On the Theory and Practice of Fiscal Decentralization

Wallace E. Oates*

IFIR Working Paper No. 2006-05

May 2006

*Professor of Economics, University of Maryland, College Park, MD. I am grateful for helpful comments on an earlier draft of this paper to Emma Galli, Charles McLure, Fabio Padovano, Harvey Rosen, John Wallis, and David Wildasin.
On the Theory and Practice of Fiscal Decentralization

Abstract

Wallace E. Oates*
Department of Economics
University of Maryland
3105 Tydings Hall
College Park, MD 20742
oates@econ.umd.edu

*Corresponding author
On the Theory and Practice of Fiscal Decentralization

Wallace E. Oates

The traditional theory of public finance has made a strong case for a major role for fiscal decentralization. This case is based on an improved allocation of resources in the public sector. And it has four basic elements. First, regional or local governments are in a position to adapt outputs of public services to the preferences and particular circumstances of their constituencies, as compared to a central solution which presumes that one size fits all. Second, in a setting of mobile households, individuals can seek out jurisdictions that provide outputs well suited to their tastes, thereby increasing the potential gains from the decentralized provision of public services (Tiebout 1956). Third, in contrast to the monopolist position of the central government, decentralized levels of government face competition from their neighbors; such competition constrains budgetary growth and provides pressures for the efficient provision of public services. And fourth, decentralization may encourage experimentation and innovation as individual jurisdictions are free to adopt new approaches to public policy; in this way, decentralization can provide a valuable “laboratory” for fiscal experiments.

However, this basic economic rationale for decentralization of the public sector is not quite so simple and compelling as it appears. Some of the more recent literature provides, first, a thoughtful and provocative critique of the traditional view of fiscal decentralization, and, second, some new approaches that reveal its dark side, especially in practice. There is emerging, in short, a broader perspective on fiscal decentralization that raises some serious questions about its
capacity to provide an unambiguously positive contribution to an improved performance of the public sector.

My purpose in this paper is twofold. First, I want to review the basic theory of fiscal decentralization. There are some loose ends to the traditional argument that open up some intriguing issues. Second, I want to turn to some of new literature on fiscal discipline in multi-level government. This literature has focused attention on some basic and destructive forces that can undermine the economic performance of a relatively decentralized public sector. I find it helpful to begin by revisiting a Decentralization Theorem that I formulated long ago. As a point of departure, I want to explain briefly why I introduced the proposition and the rationale for its particular form and proof.

1. On the Decentralization Theorem and Its Motivation

In my book (1972), I proposed a straightforward decentralization theorem that formalizes the basic efficiency argument for the decentralized provision of certain kinds of public goods. The theorem lays out a set of sufficient conditions for the decentralized provision of these goods to be Pareto-superior to a centralized determination of public outputs. It goes as follows:

*The Decentralization Theorem:* For a public good—the consumption of which is defined over geographical subsets of the total population, and for which the costs of providing each level of output of the good in each jurisdiction are the same for the central or for the respective local government—it will always be more efficient (or at least as efficient) for local governments to provide the Pareto-efficient levels of output for their respective jurisdictions than for the central government to provide any
specified and uniform level of output across all jurisdictions (p. 35).

The theorem seems almost trivially obvious. But I thought it useful at the time to have a more precise statement of the general idea for two basic reasons. First, it provides an explicit set of conditions under which decentralized provision is to be preferred (on efficiency grounds) to centralized control. As it turns out, some of the ingredients of the theorem are more problematic than I appreciated at the time. A reconsideration of them has opened up a whole range of alternative approaches and new interpretations. And, second, as explored in the appendix to the chapter (1972, pp.54-63), one approach to proving the theorem provides a straightforward algorithm (taken from Barzel 1969) for determining the magnitude of the welfare gains from decentralization. This has the added virtue of identifying the key parameters that determine the size of the welfare gains from decentralized finance. David Bradford and I (1974) actually employed this approach in an exercise in which we measured the welfare gains from the decentralized determination of school budgets in New Jersey.

Let me begin by going back to the conditions I imposed in the theorem. There are three issues that I want to address here. The first, and perhaps the most straightforward, is the matter of interjurisdictional spillover effects (or externalities). Second, there are the closely interrelated issues of the mobility of individuals across jurisdictions and the precise nature of the public good. The third matter, one (which for good reason) has gotten much attention, is the assumption in the theorem that the central government is constrained to provide a uniform level of output across all jurisdictions. I will take up each of these issues in turn.

First, the theorem assumes that the benefits from the consumption of the public good are limited to those individuals within the jurisdiction where it is provided. There are no
interjurisdictional spillover effects associated with the good. Note that it is still a public good in the sense that it is jointly consumed (i.e., consumed in the same quantity) by all residents of the jurisdiction, but it has no impact on the well-being of those outside its borders. We are all familiar with the allocative distortions that typically occur when externalities are present, and this assumption simply rules out any such external effects.

There is an interesting and related matter that arises when this condition is violated. Suppose we are considering an assignment problem for a particular public good. The issue is whether or not to make the responsibility for its provision a central, or a “local,” government responsibility. If there are interjurisdictional spillover benefits present, this argues for assigning the function of providing the public good to the central government. But there is a cost here. If the central government provides a uniform output across all jurisdictions, then the welfare gains from diversifying outputs in accord with local preferences and conditions are lost. There is thus a basic tradeoff in making this determination between the gains from accounting for the spillovers through centralized provision versus the losses from tailoring outputs to local circumstances. In my book, I provided an informal diagrammatic treatment and discussion of this tradeoff in chapter 2.

