opportunities Reconciliation Act of 1996, known more generally as the Welfare Reform Act of 1996. The act has two primary objectives: 1) to reduce welfare roles by providing families currently on welfare with the means and the incentives to seek work; and 2) to end welfare's 60-year status as a nationally funded entitlement. Of central interest to us here is the second objective, since it potentially represents a significant shift in the responsibilities of different levels of government for the provision of low income assistance, and how Washington may wish to manage our federal relations generally.

The Welfare Reform Act of 1996 has two potentially important consequences for federal-state relations in the provision of low-income assistance; it saves the federal government money and it breaks the federal-to-state-to-recipient entitlement for low income assistance. Prior to the passage of the Welfare Reform Act, the federal government shared directly in the financing of states' decisions for welfare through an open-ended matching grant for aid to families with dependent children (AFDC), for job training (JOBS), and for aid for homeless children (Emergency Assistance). If the state spent more on welfare, the federal government shared in that expenditure at the federally set matching rate. The Welfare Reform Act replaced these matching aid programs with a single block grant called Temporary Assistance for Needy Families (hereafter TANF). TANF breaks the direct call of states and their low income households on additional central government dollars. Importantly, this shift from matching to block grants was not necessary to achieve the budgetary savings of the Welfare Reform Act. Of the $54 billion that welfare reform will save over the next six fiscal years, $46 billion will come from provisions which deny Social Security Income (SSI) supplements and food stamps to legal immigrants (U.S. House Committee on Ways and Means, 1996, p. 1332). The savings from the consolidation of AFDC, JOBS, and Emergency Assistance into the TANF block grant equals only $7.8 billion over six years. To put it another way, the

---

14 For an analysis in this journal of the possible success of the Welfare Reform Act in reaching its first goal, encouraging work by current welfare recipients, see Blank (1997).
same level of savings in federal spending from TANF could have been achieved with an (approximate) 8 percent cut in federal matching rates, say from .50 to .46 for rich states and from .78 to .72 for poor states.\textsuperscript{15}

These cost figures, together with much political rhetoric, suggest that the real fiscal target of the Welfare Reform Act of 1996 was not lower federal spending, but the federal-state relationship for how poverty dollars are budgeted. Since its inception during the Depression, AFDC had guaranteed each eligible individual (originally children only, but later mothers were included) a state-determined stipend with state spending supported by an open-ended matching grant from the federal government. Federal standards for a minimum level of the stipend and funding through a federal, open-ended matching grant guaranteed at least minimal AFDC payments to all eligible households. Higher state spending is allowed and the federal government will match that spending. Originally, the federal government matched state spending at a uniform rate of $1 of federal money for each $2 of state spending, implying a federal "matching rate" of .33. Subsequent reforms have raised federal matching rates to .50 ($1 federal for each state $1) for the very richest states and to .78 for the poorest states ($3.50 federal for each state $1). The welfare reform ends this federal guarantee of fiscal support for state spending by consolidating AFDC, JOBS, and Emergency Assistance into the TANF block grant. Replacing welfare matching aid with the TANF block grant frees the federal budget from the welfare entitlement and makes states responsible for the full cost of each additional dollar spent on low income assistance.

Further, the Welfare Reform Act of 1996 gives states wide latitude to determine program eligibility and benefits, and largely removes federal government regulations for low income assistance. TANF dollars can be reallocated away from direct income support to programs which promote job training, child care, prevention of teen pregnancy, and marriage. Given these alternative uses of the TANF block grant, TANF monies are likely to be highly fungible out of direct income support, if that is what a state's politics prefers.\textsuperscript{16} The only requirements imposed on the

\textsuperscript{15} This is only approximate, because a cut in the federal matching rate for state spending will reduce the states' own spending and thus have "second-order" effects further reducing total federal outlays. These second-order effects would be small. Craig and Inman (1986) estimate the price elasticity of own state spending with respect to changes in the federal matching rate to equal .12; thus an 8 percent reduction in the matching rate will reduce state spending by less than 1 percent.

\textsuperscript{16} Legally, a state will be free to reallocate those dollars over 80 percent of the state's TANF allocation, as the Welfare Reform Act of 1996 requires that 80 percent of the TANF allocation be spent on approved welfare-related activities. Once the 80 percent target is met, TANF monies may be allocated to other state activities outside the welfare budget, including general tax relief. However, state and local governments are very clever in labelling programs to circumvent federal regulations—one Pennsylvania high school district reclassified fourth year AP Spanish as a bilingual language program to qualify for federal low-income education aid—so that the Welfare Reform Act's maintenance of effort regulation may prove only a weak constraint. If so, state welfare spending will become fully fungible and each additional dollar spent on welfare will imply an opportunity cost to the state of $1. Finally, TANF funding may diminish in importance over time. While generous in the near-term, TANF funding is not indexed and there is only modest protection if welfare roles rise during a deep recession. WRA 96 establishes a contingency fund with $2 billion in reserves to be allocated through TANF when a recession occurs. The fund is
states are those which restrict the size of the state welfare roles. TANF funds cannot be used to support assistance for families whose members have received assistance for five or more years, although states can exempt 20 percent of their caseload from this requirement. Adults receiving TANF funds must "engage in work" within two years; states may choose a shorter grace period. By 1997, 25 percent of single parents receiving TANF funds (50 percent of two-parent families) must be engaged in work for at least 20 hours per week; by 2002, 50 percent of single parents (90 percent of two-parent families) must be working, although states retain some flexibility in defining "work."

What will happen to low-income assistance now that fiscal responsibility has been moved back to the states? Only tentative predictions can be drawn from the numerous studies of state financing of AFDC spending. Two offsetting incentives are at work. First, the Welfare Reform Act of 1996 eliminates AFDC's matching rates for additional state spending, which raises the effective price (the net of matching aid price) of an additional dollar of low-income assistance for a state from \((1 - \text{matching rate})\) to 1; poor states will see the effective price of an additional dollar of poverty assistance rise from .22 \((= 1 - .78)\) to 1, while richer states will see their prices rise from .5 \((= 1 - .5)\) to 1.17 Even though most studies estimate modest price elasticities of demand for welfare spending, ranging from \(\text{-.02 to -.50} (\text{Ribar and Wilhelm, 1996})\), having the effective price of welfare increase by 354 percent (for poor states) or 100 percent (for rich states) implies potentially large consequences. For example, a price elasticity of -.20 would imply nearly a 70 percent cut in spending in poor states and a 20 percent reduction in richer states.

On the other side, although the welfare reform removes the price incentive for state welfare spending, it substitutes a block grant of a nearly equal dollar amount. Direct estimates of the elasticity of welfare spending to lump-sum grants range from .01 to .30, with lower estimates corresponding to more "fungible" grants-in-aid (Craig and Inman, 1986; Inman, 1979). Using these estimates we expect the TANF block grant equal to the state's current loss in matching aid to increase welfare spending in poor states from 4 percent (for a block grant elasticity of .01) to perhaps as much as 30 percent (for a block grant elasticity of .30) and from 2 percent (.01 block grant elasticity) to 20 percent (.30 block grant elasticity) in rich states.18

Combining these effects, a price elasticity of -.20 implies a decline of welfare spending in the poor states from 40 to 66 percent (that is, a drop of 70 percent from the price effect and an increase of either 4 or 30 percent from the TANF

---

17 The precise specification of the price of welfare spending is a bit more involved, depending on which political economy model and policy interactions are being estimated. For a useful review of the various price specifications for state welfare spending, see Chernick (1996).

18 The above-estimated percentage changes following a TANF block grant are based upon these grants elasticities and the average welfare spending and grants levels reported in Craig and Inman (1986). The larger effects for poorer states reflect the fact that they will receive a larger percentage increase in block grant aid from the new TANF grant.
block grant) and a fall in the rich states from 0 to 18 percent (a 20 percent decline
from the price effect plus either a 2 or 20 percent increase from the TANF grant
effect). The larger declines occur when federal welfare aid is highly fungible from
welfare spending into other areas of the state budget. It remains to be seen how
much of a decline, if any, will finally occur. That there are “maintenance of effort”
provisions in the Welfare Reform Act, requiring that 80 percent of TANF grants be
spent on welfare-related programs, suggests that at least a majority in Congress is
concerned that declines may be significant.\footnote{19}

Of the three principles of federalism specified above, the principle of coop-
erative federalism seems to us to provide the strongest rationale for the federalism
reforms undertaken by the Welfare Reform Act of 1996. For example, the principle
of economic federalism, aimed as it is at internalizing all relevant externalities,
would advocate either full and direct central government provision or the use of
matching aid as the most economically efficient ways of adjusting for fiscal spillovers
between jurisdictions. When providing low-income assistance, there are two sets of
spillovers to be internalized, those on the cost side and those on the benefit side.
On the benefit side, Boadway and Wildasin (1984) emphasize that residents of one
state may derive benefits from the provision of low income assistance in another
state. On the cost side, Gramlich (1985) emphasizes that mobile taxpayers and
mobile poor drive up the costs to states of providing a dollar of benefits to their
own poor residents, since higher taxes and benefits may drive away the residential
or business tax base while attracting additional poor. For both reasons, the private
marginal cost to a state of providing $1 in additional welfare benefits will be larger
than the social marginal cost. If grants are used, the matching rate should be set
to internalize for state politicians these two types of spillovers.\footnote{20}

\footnote{19} Even so, we remain skeptical that this maintenance of effort provision alone will have much effect in
preventing states from allocating resources away from poor households if such reallocations are the
politically preferred allocation. Central government rules on state budget allocations are typically very
difficult to enforce.