It is interesting that a tradeoff of a similar spirit emerges from some of the more recent literature that is developing a so-called “second-generation theory of fiscal federalism.”¹ Drawing on a framework in which there is asymmetric information and incomplete contracts, Paul Seabright (1996), for example, finds that centralization can increase welfare by improving “coordination” (by accounting for spillover effects), but this comes with a loss of

¹For a survey of this new literature, see Oates (2005).
“accountability” (local control over local outcomes). As Timothy Besley and Stephen Coate (2003) put it, “All of this notwithstanding, the key insight remains that heterogeneity and spillovers are correctly at the heart of the debate about the gains from centralization” (p. 2628). At any rate, my theorem disposes of the problem by simply assuming that there are no external effects across jurisdictional lines.

A second key issue in the theorem has to do with the precise nature of the public good. Often when treating public goods in general or those that are provided by a central government, the assumption is made that they are pure public goods in the Samuelsonian sense that an additional person can consume the output of the good without reducing the consumption of anyone else; in standard parlance, consumption is said to be “non-rival.” National defense is the ubiquitous example. Most of the literature in local public finance, however, has adopted a different conception of public goods. So-called “local public goods” (as envisioned, for example, in the Tiebout model) are not pure public goods: they are subject to costs of congestion. Often the assumption is made that they are “fully congestible” (i.e., doubling the size of the group requires a doubling of inputs in order to maintain the level of consumption unchanged).

---

2The more recent literature provides a number of new perspectives on this assignment issue (e.g., Lockwood 2002). In some instances, they provide a somewhat different slant on the tradeoff between centralization and decentralization. Eckhard Janeba and John Wilson (2003), for example, construct a model in which in which tax competition among local governments is the source of inefficiency in local provision, while inefficiencies at the central level have their source in decisions by a minimum winning coalition in a central legislature. In another approach, Jan Brueckner (2004) poses the tradeoff in terms of the inefficiencies stemming from local taxation versus the gains from Tiebout sorting under local provision.

3Another (and traditional) way of addressing this problem is through a set of (Pigouvian) taxes and subsidies that serve to internalize the external effects. A system of properly designed matching grants, for example, can provide the appropriate incentives to expand outputs of public services to encompass the benefits to those outside the jurisdiction.
It is interesting that the Decentralization Theorem says nothing about this matter. It need not, because the theorem is sufficiently general to encompass a whole range of types of public goods, including both pure public goods and local public goods; they can be rival or non-rival in consumption. All that the theorem requires is that whatever the nature of these public goods, the cost of providing a given level of output in a particular jurisdiction is the same, be the provider the central or local government. This does, of course, rule out any economies of scale from central government provision.

However, this issue does raise some fundamental matters about how we conceptualize (or model) the decentralization problem. Of particular importance is what we assume about the mobility of economic units across jurisdictional boundaries (my second issue). The theorem fineses this matter by assuming a complete absence of mobility. One of the conditions specified in the theorem is that “the consumption [of the public good] is defined over geographical subsets of the total population.” This implies that individuals are not able to move across jurisdictional lines; were this not so, the geographical subsets of the population would change as people altered their location. This condition is a bit stronger than it need be. What is required here is that there be no mobility in response to changes in fiscal parameters (i.e., changes in jurisdictional outputs of public goods or levels of taxes). In short, this is explicitly a non-Tiebout kind of world.4

In this respect, the Decentralization Theorem is quite restrictive. Several strands of the literature address the mobility issue explicitly and reach quite different sorts of results. Where

4An interpretation of the theorem that allows mobility with respect to other non-fiscal variables does complicate matters a bit, for it makes the population of a given jurisdiction (and hence the pattern of efficient outputs of public goods) specific to a particular locational equilibrium.
mobility with respect to fiscal (and perhaps other) parameters exists, the precise character of the public good becomes crucial. In the Tiebout model, for example, there is costless mobility; individuals seek out a jurisdiction that provides exactly the level of output of the public good that they wish to consume. In so doing, they reveal their preferences for “local” public outputs and generate a Pareto-efficient outcome in the public sector. But the properties of the public good are critical here. For the Tiebout model to work, the public good must be fully subject to congestion (i.e., for each level of output, if we double the size of the population, we must double the quantity of inputs to maintain output at an unchanged level), and individuals must pay a tax (equivalent to a price) equal to the marginal cost of extending the output of the public good to encompass an additional consumer. Moreover, individual location decisions must not be constrained by anything other than fiscal variables (not, for example, by location of employment). For these reasons, the Tiebout model is typically taken to apply to metropolitan local finance, a setting in which individuals work in a particular urban setting that provides a wide choice of city and suburban jurisdictions in which to reside.\(^5\)

If the set of local tax-prices facing mobile individuals does not faithfully reflect marginal costs, then problems arise because location decisions can involve external costs on other persons in both the originating jurisdiction and destination (Buchanan and Wagner 1970). In an influential paper, Flatters, Henderson, and Meiszkowski (1974) departed from the Tiebout assumption by taking the public good to be a pure public good within the jurisdiction in which it is provided and by requiring that individuals must work in the same jurisdiction in which they reside.\(^5\)

\(^5\)For a review and assessment of the Tiebout model, its evolution, and its contribution, see Oates (2006).
This issue has received extensive attention in Canada, where the treatment of oil revenues in the oil-rich provinces has been a contentious matter at both national and regional levels. This has provided some of the motivation for a major Canadian program of fiscal equalization among the provinces. For a formal analysis of the problem, see Robin Boadway and Frank Flatters (1982).

My basic point here is that the assumptions we make about mobility and about the nature of the public good interact in important ways in determining the kind of outcomes we obtain. And we must think carefully about the settings in which we apply these models. The Tiebout model, for example, may provide a reasonably good description of behavior in a metropolitan setting (Fischel 2001). In contrast, the FHM model has been helpful in understanding some of the issues where different regions have widely varying endowments of natural resources that affect both the productivity of private agents and the availability of tax bases.6

My third issue concerns the nature of central government provision of local public goods. It may help to be a bit more precise on this. In equation (1), let $G$ be a vector of outputs of public goods where an element, $g_i$, is the level of public output in the $i^{th}$ jurisdiction:

\[
G = \{g_1, g_2, ..., g_n\}
\]

---

6This issue has received extensive attention in Canada, where the treatment of oil revenues in the oil-rich provinces has been a contentious matter at both national and regional levels. This has provided some of the motivation for a major Canadian program of fiscal equalization among the provinces. For a formal analysis of the problem, see Robin Boadway and Frank Flatters (1982).
The Decentralization Theorem equates “centralized provision” with a uniform level of public outputs (say $g_0$) across all jurisdictions; the theorem thus requires that $g_i = g_j = g_0$ for all $i,j$. This obviously served my purposes quite well in the theorem by providing a benchmark against which to compare (favorably) a varied, decentralized pattern of outputs of public goods. But this condition has been contested in the literature. Why should we expect the central government to provide equal levels of public outputs in all jurisdictions? Why can’t the central government simply provide the efficient level of output in each jurisdiction?

The rationale (or defense) for this constraint has taken two forms. First, there is an information issue: the argument is that local governments are closer to their constituencies. They have a knowledge of local preferences that the center cannot easily come by. It is hard for a central government to know the diverse preferences in the myriad of jurisdictions that constitute the country as a whole. In short, there is an asymmetry of information: local governments know local preferences; the central government does not. The second line of defense for this assumption is more political in character. It suggests that there are political constraints on the center that prevent it from varying local outputs. There is a national sense of equal treatment that makes it difficult for a central government to provide more generous levels of public outputs in some jurisdictions than in others.

But these arguments are not fully convincing. We see in the real world a variety of “pork-barrel” projects under which central agents deliver special public programs to their own constituencies. But even at a more formal level, there is really nothing, in principle, to prevent

\footnote{For purposes of measuring the welfare gains from fiscal decentralization, I took $g_0$ to be arithmetic mean of the demands of all the individuals in the country (Oates 1972, pp. 59-63.)}
the central government from obtaining the needed information on local cost functions and preferences (albeit costly to gather such information) and then to provide the Pareto-efficient level of output in each jurisdiction. In fact, central government responsibility for local public goods does not, in practice, rule out such sensitivity to local conditions. Under some administrative structures, the center may confer on their local agents both the responsibility and the authority to be responsive to local circumstances in their determination of local programs. Albert Breton and Pierre Salmon (forthcoming) describe the French system of administration under which prefects (appointed by the center) have the capacity and incentive to adapt national policies to local conditions. Thus, an outcome that is decentralized in spirit can, in principle, emerge from a structure of government with little in the way of real local autonomy.8

Some of the more recent literature has taken an entirely different tack on this whole issue. Lockwood (2002) and Besley and Coate (2003), for example, envision the determination of local outputs under a centralized system as taking place through the decisions of a central legislature, whose members are elected representatives of the individual jurisdictions. In such a setting, the elements of the vector $G$ in equation (1) are set by the central legislative process, which can involve various kinds of logrolling and electoral maneuvers. This can give the problem a quite different character. If, for instance, the vector $G$ is determined centrally by a

---

8There is a kind of semantic issue here. In my earlier book (1972), I characterized an economic definition of federal government as one in which outcomes at decentralized levels “are determined largely by the demands for these [public] services of the residents...of the respective jurisdiction” (p. 17). I contrasted this with political definitions of federalism that focus more on political autonomy, often expressed in the form of a constitution. The point here is that one can associate “decentralization” either with the nature of the outcomes or with the process by which these outcomes are determined. In this sense, I suppose one could characterize the prefect system (from an economic perspective) as possessing important “federal elements.”.
minimum winning coalition in the legislature, allocative distortions occur that work to the
disadvantage of those jurisdictions outside the coalition. Nevertheless, if spillovers in the
 provision of local public goods are sufficiently pervasive, a centralized outcome may still
dominate a decentralized one. The more general point here is that the Decentralization
Theorem invokes a very strong condition or constraint on the central provision of public goods
that makes it fairly easy to establish a presumption in favor of the decentralized provision of
these goods (assuming an absence of significant spillover effects or large economies of scale).
The problem of centralized versus decentralized provision can become much more complex
when centralized outcomes are characterized in a more thoughtful way.

2. On Measuring the Welfare Gains from Fiscal Decentralization

The Decentralization Theorem also suggests a straightforward way to generate a cardinal
measure of the welfare gains from the decentralized provision of public goods. Using an
approach suggested by Yoram Barzel (1969), one can provide a proof of the theorem by simply
maximizing the sum of consumer surpluses from provision of the good (Oates 1972, pp. 59-63).
This approach can be employed to measure the welfare gains from decentralization by simply
calculating the difference between the level of aggregate consumer surplus under decentralized
and centralized provision of the good.

Such an exercise, in addition to providing some sense of the size of the potential welfare
gains from fiscal decentralization, has the further virtue of identifying the key parameters upon
which these gains depend. Returning to the theorem, suppose that the costs of providing a unit
of the good per person are the same across all jurisdictions. The only source of welfare gains in
this case comes from diversifying outputs across jurisdictions in accordance with the differing demands for the local public good. It is straightforward to show for this case (Oates 1997) that the magnitude of the welfare gains depends on the variation in demand across jurisdictions and on the price elasticity of demand for the good. The greater the differences in the efficient levels of output across the jurisdictions themselves, the more there is to gain from allowing each jurisdiction to provide its own efficient output level. Perhaps less obvious is the fact that the price elasticity of demand is a key parameter here. In particular, the more price inelastic the demand for the public good, the greater the gains from decentralization. 9

David Bradford and I (1974) employed this method in an exercise in which we estimated the potential welfare losses from moving from a system of decentralized school finance to one which imposed equal spending per pupil across all jurisdictions. Using a body of data on school finance from New Jersey, we first estimated a demand curve for per-pupil expenditure. We then took the observed spending per-pupil in each school district to be that which the local population desired (i.e., the efficient level). As our benchmark for centralized provision, we simply aggregated the school budgets across all jurisdictions and divided by the total number of pupils to get our figure for per-pupil expenditure for the case of centralized school finance. Since there was wide variation in expenditure per pupil within our sample ($295 to $547) and since our

9It is interesting to note that this is precisely the opposite result from the case of deadweight losses in taxation. For the tax case, we all learn that the welfare losses from an excise tax on a specific good (at least in a partial-equilibrium framework) become larger as demand curves become more price elastic, because a given tax causes a larger response (distortion) in terms of the change in quantity purchased. In contrast, the welfare losses from fiscal centralization are greater, the more price inelastic is the demand curve, because given changes in quantities on the horizontal axis moving away from an efficient level generate greater losses in consumer surplus where demand curves are relatively steep.
estimate of the price elasticity was low (-0.36), it is not surprising that we found the welfare losses from the centralization of school spending to be quite large. Our calculations suggested that the average deadweight loss associated with moving a dollar of expenditure from a high-spending district to a low-spending district was on the order of 50 cents.

This estimate assumes, of course, that there are no external effects associated with local school spending, an assumption one might well question. More generally, this approach has an admitted tendency toward large estimates of the welfare losses from fiscal centralization. First, it uses as a benchmark the assumption of uniform provision across jurisdictions under a centralized regime; it thus provides no scope whatsoever for the central authority to adapt outputs to local circumstances. And, second, it assumes fully efficient local decision-making by presuming that the observed outputs under decentralization are at Pareto-efficient levels. It is worth noting, in this context, that these kinds of measures are also likely to produce large estimates of welfare losses from centralization because the existing estimates of the price elasticity of demand for local public goods typically suggest highly price-inelastic demands. There is now a large body of econometric estimates of demand functions for a wide range of local public services, and they typically find price elasticities on the order of -0.3 to -0.5.10

Even if there is no variation in demands across jurisdictions, fiscal decentralization can produce welfare gains where costs vary, since with given demands, differing costs will result in differences in efficient levels of output.11 Such variation in the costs of providing local public

---

10 For surveys of this econometric literature, see Rubinfeld (1987) and Oates (1996).

11 The Decentralization Theorem assumes that cost functions for local public services are identical across all jurisdictions. It thus focuses solely on differences in demand as the source of welfare gains from fiscal decentralization.
services can come from either of two sources. First, there may be actual differences in the production functions among jurisdictions. The costs, for example, of providing clean, clear roads are likely to be lower in areas where there is relatively mild weather than where winter (and summer) storms make road-clearing operations a more formidable task. Second, there is an altogether different and interesting source of variation in the cost per person of providing local services: the congestion properties of the public service. Suppose, for example, that the local public good is a pure public good within the jurisdiction in which it is provided. Then it will be less expensive per person to provide a given output in a relatively populous jurisdiction. And, hence, the efficient level of output, other things equal, will be higher in such a jurisdiction.

It is easy to see for the case of varying costs that the gains from fiscal decentralization depend on the magnitude of the variation in costs across jurisdictions and, once again, on the price elasticity of demand (Oates 1997). But for this case, these welfare gains are greater, the more price elastic are the demand functions—just the opposite of the case where differences in efficient levels of output have their source in variation in demand. Thus, the impact of the price responsiveness of demand on the magnitude of the welfare gains from decentralization depends on whether the divergences in efficient levels of local outputs have their source in differences in demands across jurisdictions or differences in costs.

An especially dramatic case of welfare losses from centralization with their source in cost differentials emerged a few years ago in the debate over the Environmental Protection Agency’s

12This is like the case of distorting taxes in that the source of the variation manifests itself on the price, rather than the quantity, axis.
new “arsenic rule” for safe drinking water. The removal of contaminants from drinking water involves processes with enormous economies of scale (in terms of population). It is typically much less expensive per person to provide safe drinking water in heavily populated areas than in those with fewer people. In this case, the EPA introduced a new and much tougher standard for arsenic concentrations in drinking water. The estimated benefits were quite modest. The problem was that the average cost per household to attain the standard varied from under $1 per household per annum in large water systems (like New York City) to over $300 per household in the smallest systems. In benefit-cost terms, the new rule may have been defensible for large water districts, but (based on then-available estimates) the rule promised large welfare losses for smaller districts (Oates 2002). The setting of environmental standards in the U.S. often involves a single standard for all jurisdictions (a one-size-fits-all approach); following this precedent, the EPA’s new standard was imposed on all water districts. However, as this case makes clear, a more decentralized approach to the setting of standards for many dimensions of environmental quality makes lots of economic sense.

The traditional economic case for fiscal (and regulatory) decentralization is thus founded on the potential welfare gains from diversifying local public outputs in accordance with local circumstances (encompassing both differences in preferences and costs). And the existing (if small) empirical literature suggests that these gains can be quite large. An optimal fiscal structure from this perspective is one in which a central government provides outputs of those

---

13 More generally, the removal of contaminants from drinking water involves processes that exhibit huge differences in cost per person across water districts of different population sizes. See U.S. Congressional Budget Office (1997) and Dinan, Cropper, and Portney (1999) for studies that explore the case for decentralized standard setting for safe drinking water.
public goods that are truly national in scope and decentralized levels of governments determine the appropriate levels of outputs of those public services whose consumption is limited to residents of their respective jurisdictions. In practice, however, there is much more to the story. To achieve efficient outcomes requires a set of public decision-making institutions that provides the needed incentives for effective budgetary choices. Much of the more recent literature suggests that this much easier said than done. And to this I turn next.

3. On Fiscal Institutions and Decentralization

The earlier literature in fiscal federalism recognized the potential for distorted outcomes under certain spatial patterns of benefits from local public services and particular forms of taxation. In his seminal work on tax-exporting, Charles McLure (1967, 1969), for example, provided estimates of the extent to which various taxes levied in one state (or jurisdiction) were actually borne by residents elsewhere. McLure went on to develop the welfare implications of tax-exporting, pointing out that since local decision-makers could expect outsiders to finance some portion of their local budget, there would exist incentives for excessive local expenditure. But it is in the more recent literature that the “dark side” of fiscal decentralization has been spelled out more fully. 14 This has been prompted, it seems to me, by two developments. First, we have seen several major fiscal and economic collapses (such as those in Argentina and Brazil), as well as more low-grade but persistent fiscal malfunctioning, that have important roots in the vertical structure of the public sector and, more particularly, in the behavior of

14Remy Prud’homme (1995) and Vito Tanzi (1996) raised a range of specific concerns regarding the dangers inherent in fiscal decentralization.
decentralized fiscal authorities. This has prompted a careful examination of just what has gone wrong. Second, there have been important advances in the theory of fiscal structure and its operation that draw upon the new theory of the firm (e.g., Cremer, Estache, and Seabright 1996). The recognition that asymmetric information, incomplete contracts, and the associated “agency costs” manifest themselves in public, as well as private, sector institutions has opened up a whole new perspective on fiscal decentralization.

A major issue in this literature is the “raiding of the fiscal commons.” This is not just a matter of corruption.\textsuperscript{15} Such destructive behavior derives directly from the incentives that the political and fiscal system create for both public officials and the electorate. The very structure of intergovernmental finance in certain countries makes such behavior almost inevitable.

The source of these “raids” is the presence of “soft budget constraints.” The term itself comes from the seminal work of Janos Kornai (1979, 1980). Kornai originally introduced the term to describe the setting of state-owned enterprise in socialist countries, where managers could depend on the central government to bail them out of financial difficulties. This created an environment in which these managers could preside over chronic financial losses with little to fear from higher authorities. The scope of the term has, more recently, been broadened to encompass a range of economic entities whose financial losses will be underwritten by some form of “supporting organization” (Kornai, Maskin, and Roland 2003). For my purposes here, the term can be understood to refer to lower levels of government who operate with the

\textsuperscript{15}There has emerged a large literature on corruption in the public sector, much of it empirical in character. One of the important questions posed in this literature is whether corruption is likely to be a more serious problem at the central or at decentralized levels of government. My reading suggests the jury is still out on the answer to this question.
expectation that their fiscal deficits will be covered by a higher level (often the central) government. This is thus a world in which provincial (state) governors or local mayors can expect fiscal bailouts from higher-level authorities; they need not, in short, keep their fiscal houses in order.

The key issue here is the source of such expectations. Since such bailouts can obviously undermine responsible fiscal decision-making, how is it that a fiscal system can come to embody such a perverse set of expectations? The literature typically explains this phenomenon in terms of some kind of sequential, game-theoretic framework (e.g., Wildasin 1997; Qian and Roland 1998; Goodspeed 2002). In the first stage of the game, the central government commits itself to a sensible, no-bailout position; it states explicitly that it will not come to the fiscal rescue of lower levels of government running deficits in their budgets. In the second stage of the game, decentralized officials must decide whether or not this claim is credible. And, as the literature makes clear, there are plenty of reasons why these officials may not find the center’s claim to be persuasive. First, the central government presumably cares about the welfare of the citizenry (either for altruistic reasons or in order to be re-elected). The failure of a provincial or local government can have serious consequences not only for the well-being of its residents, but for others as well. As David Wildasin (1997) has argued, such fiscal malfunctions can have important spillover effects on other jurisdictions. Moreover, in the complexity of hierarchical politics, governors or mayors may be in a position to shift the blame for their fiscal crises onto central public officials. Thus, the very political survival of central incumbents may well depend on their coming to the aid of lower-level fiscal authorities.

This is further complicated by the potentially important “insurance role” of the central
government. As Lockwood (1999) and others have argued, there will be times when decentralized governments come under fiscal distress that has purely external origins. The central government can, under such circumstances, provide valuable assistance that will soften the impact of these external “shocks.” But (as with insurance in the private sector) there can be a real moral-hazard problem here. Especially in a contentious political setting, it may not always be easy to distinguish clearly fiscal deficits that have their source in external shocks from those that result from poor fiscal management.

For a complex of reasons, decentralized governments may thus decide that the center’s claim to a no-bailout position is not credible. In consequence, they may proceed in the second stage of the game to expand the budget beyond levels for which they have funding in the expectation that central officials will come to their rescue with grants or loans. [If, alternatively, decentralized governments find the center’s commitment to a no-bailout strategy to be persuasive and behave responsibly in fiscal terms, the game ends at this juncture.] Confronted by these provincial or local deficits, the central authority must then decide whether or not to come to the rescue with additional financial resources. And, as we have discussed, there may often be compelling reasons (both economic and political ones) for a bailout.

This raises the crucial issue of the structural source of these perverse expectations. What are the elements in a fiscal and political system that create an environment which undermines fiscal discipline? There is a not a simple answer to this question. In a fascinating and enlightening set of case studies, Jonathan Rodden, Gunnar Eskeland, and Jennie Litvack (2003) suggest that soft budget constraints typically have a multiplicity of sources that encompass existing fiscal institutions, the structure of the political system, the absence (or weakness) of certain key
markets, and the specific history of intergovernmental fiscal relations in the country. In short, each nation has its own story to tell. But there are some important lessons here. For example, as Rodden at al. point out, soft budget constraints frequently arise in settings where fiscal responsibility is ill defined. In some countries, there exist serious ambiguities about which level of government is responsible for providing certain services (such as health care or pensions) or at least the funding of them. Where spending and revenue authority and responsibility are not clearly defined, there may be good reason for governors or mayors to expect fiscal assistance from higher levels. In short, Rodden et al. find that “unclear or shared responsibilities have a cost in terms of accountability and incentives” (p. 16).

Soft budget constraints are also more likely where decentralized levels of government have weak tax systems and rely heavily on higher levels of governments for fiscal transfers. Rodden et al. describe such a setting as involving “transfer dependency.” In order to make the tough fiscal decisions and weigh the benefits against the costs of new or expanded programs, public officials need to be in a position of raising the monies from their constituencies through their own state and local tax systems. A heavy reliance on transfers creates incentives for turning to an expansion of these transfers rather than increasing taxes in one’s own jurisdiction.

Another major element of soft budget constraints is often debt finance. One way to fund a potential deficit (if assistance is not directly forthcoming from above) is through the issue of bonds. There are numerous instances, for example, in which powerful provincial governors (in countries like Argentina) have had access to the public banking system (or other state-owned enterprises) to absorb their bond issues. In such a setting, public officials can simply sell the requisite bonds to regional or national banks to cover prospective deficits. Soft budget
As Wildasin (2004) discusses, there is typically an intimate relationship between intergovernmental transfers and public debt. An increase in transfers from the center may often be financed by an increase in the center’s deficit so that central borrowing may effectively substitute for borrowing at decentralized levels of government.  

The absence of a strong and effective set of private markets constitutes a further impediment to a system of hard budget constraints. Well developed, and efficient capital markets, for example, can provide needed fiscal discipline by imposing higher borrowing costs and limiting access to credit for provincial or local governments that perform poorly. Similarly, an effective set of land markets in the presence of mobile factors of production can support sound fiscal decision-making through the capitalization of superior or poor fiscal performance (encompassing the quality of local outputs and levels of local taxation) into local property values. Thus, a strong system of private markets can itself be an important contributor toward a hardening of budget constraints.

In addition, there is the critical matter of history and precedents. The U.S. experience is instructive in this regard. Reviewing the U.S. history with a system of relatively hard budget constraints, Robert Inman (2003) argues that the U.S. record which, since the founding of the Republic and the War of 1812, is essentially clean of any strategic bailouts, owes much to an important historical episode. In the 1840s, a wave of defaults occurred encompassing eight states and the Republic of Florida as a result of poorly conceived public investments in transportation and banking projects. The federal government explicitly rejected efforts by these

---

16 As Wildasin (2004) discusses, there is typically an intimate relationship between intergovernmental transfers and public debt. An increase in transfers from the center may often be financed by an increase in the center’s deficit so that central borrowing may effectively substitute for borrowing at decentralized levels of government.
Another mechanism for the hardening of budget constraints is a system of explicit rules on fiscal behavior that introduces measures such as a balanced-budget restriction on the current account. The limitations on deficit spending and levels of public debt embodied in the Stability and Growth Pact as part of European monetary union appear to have had some impact on restraining public spending in various EMU countries (Franco et al, 2003). There is also a large literature on fiscal rules in the U.S. states that suggests that appropriately designed measures can have some effect in constraining budgetary behavior.

Another mechanism for the hardening of budget constraints is a system of explicit rules on fiscal behavior that introduces measures such as a balanced-budget restriction on the current account. The limitations on deficit spending and levels of public debt embodied in the Stability and Growth Pact as part of European monetary union appear to have had some impact on restraining public spending in various EMU countries (Franco et al, 2003). There is also a large literature on fiscal rules in the U.S. states that suggests that appropriately designed measures can have some effect in constraining budgetary behavior.
two sides of the ledger. A condition of vertical fiscal imbalance (or “transfer dependency”) is said to exist where own-revenue systems are weak and lower level governments rely heavily on transfers from above.

The issue here is the extent and nature of a sensible system of intergovernmental grants. It is important to recognize at the outset that nearly all systems of multi-level government finance incorporate significant grant programs and that there is an explicit, and quite compelling, rationale for their role. Public-finance economists have traditionally seen such grants as serving two basic purposes. First, in the Pigouvian tradition, they provide a needed incentive for lower level governments to expand public outputs that have important spillover effects into other jurisdictions. The appropriate fiscal instrument here is a matching grant that serves to internalize the spillover benefits. Second, there exist major programs of fiscal equalization in most federal countries (although not at the federal level in the U.S.) that serve to transfer resources to relatively poor regions. These are typically justified on redistributive grounds or, more recently, in terms of providing a more “level playing field” for interjurisdictional competition. In addition, to these two functions, it is often pointed out that the central government may have a kind of comparative advantage in raising tax revenues. Nationwide taxes (with uniform provisions and rates), so the argument goes, have fewer distortionary effects on flows of mobile resources than do state and local taxes with differing rates, definitions of tax base, etc. Moreover, this permits a higher degree of progressivity in the tax structure, which promotes distributive objectives. There may also be some (modest) economies of scale in tax administration. So the argument here is that we can have a better overall tax system if we rely more heavily on the central government and use transfers to provide some portion of state and
local funds.

The tough question here is whether or not an extensive system of such intergovernmental transfers is consistent with a fiscal structure with hard budget constraints that encourages efficient fiscal decisions. This is not an easy question to answer. Let me consider first the case against intergovernmental transfers. We have already seen the concern raised by Rodden et al. and others that such transfers can easily become an instrument that governments abuse in the pursuit of fiscal bailouts. In addition, there is some troubling evidence that intergovernmental grants often do not function as the normative theory would have them do, even in the context of a system of relatively hard budget constraints. Studies in the U.S., for example, suggest that where such transfers presumably have a role in internalizing interjurisdictional spillover effects, they have not done so very effectively. The design of many of these programs runs counter to this objective (Oates 2004). For example, for many years, the federal matching grants for interstate highway construction (a major grant program) involved a cost-sharing formula under which the federal government’s share was 90 percent of construction costs. Such a federal matching share surely far overstates the fraction of benefits accruing to other states. Moreover, most federal matching-grant programs (in the U.S. at least) have had “caps” at fairly modest levels of expenditure. Once the cap is reached, federal cost-sharing ceases, and there is no incentive at the margin for any expansion of the program. More generally, Inman (1988), in a study of U.S. federal grant programs to state and local governments, found that the basic economic objectives of efficiency and equity do not take us very far in explaining the actual structure and functioning of the federal grant system. He had much more success in explaining
(econometrically) the pattern of grant allocations using a model of political influence.\textsuperscript{18}

As we have noted, another basic role for intergovernmental transfers is fiscal equalization. In many federal countries like Canada, Australia, and Germany, there are major programs to transfer resources from more wealthy jurisdictions to those with relatively small tax bases (and, perhaps, high fiscal need). But these programs too can have perverse and unintended consequences. In an intriguing recent study, Fabio Padovano (forthcoming) looks at the process of income convergence among regions, where he contrasts the experience of the U.S., a country with essentially no fiscal equalization by the federal government, with that of Italy, where there have been large fiscal transfers from the wealthy North to the South. Padovano finds that in the U.S. the process of income convergence has proceeded expeditiously as economic theory would predict with a movement of industry and employment to relatively low-wage areas and a consequent narrowing of interregional income differentials over time. This process of income convergence has not happened in Italy. Padovano argues that this is largely the result of the transfer system which has muted the incentives for the factor movements that generate convergence.\textsuperscript{19}

While a more systematic and comprehensive review of the experience with intergovernmental transfers goes beyond the scope of this paper, there is surely enough here to suggest that these grant systems have often not been designed properly, frequently have not functioned very well, and in some cases have had perverse, if unintended, consequences. The case against grants,

\begin{itemize}
  \item[\textsuperscript{18}]
  See also Shama Gamkhar (2002).
  \item[\textsuperscript{19}]
  Likewise, Ronald McKinnon (1997) and McKinnon and Thomas Nechyba (1997) have suggested that fiscal equalization has, in several countries, impeded the standard process of “equalization through competition.”
\end{itemize}
based on their actual use, is not easily dismissed.

At the same time, I am reluctant to throw out intergovernmental transfers as simply a bad idea in practice. They have an important *raison d’etre*, and there are ways to address their ills. Moreover, they are firmly entrenched; we probably are destined to live with them, whether we want to or not. The issue is how to make them work better. The two key matters here are those of proper design and proper use. In terms of design, they must provide a set of incentives consistent with their objectives. Most notably, they must not encourage poor fiscal practices. As one example, Rodden et al. (2003) point out that the mandated German formula for fiscal equalization makes it clear to the smaller and poorer states that poor fiscal performance will be rewarded with increased transfers (p. 443). This is obviously a recipe for undermining fiscal discipline. In practice, moreover, a grant system needs to be transparent and predictable. There needs to be a well understood set of rules such that the system cannot be “gamed” to one’s own advantage. This may well involve limiting the discretion of the central government in the distribution of funds so as to preclude political manipulation of grant distributions. This is obviously a topic for an extended treatment in its own right. The design and operation of an effective system of intergovernmental grants ranks, in my view, right at the top of the research agenda in fiscal federalism.

A second issue on which I want to comment briefly is the intimate relationship between fiscal behavior and politics. The fiscal system is embedded in a larger political system, and the way in which fiscal decentralization works clearly depends on political, as well as fiscal, institutions. This has been the subject both of some important theoretical and empirical work. I mentioned earlier the papers by Lockwood (2002) and Besley and Coate (2003) in which central
legislatures, composed of locally elected representatives, determine the vector of local public outputs described in equation (1). Depending on how the legislature behaves (e.g., a minimum-winning coalition or a “cooperative” outcome), the pattern of local outputs will exhibit different sorts of distortions. Paul Seabright (1996) has taken a somewhat different tack in a model in which an incumbent central government determines the vector of local outputs in such a way as to maximize the probability of its re-election. The crucial variable for each local jurisdiction in this approach is the probability that the jurisdiction is decisive (or “pivotal”) in determining the election of the central government. Localities for which this probability is relatively high obviously tend to get more favorable treatment in terms of relatively high levels of local public outputs. These are simply a few examples of a larger body of ongoing theoretical work in which we find that local fiscal outcomes depend in crucial ways on how the political system works.

The critical role of political structure also comes through clearly in recent empirical work. In the illuminating collection of country studies to which I have made frequent reference, Rodden et al. (2003) find that certain properties or characteristics of the political system have profound implications for the presence or absence of hard budget constraints for decentralized levels of government. Most notably, they find that in countries, where the central legislature is itself a loose coalition of regional interests, budget constraints tend to be relatively soft with sometimes devastating results (as in the cases of Argentina and Brazil). In such a setting, fiscal outcomes have sometimes been the result of extensive logrolling behavior among regional interests that result in fiscal bailouts for decentralized authorities. This work leads to the rather paradoxical and ironical finding that for fiscal decentralization to function effectively, there is a real need for a strong central government with a sufficient presence to be able to resist opportunistic moves by
provincial and local interests. A weak center, in short, is often an integral part of a fiscal system with soft budget constraints and poor performance.

My third, and final, topic is the even bigger picture that brings together the whole constellation of fiscal structure, the political system, and market institutions. The general point here is that, in the end, fiscal performance emerges from this larger system. Barry Weingast and his co-authors have stressed the fundamental interdependence of the private and public sectors in their work and have been able to tie various elements together in terms of what they call “market-preserving federalism.” On the one hand, an effective system of markets requires a supportive political and fiscal environment. One of the basic threats to efficient markets is the encroachment of powerful political interests. “Preserving markets requires that the state be effective yet limited” (Qian and Weingast, pp. 83-4). On the other hand, efficient markets contribute to hard budget constraints and improved fiscal performance. From the Weingast et al. perspective, this crucial interdependence is best addressed in terms of a decentralized public sector in which regulatory and fiscal powers are diffused among state and local entities. Competition among these decentralized authorities restricts their capacity to intrude on the efficient operation of markets. Efficient markets, in turn, constrain the fiscal activities of public-sector agents. As we noted earlier, in a setting with mobile factors and relatively efficient land and credit markets, poor fiscal performance will penalize a jurisdiction through restricting local government access to credit markets, raising the interest rates on local public debt, and through the capitalization of the poor quality of local outputs and high taxes into local property values.

20See, for example, Weingast (1995), Qian and Weingast (1997), McKinnon (1997), and Weingast (forthcoming).
There is thus a kind of mutual reinforcement between markets and fiscal decentralization that enhances the performance of both the private and public sectors of the economy and that gives them a “self-enforcing” character. Weingast (1995) suggests that we can find historical examples of such systems in 18th century England and in the United States in the 19th century.

The essence of such systems is a combination of fiscal and market institutions that provides a set of incentives to individual agents for efficient behavior. A relatively decentralized public sector, characterized by competition among jurisdictions in a setting of a common market without barriers to trade, can provide a powerful inducement for public decision-makers to behave in ways that promote the welfare of their constituencies and sustain the efficient performance of private markets.

There is, however, no detailed, universal blueprint for the optimal system. As we have seen, the construction of a viable and efficient structure involves a delicate balancing act. A self-sustaining system must, on the one hand, be sufficiently decentralized to promote competition among jurisdictions and to limit the capacity of the center to undermine the efficient operation of markets. At the same time, however, the central government must be strong enough to resist credibly moves by decentralized agents to raid the commons. The right balance is likely to vary somewhat from nation to nation depending on the particular history and character of the country.\textsuperscript{21} The political economy of fiscal decentralization is thus a very complicated and

\textsuperscript{21}In their illuminating study of fiscal decentralization in Canada, Richard Bird and Almos Tassonyi (2003) argue that an objective examination of fiscal structure suggests that Canadian provinces should have relatively soft budget constraints. Yet provincial fiscal behavior has not been characterized by opportunistic raiding of the commons. They suggest that this is largely the result of a political system that limits provincial influence in the central legislature and of strong markets. In addition, they note that national temperament may itself be important. They contend that “...generally prudent fiscal behavior has become an institutionalized norm in Canada, in part
challenging issue. As Wildasin (2004) indicates, it offers us a rich research agenda with a potentially enormous payoff.
References


The IFIR Working Papers Series

Other titles in the series:


IFIR Working Papers contain original research contributed by scholars affiliated with the Institute for Federalism and Intergovernmental Relations at the Martin School of Public Policy and Administration at the University of Kentucky, Lexington, Kentucky. Visit the IFIR web site at http://www.ifigr.org to download IFIR Working Papers and for other information about IFIR.