\footnote{20} One can offer a back-of-the-envelope calculation that the recent matching rates used in AFDC were
reasonable adjustments for expected spillovers. The appropriate formula for a matching rate that corrects
for spillovers is: $m = (1 - 1/ \rho \phi)$, where $m$ is the matching rate, $\rho$ is the factor (presumably greater than
1) by which one must multiply the private marginal benefit received from a change in welfare benefits
to account for the benefits felt by those in other jurisdictions, and $\phi$ (also presumably greater than 1)
is the factor by which one must multiply the costs for a state of changing welfare benefits to make up
for the cost spillovers incurred by scaring off or attracting businesses and the poor. Helms (1985) esti-
mates the elasticity of state incomes to a tax-financed increase in welfare benefits to be about .1 and
Blank (1988) estimates the elasticity of welfare immigration to an increase in benefits to also be about
.1. These two effects have an additive impact, so that the private cost to a state of providing an extra $1
of social assistance will be about 20 percent higher than the $1 social cost, or about $1.20; thus $\phi$
can be taken to equal 1.2. On the benefit side, if we assume that if one state spends an extra $1 on benefits,
it raises the utility of each of the other 49 states by an average of 2 cents, then the social marginal benefit
of having one state spend an additional $1 will be nearly $2 (the $1 spent plus a spillover of
49 × 2 cents). Thus, $\rho$ equals about 2. Using these estimates of $\phi$ and $\rho$, the matching rate formula
implies an efficient welfare matching rate to control spillovers of .58, close to AFDC’s average matching
rate before the welfare reforms.

The more general conclusion that matching rates are an appropriate policy is not very sensitive to
The principle of democratic federalism also will support the continuation of direct central government provision or matching aid, again to internalize interstate spillovers from welfare, though this principle might well advocate reductions in direct support or matching rates if the evidence points to inefficient current policies because of political logrolling. After all, welfare spending subsidizes poor families. In this light, the politics of these subsidies should be no different than the politics of all subsidies: tax deductions for charitable giving, farm price supports, tariffs, or low cost loans to cities, states, and non-profit institutions. For some evidence that welfare has been part of this wider subsidy logroll, see Ferejohn (1983). While lower matching rates might be justified by the principle of democratic federalism, the continued presence of cost and benefit spillovers when states provide welfare still recommends direct central government provision or a targeted matching rate greater than zero.

Nor can the fiscal reforms in the Welfare Reform Act of 1996 be well-explained by a shift towards greater concern for the competing constitutional values of political participation and the protection of individual liberties. Though one might justify the back-to-work welfare reforms from a principle of personal liberty, it is hard to see how the fiscal reforms replacing matching aid with a fungible block grant enhances personal freedoms. Nor is overall political participation likely to be greatly affected by the reforms. On one side, increasing state discretion over welfare policies does take an important step towards local control of policies, which should enhance participation. However, setting the matching rate at zero significantly raises the cost to the middle class voters of including lower-income families in any budget coalition. This may well reduce overall participation in deciding state fiscal choices. Finally, as long as efficiency is still a valued social objective and the central government is allowed to set policies, then matching aid, not fungible block grants, is the better policy.

The principle of cooperative federalism seems to account best for all the major fiscal reform features of the Welfare Reform Act of 1996. Cooperative federalism advocates the decentralization of policy to the lower tiers of government with the federal government's role limited to encouraging efficient agreements between jurisdictions. The Welfare Reform Act seems to do just that. First, states, not cities or counties, are chosen by the act as the appropriate unit of government to receive the new welfare responsibilities. This assignment recognizes the significant mobility of lower income families and tax base within metropolitan areas (Inman, 1992) as well as the possibility of significant benefit spillovers between local governments (Pauly, 1973). States are generally the right unit of government to internalize such metropolitan-wide spillovers from welfare policies. Further, the Welfare Reform Act

---

the numbers chosen for $\rho$ and $\phi$. If the benefit of spending $1$ in one state raises utility by only half a cent in 49 other states, and the cost spillover parameter is 1.1, a still significant matching rate of 0.27—almost the matching rate when AFDC was first approved in 1935—is justified. As state economies become more open, as they surely have over the past 60 years, then the spillover parameters $\phi$ and $\rho$ will increase and the efficient AFDC matching rate will rise as well, as it has over the past 60 years.
directly removes a possible interstate spillover from welfare policy by allowing states to impose residency requirements for the receipt of benefits, although this provision seems sure to be challenged in court. In short, the Welfare Reform Act seems designed to assign welfare responsibility to those lower-tier governments (states) which internalize as many welfare externalities as possible. For any welfare spillovers that might remain, the act grants states wide latitude to set welfare policies, discretion which should facilitate interstate agreements. Finally, and perhaps most decisively for the view favoring cooperative federalism as the rationale for welfare reform, the act drops matching grants in favor of fungible block grants, thereby removing an important impediment to efficient bargaining between states.  

Seen in the light of our three principles of federalism, the recent fiscal reforms in welfare policy appear to be a significant institutional experiment with an alternate paradigm of federalism, one which emphasizes the ability of states, not the central government, to handle cross-jurisdiction spillovers. If the experiment proves successful, then other central government policies like Medicaid (another entitlement poverty program), environmental and business regulation, infrastructure spending, and perhaps even Social Security and Medicare may become candidates for fiscal decentralization too.

Conclusion

Rethinking federalism means rethinking the terms under which sovereign citizens or states join together to form a "more perfect union." Whether one is struggling to form a political union for the first time (European Union, South Africa, Russia), deciding to break away from an existing union (Quebec), or to reform a stable one (United States), decisions must be made along each of the institutional dimensions which define the federal constitution: the number of lower-tier governments, their representation to the central legislature, and the assignment of policy responsibilities between the center and lower tiers. Whatever federal constitution is selected will have important implications for the valued goals of government: economic efficiency, political participation, protection of rights.

Most countries will want to mix and match their federal institutions, depending on how they view the performance and weigh the goals of government. For example, for economic efficiency, lower-tier governments might best be assigned responsibility for congestible public services; both economic theory and the available evidence seem to support this allocation. For the problems of pure public goods and interjurisdictional spillovers, the principle of economic federalism recommends that these goods be assigned to the central government. However, the prin-

21 As first noted by Buchanan and Stubblebine (1962), central government using taxes or subsidies alone to internalize spillovers alters the marginal incentives of agents who are bargaining, leading to an inefficient, post-bargain allocation. In other words, Pigovian taxes without compensation is a substitute for, not a complement to, the Coasian bargain.
principles of cooperative federalism and democratic federalism are less clear on the point. The ability of central governments to provide pure public goods efficiently may depend crucially on how representatives are selected for the national legislature. Locally chosen representatives may place parochial interests above the collective interest in efficient public goods provision. Cooperative federalism argues for assigning pure public goods and spillovers to the local level, much as the new welfare reforms have done; cooperative federalism counts on the ability of interjurisdictional bargaining to allocate such goods better. The principle of democratic federalism retains central government assignment for pure public goods and spillovers but argues for a more rough and ready representation of local interests at the national level, like recommending nationally elected representatives to a majority-rule national legislature. Finally, other goals for government will be considered too, such as political participation and protection of individual rights. These important constitutional values might favor extensive local representation in the national legislature, even allowing for the potentially significant efficiency costs imposed by a large constituent-based legislature, or they might favor many local governments and the local assignment of public goods and spillovers, even recognizing the possibly large inefficiencies imposed by spillovers left unresolved because of incomplete interjurisdictional bargains.

As we rethink federalism, we must recognize—as did our Founding Fathers—that the selection of the institutions of federalism necessarily carries with it a balancing of these competing social goals of economic efficiency, political participation, and the protection of individual rights and liberties.

We would like to acknowledge with thanks the financial support by the National Science Foundation (Inman) and the Fisher Center for Real Estate and Urban Economics (Rubinfeld). We must also mention the financial and intellectual support of the Center for Advanced Study in the Behavioral Sciences (SES-9022192), which provided us with a unique opportunity to begin our own rethinking about federalism. Thanks too to Richard Bird, Steve Coate, Brad De Long, Michael Fitts, Geoffrey Garrett, Alan Krueger, Wallace Oates, Eric Posner, Harvey Rosen, Timothy Taylor (particularly), and Michael Wiseman for their helpful comments on an earlier draft.

References


Bergstrom, Theodore, Judith Roberts, Daniel L. Rubinfeld, and Perry Shapiro, "A Test for Efficiency in the Supply of Public Education," *Jour-...


This article has been cited by:

