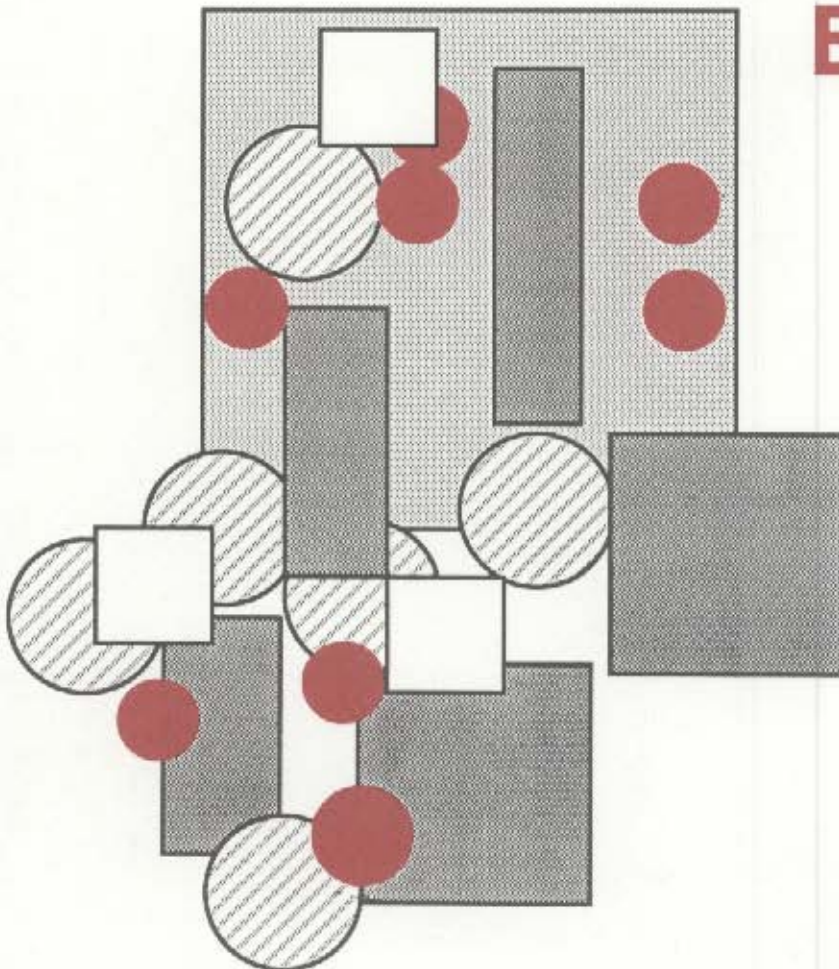


A Commission Report

The Organization Of Local Public Economies



ADVISORY
COMMISSION ON
INTERGOVERNMENTAL
RELATIONS

Washington, DC 20575
December 1987 A-109

Current Members of the Advisory Commission on Intergovernmental Relations

(December 1987)

Private Citizens

James S. Dwight, Jr., Arlington, Virginia
Daniel J. Elazar, Philadelphia, Pennsylvania
Robert B. Hawkins, Jr., *Chairman*, Sacramento, California

Members of the U.S. Senate

David Durenberger, Minnesota
William V. Roth, Jr., Delaware
James R. Sasser, Tennessee

Members of the U.S. House of Representatives

Sander Levin, Michigan
Jim Ross Lightfoot, Iowa
Ted Weiss, New York

Officers of the Executive Branch, U.S. Government

Gwendolyn S. King, Deputy Assistant to the President,
Director of Intergovernmental Affairs
Edwin Meese, III, Attorney General
Vacancy

Governors

John Ashcroft, Missouri
Ted Schwinden, Montana
John H. Sununu, *Vice Chairman*, New Hampshire
Vacancy

Mayors

Donald M. Fraser, Minneapolis, Minnesota
William H. Hudnut, III, Indianapolis, Indiana
Robert M. Isaac, Colorado Springs, Colorado
Henry W. Maier, Milwaukee, Wisconsin

Members of State Legislatures

John T. Bragg, Deputy Speaker, Tennessee House of Representatives
Ross O. Doyen, Kansas Senate
David E. Nething, Majority Leader, North Dakota Senate

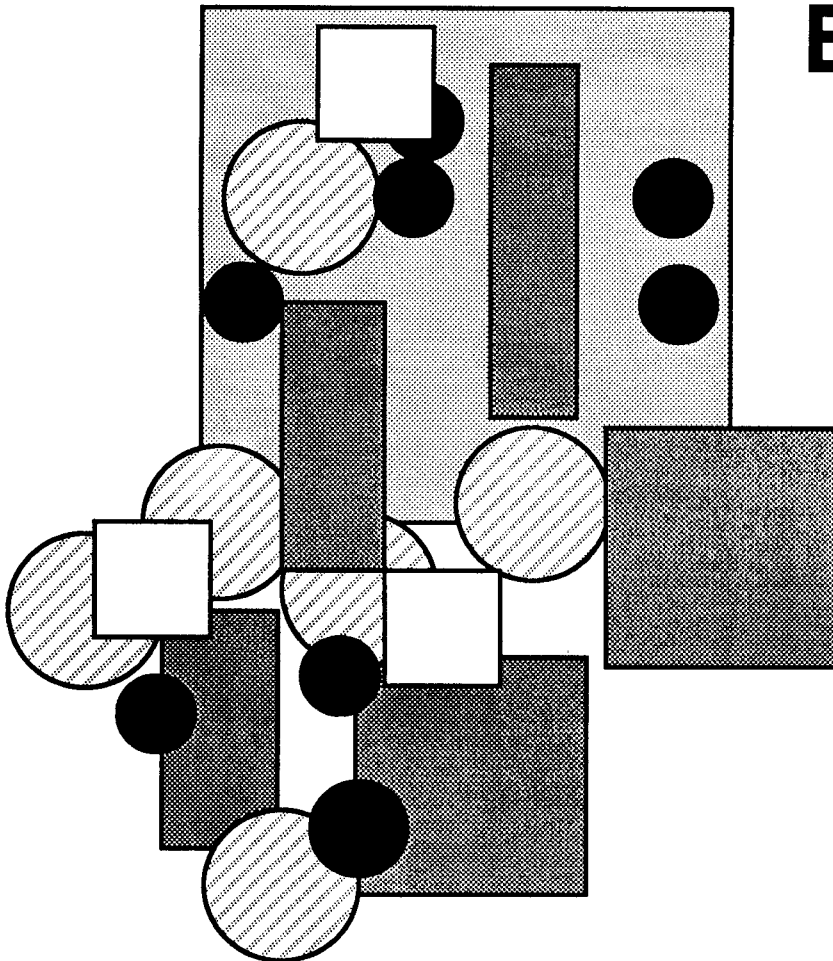
Elected County Officials

Gilbert Barrett, Dougherty County, Georgia, County Commission
Philip B. Elfstrom, Kane County, Illinois, County Commission
Sandra Smoley, Sacramento County, California, Board of Supervisors

A Commission Report

The Organization Of Local Public Economies

Reprinted September 1988



**ADVISORY
COMMISSION ON
INTERGOVERNMENTAL
RELATIONS**

Washington, DC 20575

December 1987 **A-109**

Preface

This policy report is the first publication from a Commission project on "Rethinking Local Self-Government in a Federal System." While drawing upon a venerable tradition of thought that extends back through several centuries of American experience with local government, this report challenges more recent conventions of thought that have been intellectually, if not politically, dominant throughout much of this century. For decades, many local government analysts have warned citizens of the dire consequences of "fragmentation" and of the "proliferation" of local governments, especially in metropolitan areas. Citizens have been urged to consolidate their local governments and to simplify local institutional arrangements. At times, state legislatures and even the federal government have been urged to effect metropolitan reorganization on behalf of local citizens. With this report, the ACIR comes down on the side of local citizen choice and of variety in local governmental organization.

Local citizens, when given a choice, have seldom elected to eliminate much of the basic infrastructure of "fragmented" local government in America. With this study, the commission explores a way of thinking and a range of evidence that lend intellectual and empirical support to these historic citizen choices. What is more, the report offers preliminary guidance on how to organize "local public economies" to achieve both greater responsiveness to citizen preferences and greater efficiency in fulfilling those preferences. The principal focus of these efforts should be on the "rules of the game," supplied in great part by state legislatures, that provide the institutional context within which local citizens and their elected officials can make choices that serve their best interests. This approach, when coupled with mechanisms to redistribute resources to less advantaged communities, can also

help to relieve inequities, but without destroying the political independence of the disadvantaged.

Accepting the challenge of local self-governance involves, among other things, changing the very language that we use to speak of local organization. New distinctions and new concepts can only be conveyed with new language. This report asks its readers to make a distinction, for example, between the "provision" and "production" of local goods and services. It asks readers to think of a cluster of local governments in a particular region not as a fragmented maze but as a "local public economy." It suggests that metropolitan "governance" may not depend upon the creation of a metropolitan "government." Likewise, pejorative terms, such as "fragmentation" and "proliferation" are replaced by such terms as "differentiation" and "growth." How we label things has a great deal to do with how we understand things.

This report concentrates, therefore, on certain fundamental elements of how we think about governing local communities in metropolitan areas. Such a rethinking is necessary if America is to remain true to its heritage of local self-governance and to the evident preferences of citizens in a democratic society. Such a rethinking is also necessary if America is to meet the challenges that will face governments, especially local governments, during the third century of American federalism. Changes in the federal aid system, the rise of international economic competition, continuing metropolitanization, and fluctuations in the performance of regional economies are among the challenges that require all of our institutions to be more productive and innovative. The vitality of local government lies in part in the kind of diversity and differentiation that can create opportunities for genuine self-government and for the productivity and resourcefulness that can come from citizens working together in manageable, human-scale communities.

Robert B. Hawkins, Jr.
Chairman

Acknowledgments

The principal author of this report is Ronald J. Oakerson, Senior Analyst and Assistant Director of Research. He was assisted in preparing portions of *Chapter Two* by Gary M. Anderson, former Analyst. General supervision and other assistance was supplied by John Kincaid, Director of Research. Lawrence A. Hunter, former Director of Research, supervised the initial phases of this project. Dolores T. Martin, former Senior Fellow, also contributed much to the project during its formative period. Further editorial assistance was supplied by Joan Casey, Information Officer, and secretarial assistance, by Anita McPhaul and Lori O'Bier.

In addition to ACIR staff, a number of persons directly contributed portions of the research base for this study while under contract to the Commission. Robert M. Stein of Rice University, together with Martin, contributed the statistical data analysis of municipal contracting that forms part of the discussion in *Chapter Two*. Roger B. Parks of Indiana University, assisted by Henry A. Bell, together with Oakerson, performed the field research and data analysis undertaken in the St. Louis, Missouri, metropolitan area

that helps to support the theoretical analysis outlined in this report. Parks also read and commented upon numerous drafts of the report.

Preliminary plans for this study were reviewed by an academic advisory board consisting of the following persons: Louis de Alessi, Robert Bish, James Ferris, Bryan Jones, Ted Kolderie, Daniel Mandelker, Stephen Mehay, Elinor Ostrom, Paul Peterson, Derek Shearer, and Richard Wagner. Initial findings were reviewed at a critics' session on April 22, 1987, attended by Henry A. Bell, Ann Cole, William Colman, Edwin Connerly, Lawrence A. Hunter, Susan Lauffer, Michael Libonati, Dolores Martin, James L. Martin, Richard Moore, Steve Moore, William Niskanen, Roger Parks, Doug Peterson, Mark Schneider, John Thomas, and Louise White. Others who read and commented upon drafts of the report include Wayne Anderson, Vincent Ostrom, Gordon Whitaker, and Joseph Zimmerman. The Commission and its staff are very grateful for the help of all those who contributed advice and criticism, but of course retain full responsibility for the content of the report.

John Shannon
Executive Director

John Kincaid
Director of Research

Contents

Summary	1
Chapter 1 Distinguishing Provision and Production	5
Introduction	5
Conceptual Foundations	6
Organizing the Provision Side	7
Problems of Provision	7
Transaction Costs as a Source of Constraint	9
Organizing the Production Side	10
Production Problems	10
Coordination Costs as a Source of Constraint	11
Linking Provision with Production	11
Governance	12
Equity	12
Implications: The Remainder of this Report	13
Chapter 2 Separating Provision and Production	15
Introduction	15
Contracting Options	15
Current Practice	15
Explaining the Variation in Contracting	16
<i>Table 2.1</i> Logit Estimates for Degree of Municipal Contracting, 1982–83	16
The Relevance of Provision Arrangements	17
“Pure Provision” Units	18
Are Small Local Governments Nonviable?	18
Do “Pure Provision” Units Have Advantages on the Provision Side?	19
Conclusion	20
Chapter 3 Differentiating Provision	23
Introduction	23
Types of Provision Units	23
Three Basic Types	23
The Special Case of Special Districts	24
Neighborhood Provision Units	25
Performance Differences among Provision Units	26
Size Effects	27

How Much Differentiation?	27
Conclusion	28
Chapter 4 Differentiating Production	31
Introduction	31
Separable Production Tasks	31
Service Components	31
Transaction Costs: A Constraint	31
The Extent of Differentiation and Its Effects	32
Conclusion	33
Chapter 5 Governance, Change, and Civil Community	35
Introduction	35
Choosing Boundaries	35
Dynamics and Adaptive Capabilities	37
The Limits of Interlocal Competition	38
Local Government Constitutions	39
Fragmentation and Civil Community in Metropolitan Areas	40
Conclusion	41
Chapter 6 Equity	43
Introduction	43
Alternative Approaches to Fiscal Disparity Problems	43
Empirical Research	44
Political Choice for Disadvantaged Communities	46
Chapter 7 Conclusion: Research and Policy Directions	49
Introduction	49
Metropolitan Governance	49
Implications for Further Research	50
Immediate Policy Implications	51
Conclusion	52
Commission Findings and Conclusions	53
1. Variety and Diversity	53
2. Distinguishing Provision and Production	53
3. Opportunities for Initiative, Creativity, and Cooperation	53
4. Citizen Choice and Consent in Local Governance	53
5. Equity and Community Choice	54
6. State Legislatures: Trustees of Local Governance	54
Commission Recommendations	55
<i>Recommendation 1: Nonviable Governments</i>	55
<i>Recommendation 2: Special-Purpose Governments</i>	55
<i>Recommendation 3: Boundary Review Agencies</i>	56

Summary

This inquiry into the organization of local governments argues that a multiplicity of general-purpose and special-purpose governments in a metropolitan area is not an obstacle to good government or to metropolitan governance. On the contrary, a diversity of local governments can promote key values of democratic government—namely, efficiency, equity, responsiveness, accountability, and self-governance. A multiplicity of differentiated governments does not necessarily imply fragmentation; instead, such governments, interactively linked through a variety of arrangements, can constitute a coherent local public economy.

A key concept in this study is an elementary distinction between the “provision” and “production” of public goods and services. Basically, provision refers to decisions that determine what public goods and services will be made available to a community. Production refers to how those goods and services will be made available. For example, a local community may choose to provide its residents with trash collection twice a week. This is a provision decision. The community can arrange for the production of this service in essentially one of two ways: (1) it can organize and operate its own trash collection service or (2) it can arrange for others, such as another government or a private firm, to deliver the service, possibly under contract.

Provision covers taxing and spending decisions about public goods and services, determinations of service standards, the monitoring of service delivery, and the securing of accountability to standards. Production refers to the process of combining resource inputs to make a product or render a service. The point is that provision and production are separable activities that can be linked in a variety of ways. Most often local governments both provide and produce goods and services. This is the traditional model of what a local government is supposed to do. Increasingly, however, local governments act as providers, but not as producers. That is, communities have elected to provide themselves with a range of services through local governments, which then contract out to private

firms or to other governments for the production and delivery of those services.

The implications of distinguishing provision and production, however, go well beyond contracting, and extend to other aspects of organizing a “local public economy.” The latter term refers to all of the local governments and private contractors that operate within a particular region or metropolitan area. A local public economy can be viewed as having a provision side and a production side, each of which can be organized in quite different ways. The criteria for organizing provision differ substantially from the criteria for organizing production. Provision criteria are concerned with how best to satisfy the preferences of citizens; production criteria have to do with the efficient management of human and material resources. In particular, the appropriate scale of organization for provision is frequently quite different from the appropriate scale for production.

Efficiency criteria suggest that provision ought to be differentiated among a number and variety of provision units in a local public economy. This is equivalent to suggesting, in traditional terms, that local government ought to be fragmented—arranged in multiple, sometimes overlapping, jurisdictions. The variety of provision units can include both general-purpose and special-purpose governments. Multiplicity serves a number of useful purposes: it increases the sensitivity of local government to diverse citizen preferences; it increases efficiency by matching the distribution of benefits more closely to the economic demand of communities; and it enables citizens to hold public officials accountable to a specific community of interest.

A growing number of local government units, particularly special districts, has usually been viewed as undesirable “proliferation” within metropolitan areas. It is possible to have too many local governments, but it is also possible to have too few. An “optimal” number of provision units has to be based on a tradeoff between the values associated with more responsive provision and the costs of organizing and operating each additional unit. These costs, best construed as

the transaction costs associated with making collective decisions, necessarily limit the number of provision units that can be organized in a metropolitan area. Yet neither growth nor reduction in the number of independent provision units in a local public economy can be taken as evidence that any particular number is suboptimal.

Much the same sort of calculation must occur on the production side. Different services, and even different components of the same service, exhibit different economies of scale—referring to the size of organization that can most efficiently produce or deliver a service. Production therefore ought to be differentiated, to some extent, among a variety of independent production units that can be linked together either through contracts or through cooperative agreements. The appropriate constraint on the number of production units consists of coordination costs—the transaction costs associated with making joint decisions on an interorganizational basis. Consolidating services or service components in a single production organization does not necessarily result in greater efficiency.

Given the need for some degree of differentiation (fragmentation) on both the provision side and the production side, a fundamental question remains: Who can make the necessary tradeoffs to arrive at an optimal number of independent units? On the provision side, the answer is citizens, acting as voters and consumers. Local citizens are in the best position to determine the appropriate number and mix of local governments required to serve their communities. This is because citizens can decide when their time and effort in organizing and monitoring an additional unit is worth as much as the extra responsiveness of provision to their preferences. On the production side, the answer in general is elected local officials, those who are directly involved in arranging for the production of specific services. Local officials are in the best position to know both the benefits of more efficient production and the costs of coordinating multiple producers, and are therefore able to make an appropriate tradeoff between efficiency in production and ease of coordination.

These conclusions have important implications for the governance of a local public economy. Governance is separable from provision and production, yet intimately related to both. The governance of a local public economy determines its basic structure—the number and variety of both provision and production units. If these structural decisions are best made by citizens, on the provision side, then governance depends upon rules that allow for citizen choice. State enabling rules that allow citizens to create (and dissolve) local governments through procedures such as municipal incorporation (and similar procedures of special district formation), can be viewed as essential aspects of local governance. Citizen participation in

making boundary adjustments through annexation and consolidation (as well as detachment) is also an integral part of this process. Similarly, on the production side, governance depends upon rules that allow local officials to work out mutually agreeable arrangements, both with one another and with the private sector, in searching for the most efficient units of production subject to coordination limits.

The rules that establish a framework of governance for a local public economy, though derived from numerous legal sources (mostly state laws and constitutions), must be viewed as a single configuration—a local government constitution. The ability of local citizens and officials to develop and maintain a local public economy in an optimal manner depends upon an appropriate configuration of rules. The wrong configuration can result in serious distortions of local structure—too many or too few units of this or that type, ill suited to the problems faced by local communities. If citizens in some areas have created too many special districts, for example, it may be because enabling rules have too tightly constrained general-purpose local governments. The remedy in this event is not to limit citizen choices, but to expand the choices that citizens and officials can make within their existing provision units. In many areas, however, the limiting factor may be too little authority in the hands of citizens to create additional units, resulting in an insufficient number and variety of local jurisdictions able to respond to citizen preferences.

Local public economies can also generate inequities among citizens organized in different communities. Traditionally, this problem has been addressed by calling for the organization of larger, more heterogeneous communities able to pool revenues on a larger scale and distribute services according to need. This approach is often inconsistent, however, with citizen choice and empowerment. An alternative approach, already followed to some extent (especially with respect to schools), is to use overlapping jurisdictions, such as state and federal governments, to redistribute resources when fiscal disparities among local jurisdictions exceed the limits of equity. This approach allows less advantaged communities to retain their political independence and, importantly, to arrange for the production of services themselves.

In general, a well-organized local public economy will consist of an array of provision units—general-purpose governments such as municipalities, townships, and counties, alongside special-purpose governments, such as school districts, fire protection districts, and park districts, to name only a few possibilities. Such provision units will arrange for production in a variety of ways—in-house production, private or intergovernmental contracting, or joint production units able to deliver services to two or more communities. Viewed as a whole, such an array is necessarily

complex. What enables this complex array to work, however, is a network of relationships among citizens and local officials, a network that is based upon their ability to make incremental choices that optimize the arrangements for both provision and production of services.

For decades, local government analysts have searched for a seemingly ideal pattern of local organization—consolidated metropolitan government, two-tier systems, or any arrangement that would reduce the governmental complexity ordinarily found in metropolitan areas. This report suggests that there is no ideal pattern suitable for all local communities and re-

gions. Rather, local citizens and officials need to be empowered to create optimal patterns of order for themselves. Perfection will continue to be elusive. Different communities will make different tradeoffs, and every community can learn from these natural experiments. Metropolitan governance need not depend upon the creation of a metropolitan government. Citizens can govern local public economies, by creating and maintaining multiple local governments within a framework of rules. It is this process of governance, and the structure that maintains it, that should be a major focus of research in state-local and interlocal relations.

Distinguishing Provision and Production

INTRODUCTION

The last two decades have seen significant changes in ways of thinking about the patterns of organization that affect local government. New ideas and concepts of local organization and metropolitan governance have generated new research, with important, often counter-intuitive, conclusions. Patterns once despised are now seen to have virtues; patterns once welcomed are now viewed with skepticism. Yet the traditional American commitment to local self-government appears to be as strong as ever. A new consensus may be emerging around a simple but powerful idea—that multiple local governments, coexisting in a given region, constitute a “local public economy” consisting of a “provision side” and a “production side,” each of which can be organized in quite different ways. The purpose of this report is to develop the implications of this new way of thinking and, in the process, examine some evidence relevant to an assessment of its potential for improving the conduct of local government in the United States.

Distinguishing provision (taxing and spending) from the production and delivery of local goods and services has far-reaching implications for the organization and governance of a local public economy. The implications explored in this report include (1) a greater reliance on both private and intergovernmental contracting to produce services, (2) a greater number and variety of local government jurisdictions to make provision for local services, (3) a more highly differentiated set of production units, whether public or private, each of which produces particular components of a public good or service, and (4) the separation of issues of “governance” from both provision and production. Each of these implications raises issues of *intergovernmental organization* that acquire a fundamental importance in the way local public economies work.

A public economy is not a private economy, and local is not national. Many of the concepts and theories of economics apply to a local public economy, but not all, at least not in a completely straightforward way. To link the concepts “public” and “economy” in a single phrase is to integrate political with economic reasoning. The basic units of organization—though not the only relevant units—in a local public economy are local governments. In this context, consumers are first of all “citizens.” Implied in the term “economy” is the presence of important patterns of exchange as an element of local organization. Yet exchange is embedded in community, a context that becomes even more important in the case of exchanges among governments and other economic units in specific localities. The governance of a local public economy, like a private economy, is a political task carried out by political means.

At the same time, local public economies are linked to private economies in a variety of ways, some of which distinguish the local scene from a national political economy. A number of observers have emphasized that local communities are much more exposed to the movement of private firms and individuals across borders, and thus more constrained than a national community by the effects of economic competition, whether distant or close at hand.¹ Local governments also pursue policies that both facilitate and constrain local economic development, shaping it to the public demands of local communities. Much of the revenue of local government is raised from commercial and industrial sources that endeavor to pass their costs on to consumers, many of whom are not local residents. Moreover, the way that local government is organized, together with decisions about what local goods and services are to be provided and how to arrange for their production, may encourage or inhibit the growth of a “public-private sector” con-

cerned with the production and delivery of services to local governments and their citizens.

Alexis de Tocqueville, whose *Democracy in America* is widely regarded as the "classic" expression of the values of American local governance, focused much of his attention upon two concepts: citizenship and productivity.² One of his basic arguments was that citizen empowerment, in the context of American local institutions, led to community productivity—better roads, better schools, better law enforcement.³ How to maintain, and in some cases recreate, this relationship between citizenship and productivity is perhaps the basic challenge of American local democracy today. Citizens in a public economy are analogous to consumers in a private economy. Economic betterment is generally associated with an expanding set of consumer choices. In a private economy, choice is power. Similarly, a developing local public economy may be associated with an expanding set of citizen choices. And again, choice is power.

Four broad avenues of research have contributed to this ACIR study: (1) an assessment of the major theoretical work integrating political and economic approaches in the study of local governance;⁴ (2) a review of recent empirical research linking the structure and performance of local public economies in the United States; (3) a case study of the St. Louis metropolitan area as an example of a highly differentiated local public economy;⁵ and (4) an aggregate data analysis of municipal contracting and joint service agreements utilizing a representative national sample.⁶

This chapter continues by examining the conceptual foundations for a study of local public economies, distinguishing the provision side from the production side. Subsequent chapters explore the implications of this distinction for basic patterns of local and inter-local organization in view of recent empirical research.

CONCEPTUAL FOUNDATIONS

Richard A. Musgrave advanced the basic distinction between provision and production as early as 1959:

Provision for public wants. . . does not require public production management, just as public production management does not require provision for public wants. Quite different criteria apply in determining the proper scope of each.⁷

Vincent Ostrom, Charles M. Tiebout, and Robert Warren, writing in 1961, were among the first to argue for the importance of this distinction in the organization of local public economies in metropolitan areas:

The separation of the *provision* of public goods and services from their *production* opens up the greatest possibility of redefining economic functions in a public service econ-

omy. Public control can be maintained in relation to performance criteria in the provision of services, while allowing an increasing amount of competition to develop among the agencies that produce them.⁸

Despite long-standing research as well as many expressions of public preference, the distinction has been slow to affect the ways in which scholars and practitioners think about the functions of local governments. Instead, thinking was long dominated by the idea that "strong" local government requires relatively large general-purpose governments able to *provide and produce* a wide range of goods and services. Production criteria have therefore been predominant in evaluations of local government organization. Discussions of local government "capacity," for example, have usually meant capacity to produce. From this perspective, a local government that contracts out for *all* of its service production is inactive or nonperforming.

The concept of a local public economy, by distinguishing a provision side from a production side, anticipates that each side will engage in distinctive activities.⁹ Provision, in general, refers to collective choices that determine:

- (1) what goods and services to provide (and what are to remain private);
- (2) what private activities to regulate, and the type and degree of regulation to use;
- (3) the amount of revenue to raise, and how to raise it (whether by various forms of taxation or by user pricing);
- (4) the quantities and quality standards of goods and services to be provided; and
- (5) how to arrange for the production of goods and services.

The basic choice on the provision side is whether to make any sort of provision at all. If no public provision is made (by any unit of government), provision of the good or service remains a private decision. Many local governments, for example, provide elaborate programs of recreation, while other communities choose to leave the provision of recreation in the private sector. If public provision is to be undertaken, how to provide is another basic choice. The two major alternatives are (1) taxing and spending to provide a public good or service or (2) regulating private activity in order to shape private decisions to a public purpose. Provision activities thus include, on the one hand, setting both tax rates and user charges and choosing how to spend public money, and, on the other hand, enacting and enforcing laws or rules that constrain private behavior according to public criteria. Finally, if goods and services are to be provided, there follows a basic choice of how to arrange for production—whether by organizing a production unit "in-house" or by selecting and hiring outside producers

(public or private). Having arranged for production, provision requires monitoring the quality and quantity of goods and services produced, representing the interests of citizen-consumers to producers, and holding producers accountable for their performance.

Provision thus consists of a wide range of decisions and activities. The focus in this report is on the public provision of goods and services and the activities associated with this type of provisioning. This is not to depreciate the importance of other provision choices, including the choice of nonprovision or privatization. Rather it is to examine one important dimension of local provision activity in depth and to trace the implications of various organizational choices for the structure of local public economies.

Production, as distinguished from provision, refers to the more technical process of transforming inputs into outputs—making a product or, in many cases, rendering a service. Part of the task of provision is to establish a link with producers, but not necessarily to organize production. In this sense, a provision unit is like a household. One of the basic household decisions, in economic terms, is to decide whether to (1) produce for oneself or (2) enter the marketplace and purchase a particular good or service. All households produce some goods and services for themselves, whether they be meals, laundry, or a self-built home. The same is true for most public provision units. Modern households may tend to be formed more on the basis of provision criteria, such as compatible preferences, than on production criteria, such as skill in housekeeping, meal preparation, and general home repair. (Households that lack access to a well-developed market economy may be quite different.) Even when service production is delegated to agents (public or private), it is frequently better viewed as “coproduction,” a process whereby both a specialized producer (e.g., a teacher) interacts with a citizen-consumer (e.g., a student) to “produce” a good (e.g., education).¹⁰ In this sense, the members of a provision unit, like the members of a household, may still be intimately involved in a production process that relies upon “outside” producers.

A local public economy consists of some array of provision units and production units. The provision units are communities organized to make collective decisions. Governmental arrangements are one way to organize provision units, but there are alternatives, chiefly some sort of homeowners association, such as a condominium or organized subdivision. Municipalities, townships, counties, school districts, and other special districts may all function as provision units on behalf of various communities, some of which overlap others. Production units are either private firms or public agencies and departments organized by a provision unit (or jointly by two or more provision units). The types of relationships formed among this array of

provision and production units are the subject of *Chapters Two, Three, and Four*.

The distinction between provision and production lays the conceptual foundation for a new understanding of the organization of local governments. Royce Hanson, Associate Dean of the Hubert H. Humphrey Institute of Public Affairs at the University of Minnesota, recently wrote:

There is a growing consensus that government has a responsibility to *provide* services and facilities through policies it makes, but services and facilities may be *produced* by either sector, based on values of cost effectiveness and equity.”¹¹

The work of local government per se increasingly is viewed primarily as provisioning, rather than producing. Different criteria apply in the choice of an organizational unit to *provide* a service than in the choice of an organizational unit to *produce* a service. It follows that patterns of organization on the provision side of a local public economy may differ substantially from those on the production side.

ORGANIZING THE PROVISION SIDE

The organization of the provision side of a local public economy encounters problems that fall into three main classes: (1) preference revelation, (2) fiscal equivalence, and (3) accountability. Each class of problems suggests criteria for assessing the *efficiency* of the provision side. The criteria are quite different from those used to assess production-side efficiency.

Problems of Provision

Preference Revelation

The problem of individual preference revelation derives from the incentives of individuals to conceal their true preferences for public goods and services if provision is organized on a strictly voluntary basis. The institutional requirement is for some process of *collective choice* from which an individual cannot simply opt out. (Individuals can, of course, move out of a local jurisdiction, but this is different from opting out of provision while continuing to live in the jurisdiction.) The collective-choice requirement can be satisfied in a variety of ways—municipalities, counties, townships, special districts, and subdivisions all provide various institutional capabilities for citizens to act collectively.

Given some form of collective choice, the problem becomes one of organizing a set of provision units that can minimize the distortion of individual preferences and facilitate the flow of relevant information to decision makers, as individual choices are aggregated to arrive at a single set of choices for relevant communities. The potential for distortion of both preferences and information is considerable in a process of collec-

tive choice. As John Stuart Mill stated the problem in his essay *On Liberty*, what we call self-government is “not the government of each by himself, but the government of each by all the rest.” Majority rule can be used to aggregate preferences no matter what those preferences are, but the outcome of this process may be arbitrary, unfair, or insensitive to variations among individuals and groups.¹² Avoiding these undesirable consequences requires structures and processes of collective choice that can take account of diverse community preferences and mobilize specific time-and-place information about local community conditions.

No particular scale of organization is implied with respect to any single provision unit as long as those individuals immediately affected by provision (or non-provision) are included within its boundaries. The necessary extent of those boundaries varies with the type of good or service in question. The “best” boundaries are those which, minimally, include the affected group of people and, at the same time, are not so large as to include groups of individuals with widely different preferences. In sum, a provision unit should, as closely as possible, define a *community of interest* among a group of people who share some contiguous part of the local geography.

Different goods and services, however, implicate somewhat different communities. The geographic meaning of “local” varies from one problem to another. “Local problems” come in various sizes and shapes. One approach is to establish boundaries that embrace the widest set of problems. Traditional prescriptions for “metropolitan government” follow this approach. The difficulty is that a single large provision unit is then created that must act on behalf of an extremely heterogeneous collection of smaller communities with distinct interests. The alternative approach is to rely upon *overlapping* provision units that allow different communities of interest to be organized in relation to different public problems.

In sum, the criterion of preference revelation is that individuals affected by common problems, who want perhaps to provide themselves with collective goods and services, must be able to act collectively. Moreover, the difficulties of collective choice suggest that the power to act collectively, rather than being concentrated in a single jurisdiction, ought to be distributed among diverse communities of interest, some of which are more inclusive than others. The basic task of local governance, then, is to distribute, in an optimal way, the powers of collective action among various, overlapping communities of interest, allowing a variety of provision units to be created.

Fiscal Equivalence

Efficiency on the provision side of a local public economy varies with the degree of “fiscal equivalence”¹³ that is attained by establishing boundaries. This criterion means simply that individuals (house-

holds or firms) and groups (neighborhoods or communities) “get what they pay for and pay for what they get.” Otherwise, individuals are led to distort their own expression of “demand” for local public goods, either by inflating their demands if others are required to pay or by depressing their demands if they have to pay more than their share. A lack of fiscal equivalence undermines the community of interest within a provision unit and generates incentives for individuals to try to improve their own lot at the expense of others rather than by joining with others to improve the welfare of the community.

For some goods and services—those subject to exclusion at the point of consumption—user charges can be utilized to generate fiscal equivalence among individuals. A public swimming pool or golf course may therefore charge admission. Benefit-based taxes can also be used to the same end. Property taxes are generally understood as benefit based.¹⁴ Unlike a user charge, however, a tax is not contingent upon service delivery. The payment of taxes cannot establish fiscal equivalence apart from the delivery of services. A jurisdiction that contains numerous and varied communities of interest, yet raises revenue on a jurisdiction-wide basis, does not automatically distribute services in a way that is proportionate to the revenue generated from various communities. In general, the greater the disparity between jurisdictional boundaries and relevant communities of interest, the more problematical fiscal equivalence becomes.

Accountability

Provision units also must deal with the potential for distortion in principal-agent relationships between citizens and officials. All communities stand in need of agents who represent the interests of members. The problem is that agents can develop their own interests distinct from the interests of those whom they represent. This is especially the case when the “principal” is a community of people who find common expression in great part through their agents, including both elected officials and civil servants. Provision units need to be organized in such a way that individual citizens are able to exercise a significant measure of “voice” so that those who function as agents can be held accountable in the conduct of community affairs.

The ease with which citizens can hold officials accountable is related both to the size and preference heterogeneity of a jurisdiction and to the number of functions undertaken by a jurisdiction. Officials in a single small jurisdiction related to a single coherent community of interest can be held accountable by citizens with relative ease. Officials in a large jurisdiction related to multiple communities of interest can be held accountable by citizens only with much greater difficulty. Single-function units exhibit lower within-unit costs of accountability than do multi-function units, although multiplying the number of single-function units

increases the costs of accountability overall. In general, the costs of accountability to citizens are minimized by provision arrangements that closely match provision units to discrete communities of interest.

Transaction Costs as a Source of Constraint

The organization of the provision side of a local public economy is thus concerned with boundary decisions used to define a community of interest, a choice of taxing instruments and the way those instruments are used in terms of fiscal equivalence, and the availability of opportunities for citizens to hold officials accountable to a community of interest. All three criteria—preference revelation, fiscal equivalence, and accountability—point to the desirability of allowing citizens to create multiple provision units so that each unit is as closely matched as possible to an underlying community of interest among citizens. Interests shared on a small scale should be organized by relatively small units; interests shared on a larger scale should be organized with reference to larger, overlapping jurisdictions.

None of the three criteria, however, suggest a general source of *constraint* on the creation of provision units.¹⁵ The limiting factor in the creation of multiple provision units consists of the *transaction costs* involved in organizing and operating such units.¹⁶ The conduct of both private and public business entails transaction costs. In the private sector, transaction costs include the time and effort devoted to such activities as searching for buyers or sellers, bargaining over price, and enforcing contracts. In the public sector, transaction costs consist of the time and effort devoted to making collective decisions, including the cost of elections, communications and meetings, and reporting to secure accountability. The costs of citizen participation must be counted in addition to the costs of official action.

To some extent, the values associated with preference revelation, fiscal equivalence, and accountability have to be traded off against the transaction costs associated with creating and then maintaining a separate provision unit. Organizing costs are present mainly at the start of an enterprise, but the marginal costs of operating a separate unit—the additional operating costs of adding a provision unit—are continuing. The *marginal* costs of operation, however, are not equal to the *total* costs of operation. Presumably, adding a unit relieves other units of part of their work (or precludes work from being added). Much of the controversy surrounding “duplication” among multiple provision units could be resolved, in principle, if the marginal cost of each unit’s operations, including the costs of citizen participation, could be estimated. This figure would indicate the price of achieving more precise preference revelation, greater fiscal equivalence, and

greater (or less costly) accountability, by organizing an additional provision unit.¹⁷

Choosing an appropriate number of provision units presents an optimizing problem for citizens. Consider, for example, the organization of school districts as separate provision units—separate, that is, from municipal, county, or township governments organized to provide a range of community services. Electing a board of education separate from a city council entails some obvious additional costs—elections, personnel, meetings—mostly transaction costs. If the same community is organized by both a school district and, perhaps, a municipality, then fiscal equivalence is not at issue. The difference must lie in the areas of preference expression and accountability. By electing a school board separately, preferences about schools do not have to be expressed simultaneously with preferences about police protection or garbage collection as voters make electoral choices. Accountability for the performance of schools can be separated from accountability for the performance of other functions. Separate accountability for schools is presumably worth the marginal cost of a separate provision unit. The alternative way that this issue is usually expressed is to keep schools “out of politics.” What that expression means is to keep educational issues from being confounded with other issues in a community.

On the other hand, a community may find that police protection and garbage collection do not warrant separate provision units within the same community. Although these two functions are quite distinct, as long as the same community is involved, it may be possible to economize on transactions costs by combining provision for both services in a single unit.

The optimal number of provision units may also change over time, as, for example, when citizens decide to create a new municipality within an existing county jurisdiction. The question in this case is whether a community finds that its interests and preferences have become sufficiently different from others in the county to justify the creation of a new provision unit. The initial costs of creating a new municipal jurisdiction are substantial, usually including at least the costs of petitioning and voting. The marginal operating costs are also apt to be positive, adding a new mayor and council to the existing county provisioning apparatus. At the same time, however, the work load of a county council and other elected county officials would decrease somewhat. The potential benefits lie in the areas of preference revelation, fiscal equivalence, and accountability. More precise preference revelation may enable citizens in the new municipality to obtain a substantially different level and mix of services. Greater fiscal equivalence may lead to a more efficient allocation of resources among competing services and communities, considered from a countywide perspective. Accountability secured with

respect to a more compact community may enable citizens to obtain a closer match between their preferences and local government performance. Citizen-participation costs entailed in obtaining accountability may actually decrease, owing to the greater accessibility of officials. The question that local citizens have to decide is whether these advantages outweigh the total organizing costs and marginal operating costs of creating and maintaining a new provision unit.

The consolidation of existing provision units poses the same question in reverse. Are the marginal gains from eliminating a provision unit greater than the costs calculated in terms of preference revelation, fiscal equivalence, and accountability? Communities must also sometimes consider a potential redistribution factor. Less wealthy communities may gain from associating with more wealthy communities in a single provision unit, a calculation that must be discounted to some extent by political risk. This factor then has to be added to the marginal gain calculated in terms of operating costs.

Because organizing and operating provision units are costly activities, one might anticipate diminishing returns as preference revelation, fiscal equivalence, and accountability are improved. Perfection on any of these criteria is out of reach. Citizens (or their representatives) must decide whether increased performance with respect to the major provision criteria is worth the added cost of organizing and operating one more provision unit. At some point as the number of provision units per citizen increases, it becomes more efficient to assign new functions to existing units rather than create new units to organize provision. Part of the logic of multi-purpose units is to economize on transaction costs. This does not mean, however, that multi-purpose units are always preferred to special-purpose units. The answer varies from case to case.

The efficiency of the provision side of a local public economy depends upon a complex set of calculations and tradeoffs—some of which are inherently subjective—subject to change over time. Apart from the choices that citizens make, there is no way of determining provision-side efficiency. Traditionally, a great deal of concern has been focused on the “proliferation” of units of local government. On this basis, rules have been proposed, and to some extent enacted, that tend to limit the ability of citizens to create new units.¹⁸ Even in the absence of such rules, however, it should not be assumed that units proliferate without constraint. There is no reason to presume that citizens are unable to take into account the costs as well as the benefits of creating an additional unit, even if neither side of the tradeoff can be quantified with precision. Some communities resist incorporation. Others choose to merge with neighboring communities.¹⁹ Growth in the number of units in a local public economy does not imply that growth is unconstrained.

ORGANIZING THE PRODUCTION SIDE

Organization on the production side, in contrast to provision, is based upon considerations having to do with the technical transformation of resource “inputs” into product “outputs.” Unfortunately, no one has been able to demonstrate a “recipe”²⁰ for producing good policing or good education. On the other hand, somewhat more is known about producing good streets and good refuse collection. Almost all local public goods and services, however, depend upon the availability of specific time-and-place information, such as neighborhood conditions, to support effective production choices.²¹ An emphasis has to be placed upon the scale and organization of a production process that allows individual producers to make locally informed judgments. This is a much different production problem than that which is involved in a typical “assembly line.”

Production Problems

Coproduction

Traditionally, production-side considerations have placed a heavy emphasis on the importance of “management,” namely, the coordination of human and material resources. Human resources are generally viewed in the context of an employer-employee relationship. Many public services also depend, however, on the productive efforts of citizen-consumers as an integral part of the production process. The participation of citizen-consumers in production, as distinguished from provision, is called “coproduction.”²² Unlike employees, citizens are not as easily subject to the direction of a professional manager. While it is well known that the productivity of local public agencies, such as schools and police departments, frequently depends upon the cooperation of citizens, it is not well understood how to incorporate citizens into a production process. Yet citizen-consumers are often a crucial source of needed time-and-place information. The need for coproduction must be joined with traditional concerns for professional management when organizing the production side of a local public economy. The willingness of citizens to contribute to the production side, as coproducers, is often a crucial factor in the ability of a production unit to address relevant community needs and preferences.

Economies of Scale

An important distinction exists between local public goods that tend to be capital intensive and local public services that tend to be labor intensive. Capital intensive goods, such as water and sewer systems, are more likely to be characterized by “economies of scale,” namely, a decrease in the average unit cost of production as the scale of production increases.²³ Labor intensive services, such as police protection and

education, are more likely to exhaust potential economies of scale quickly, in part because of the greater dependence of human services on specific time-and-place information. Different goods and services are apt to be characterized by different economies of scale—some may be very small, others quite large. This logic also extends to different components of the same service. Much different economies may also be involved in increasing the level of production per capita for a given population as opposed to extending the same level of production per capita to a larger population.

When economies of scale happen to be closely matched to the scale of organization of a provision unit, in-house production is feasible. A small municipality, for example, can efficiently produce police patrol services; important economies of scale can be captured up to about 4-5 officers, after which further expansion entails constant (and eventually decreasing) returns to scale.²⁴ Similarly, a large municipality can efficiently undertake in-house production of the more capital-intensive components of services, such as a police crime lab. Frequently, however, provision criteria may lead to the establishment of provision units, both small and large, that are not well matched to economies of scale for particular services and service components. This is the circumstance that most often leads to the *separate organization* of production.

Coordination Costs as a Source of Constraint

The production and delivery of goods and services can be broken down into a large number of components, distinguishing direct service components delivered to citizens from various support service components delivered to direct service producers. Each component may be associated with a different economy of scale (although services typically exhibit constant returns over a broad range). Yet the different components of service production require coordination to varying degrees. Coordination is costly, mainly in terms of the time and effort devoted to transactions. The transaction costs associated with coordination can be expected to constrain the proliferation of separate production units much as transaction costs constrain the proliferation of separate provision units. An optimal number of production units depends upon a tradeoff between scale economies, on the one hand, and coordination economies, on the other. Coordination costs also arise, however, *within* production units. The relevant constraint therefore consists of the *marginal* costs of coordination from adding an additional production unit rather than the *total* costs of coordination for a given production unit.²⁵

LINKING PROVISION WITH PRODUCTION

Distinguishing the provision side from the production side of a local public economy opens up a range of possibilities for *linking* a provision unit to production units. The main options are as follows:

- *In-house production.* A provision unit organizes its own production unit. This is the traditional model of local organization. Municipalities organize municipal departments for police, fire, public works, and so forth.
- *Coordinated production.* Two or more production units (organized by their respective provision units) coordinate their production activities, in whole or in part. Two police departments, for example, may cooperate to investigate a string of burglaries affecting both jurisdictions.
- *Joint production.* Two or more provision units jointly organize a single production unit. A number of small municipalities, for example, may jointly finance and administer an emergency dispatch center for their police and fire departments.
- *Intergovernmental contracting.* A provision unit contracts for production with another provision unit, which assumes responsibility for organizing production.
- *Private contracting.* A provision unit contracts with a private vendor, who is responsible for organizing a production unit.
- *Franchising.* A provision unit sets production standards and selects a private producer, but allows individual citizen-consumers to choose whether to purchase the service. For example, a municipality, instead of providing trash collection, may grant a franchise to a private collector who is then authorized exclusively to sell the service to residents in that jurisdiction.
- *Vouchering.* A provision unit sets production standards and decides on the level of provision (through its taxing and spending powers), but allows individuals (or groups) to engage different producers, public or private, at their discretion. Vouchers, distributed to individuals by government, are now most frequently used in housing and social services, and combine public finance of services with individual choice of service suppliers.

Public administration theory has long assumed that the best sort of local government was a “full-service” government—somewhat analogous to a vertically

integrated business firm—performing both provision and production functions for a full range of local public goods and services. On the provision side, it was assumed that a relatively large general-purpose government is best able to balance the demands of competing concerns and interests within a single budget document. On the production side, it was assumed that coordination and economies of scale are best achieved when a single local government manager is able to assign functions among subordinate units. Both assumptions overlook the advantages of greater variety in organization.

The potential variety in organizing both the provision side and the production side, and in relating provision to production, is much greater than a traditional view of local government would suggest. More and more local governments are now beginning to look upon their primary role as one of “provisioning” as distinct from producing, choosing to produce only when in-house production, through a government bureau or department, is the best alternative available. The ability to separate provision from production, when advantageous, also potentially frees the provision side of a local public economy to be organized on the basis of provision criteria, subject only to the constraint inherent in the costs of organizing and operating additional provision units.

GOVERNANCE

The governance of a local public economy characterized by multiple provision and production units is not likely to be vested in any single provision unit. Governance is neither provision nor production. Instead, it has to do with a choice of *rules*, establishing an institutional framework within which patterns of provision and production emerge from the choices of local citizens and officials. The governance process includes the *resolution of conflict* among participants, as well as the maintenance of agreeable and equitable arrangements. The work of governance is separable from both provision and production.

The basic task of governing a local public economy is to put together, and periodically adjust, a configuration of rules that enables citizens and their local officials to seek out and create mutually beneficial provision and production arrangements, among both individuals and communities. The relevant rules pertain to: (1) organization on the provision side (which depends upon collective choice capabilities, tax instruments, the use of “police” powers in regulation, elections and referenda, and the public accountability of officials); (2) the organization of public production; and (3) relationships between provision units and production units (especially contracting and joint production).

When conflict occurs (as for example over municipal boundaries or tax incidence), governance ar-

rangements must exist to apply general rules to specific cases and to encourage participants to reach settlements. If settlements cannot be reached on the basis of existing rules, new rules may be needed. Fiscal disparities among provision units are a potential source of conflict in most highly differentiated local public economies. Adjustments in the fiscal rules governing revenue capabilities—in particular the availability, and possible sharing, of various tax bases—are often responses to fiscal conflict. The highly differentiated local public economies usually found in metropolitan areas do require a form of metropolitan governance; but it is a form of governance that depends upon citizens to make basic structural choices within a framework of rules.

The “organization” of a local public economy can therefore be understood at two levels. One is the *governance level*—the maintenance of rules—including rule-generating, rule-applying, and rule-enforcing facilities—that establish a framework within which provision and production arrangements emerge. The other is the *provision/production level*—the maintenance of specific provision and production arrangements with respect to specific goods and services in specific communities of interest. Both levels of organization can be distinguished from performance—the patterns of goods and services actually provided, produced, delivered, and consumed.

Some of the most controversial empirical questions about local government—especially in metropolitan areas—are those that concern linkages between levels of organization and between organization and performance. Does jurisdictional fragmentation, allowed by permissive incorporation rules, result in uncoordinated service delivery? Do overlapping jurisdictions, created by allowing citizens to form special districts in addition to municipalities, engender duplication of effort and a misallocation of resources? Or does the greater variety of organization at the provision/production level, generated by a governance level that allows citizens to create multiple provision units, lead to greater efficiency and responsiveness with respect to a broad range of local services?

EQUITY

The governance of a local public economy also includes reference to efforts to maintain some level of distributional equity among citizens. Two basic approaches have been followed in dealing with this issue. One approach, consistent with the traditional view of organizing local government, is to enlarge the boundaries of provision units so as to include a heterogeneous community with respect to income. Pooling revenues over this heterogeneous community while distributing services according to need would generate equity. Public resources could be closely matched to public needs. An alternative approach is to rely upon overlapping provision units to pool revenue on a larger

scale and redistribute resources to provision units considered to be disadvantaged according to a set of criteria. Both approaches have been followed to some degree in the United States. School district consolidation is perhaps the major example of the former approach, but local district consolidation has been accompanied by further efforts to redistribute resources on a statewide basis in many states, utilizing complex formulas to divide state funds among local school districts. An intermediate approach has been pioneered in the Minneapolis-St. Paul area, where revenues from commercial and industrial property are partially shared on a metropolitanwide basis.

Any effort to redistribute resources represents a departure from the criterion of fiscal equivalence. Arranging the boundaries of provision units to encompass a deliberately heterogeneous community is inconsistent with fiscal equivalence. "Getting what you pay for and paying for what you get" is, however, also a principle of equity, albeit not redistributive equity. The alternative approach, relying upon redistribution by overlapping jurisdictions, attempts to combine the principle of fiscal equivalence, as a first-order criterion of organization, with redistributive equity as a second-order criterion. Efforts to redistribute then operate only when the first-order criteria generate patterns of performance that lie outside acceptable limits of resource distribution. A critical question, one that can only be answered by empirical research, is whether the use of the first-order criteria to generate a highly differentiated pattern of organization creates serious problems of equity, which then have to be addressed by overlapping units. In traditional language, the question is this: Does fragmentation cause inequity?

IMPLICATIONS: THE REMAINDER OF THIS REPORT

Distinguishing provision and production has far-

reaching implications for the organization and governance of a local public economy. All of the important implications cannot, however, be explored here. Neither regulatory activity on the provision side, nor franchising and vouchering as important options for linking provision to production, are further developed in this report. Instead, the remainder of this report is focused upon implications for service provision and production, examined from both efficiency and equity standpoints.²⁶

Some of the implications to be drawn from the provision/production distinction, discussed in ensuing chapters, are as follows:

- Both private and intergovernmental contracting are viable alternatives to in-house production;
- Very small local governments are potentially viable as "pure provision" units;
- An optimal mix of provision units in a local public economy is likely to range in size, from large to small, with some units overlapping others;
- An optimal set of production arrangements is apt to mix in-house production with contracting out for different components of the same service;
- An optimal pattern of organization can emerge among provision and production units without being planned by any single set of decision makers;
- Metropolitan or regional "governance" is possible without metropolitan or regional "government";
- Equity may be better served by redistributing resources through overlapping jurisdictions than by adjusting local boundaries.

NOTES

¹See, for example, Paul E. Peterson, *City Limits* (Chicago: University of Chicago Press, 1981).

²For a full discussion see Ronald J. Oakeron, "The Meaning and Purpose of Local Government," ACIR Working Paper, 1986. Tocqueville's argument is developed in *Democracy in America*, Volume I, Chapter Four.

³Randy Hamilton makes a similar point in "Self-Government through Citizen Legislators: The Bedrock of Liberty" in ACIR, *Is Constitutional Reform Necessary to Reinvigorate Federalism? A Roundtable Discussion*, M-154 (November 1987).

⁴Earlier summaries of this theoretical literature include Robert L. Bish, *The Public Economy of Metropolitan Areas* (Chicago: Markham Publishing Company, 1971) and Robert L. Bish and Vincent Ostrom, *Understanding Urban Government: Metropolitan Reform Reconsidered* (Washington, DC: American Enterprise Institute, 1973). See also Elinor Ostrom, "Metropolitan Reform: Propositions Derived from

Two Traditions," *Social Science Quarterly* 53 (December 1972), pp. 474-93.

⁵See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.

⁶See Robert M. Stein and Dolores T. Martin, "Contracting for Municipal Public Services," ACIR Technical Paper, 1987.

⁷Richard A. Musgrave, *The Theory of Public Finance: A Study in Public Economy* (New York: McGraw-Hill Book Company, 1959), p. 15.

⁸Vincent Ostrom, Charles M. Tiebout, and Robert Warren, "The Organization of Government in Metropolitan Areas: A Theoretical Inquiry," *American Political Science Review* 55 (December 1961), p. 838.

⁹This account of the distinction between provision and production is based in part on the discussion found in Vincent Ostrom and Elinor Ostrom, "Public Goods and Public Choices," in E. S. Savas, ed., *Alternatives for Delivering Public Services: Toward Improved Performance* (Boulder, Colorado: Westview Press, 1978), pp. 7-49. The distinction between provision and production has also been developed by

- Ted Kolderie, "Rethinking Public Service Delivery" in Barbara H. Moore, ed., *The Entrepreneur in Local Government* (Washington, DC: International City Management Association, 1983), pp. 43-48. Also see Kolderie, "The Two Concepts of Privatization," *Public Administration Review* 46 (July/August 1986), pp. 285-91.
- ¹⁰See Gordon P. Whitaker, "Coproduction: Citizen Participation in Service Delivery," *Public Administration Review* 40 (May/June 1980), pp. 240-46.
- ¹¹Royce Hanson, "The Urban Future: New Policies and Issues," *Journal of Housing* 44 (January/February 1987), p. 18.
- ¹²A large literature exists on problems of majority voting. Two of the basic sources are Kenneth J. Arrow, *Social Choice and Individual Values*, rev. ed. (New York: John Wiley and Sons, 1963) and Duncan Black, *The Theory of Committees and Elections* (Cambridge: Cambridge University Press, 1958).
- ¹³Mancur Olson, "The Principle of 'Fiscal Equivalence': The Division of Responsibility among Different Levels of Government," *American Economic Review* 59 (May 1969), pp. 479-87.
- ¹⁴See Bruce W. Hamilton, "A Review: Is the Property Tax a Benefit Tax?" in George R. Zodrow (ed.), *Local Provision of Public Services: The Tiebout Model after Twenty-Five Years* (New York: Academic Press, 1983). For an alternative point of view, see George R. Zodrow and Peter Mieszkowski, "The Incidence of the Property Tax: The Benefit View versus the New View," in the same volume.
- ¹⁵The criterion of preference revelation does suggest a minimum size constraint for provision units as determined by the scale of specific problems. This criterion alone would limit the proliferation of small units; otherwise, the principle of fiscal equivalence would lead toward the organization of a separate municipality for every household.
- ¹⁶The costs of organization and operation are construed as transaction costs, in the main, because the discussion here pertains exclusively to the provision side. While it is true that some of the costs of setting up a provision unit will involve capital costs (e.g., building and maintaining a city hall) and labor costs (e.g., paying a city manager), most of the capital and labor costs of local government pertain to production, not provision. In general, provision-side costs are best viewed as transaction costs—the costs of individual and joint decision making involved in sustaining social relationships. Some labor costs can be viewed as specialized transaction costs, analogous to the cost of hiring a lawyer or a negotiator in the private sector, incurred in representing the interests of individuals and groups to others. For a general discussion of transaction costs, see Oliver E. Williamson, "Transaction-Cost Economics: The Governance of Contractual Relations," *Journal of Law and Economics* 22 (October 1979), pp. 223-61.
- ¹⁷David L. Chicoine and Norman Walzer estimate the *average cost* of financial control for one additional governmental unit per 1,000 residents in Illinois at \$1.21 per capita. For an additional special district the cost is estimated to be \$2.03 per capita. Chicoine and Walzer interpret these findings as "consistent with the commonly held view of duplication." Whether such "duplication" is wasteful, however, depends upon how citizens evaluate the benefit side of an additional unit. Chicoine and Walzer, *Governmental Structure and Local Public Finance* (Boston: Oelgeschlager, Gunn & Hain, 1985), p. 223.
- ¹⁸See *Chapter Six* for a discussion of these rules.
- ¹⁹Examples of both can be found in the ACIR study of St. Louis County. Forty percent of the county population remains unincorporated. Although incorporation movements have begun in some parts of the county, in response to changing preference patterns, other parts of the county have been recognized as communities for many years, but have not chosen to incorporate. A few small municipalities, on the other hand, have chosen to consolidate. See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.
- ²⁰The technical term is "production function," a specification of how various combinations of inputs can be transformed into outputs through the application of available technologies. See Werner Z. Hirsch, *Urban Economic Analysis* (New York: McGraw-Hill, 1973). The production functions of many local public goods and services are not known.
- ²¹One type of information needed to make sound production decisions has to be specific as to time and place. General knowledge is insufficient. For example, in street services, one has to know when a pothole develops and where it is located in order to patch it. Similarly, in police services, the ability to distinguish strangers from residents, or the unusual occurrence from the commonplace, depends upon having community-specific knowledge.
- ²²See Roger B. Parks, et al, "Consumers as Coproducers of Public Services: Some Economic and Political Considerations," *Policy Studies Journal* 9 (Summer 1981), pp. 1,001-11 and Gordon P. Whitaker, *op. cit.*
- ²³The most thorough discussion of economies of scale in the public sector is found in the work of Werner Z. Hirsch, *op. cit.*, pp. 331-34. In general, these economies are described by a u-shaped relationship to the size of local jurisdictions with constant returns to scale over a very wide range. For labor intensive services, such as police patrol, economies of scale are exhausted very quickly, perhaps by a community as small as 4,000 people. This refers to the actual patrol component, however, and not to support services, such as radio communications.
- ²⁴*Ibid.*
- ²⁵This subject is developed more fully in *Chapter Four*.
- ²⁶These implications are explored in a more abbreviated fashion in Ronald J. Oakerson, "Local Public Economies: Provision, Production, and Governance," *Intergovernmental Perspective* 13 (Summer/Fall 1987), pp. 20-25.

Separating Provision and Production

INTRODUCTION

To *distinguish* provision and production conceptually is not necessarily to *separate* production from provision in practice. Distinguishing provision and production does, however, lend support to both private and intergovernmental contracting as clear options for producing local public goods and services. The structure of a local public economy varies with the use of the contracting option. What leads provision units to choose one production arrangement over another? Why do some provision units contract out more than others? Some contract out very little of the goods and services they provide, while others approximate “pure provision” units, organizing little if any production in-house. Answers to these questions can help to explain the emergent structure of a local public economy and offer guidance on issues of governance and organization.

The extent of contracting determines to a large degree the structure of the production side of a local public economy, given a set of provision arrangements. Contracting tends to generate both greater specialization and greater competition among producers. Yet the extent of contracting varies with the structure of the provision side, and pure provision units tend to emerge only under restrictive conditions.

CONTRACTING OPTIONS

Current Practice

Contracting, both private and intergovernmental, is widely practiced. According to a recent ACIR analysis, 90.8 percent of the municipalities in a representative national sample reported at least one service contract among 34 service activities.¹ Most municipalities that contract out, however, use this mode of production in only a small proportion of their service responsibilities. Although contracting is employed to some extent by almost every community, the average

municipality contracts out for the production of only about one-fourth of the services it provides. (The mean percentage of service areas with at least one contract in the ACIR sample of municipalities is 27 percent. Forty-five percent of the municipalities contract for less than 15 percent of their service responsibilities.)² In-house production remains the norm, contracting, the exception, but the use of contracting—especially with the private sector—is increasing.

Significant change in both the extent and the type of municipal contracting has occurred since the 1970s. Contracting in the period prior to 1970 was heavily biased toward the public sector. ACIR found in its 1972 study of interlocal arrangements that intergovernmental contracting (i.e., contracting with another unit of government to supply a service) was the preferred alternative to in-house municipal production.³ Perhaps because of a lack of private vendors and/or a lingering concern with corrupt practices in the awarding of contracts, municipal governments avoided private service producers in favor of governmental jurisdictions when shedding service production. This reluctance to use private vendors has diminished significantly since 1972. The proportion of communities reporting at least one private service contract now exceeds the percent of cities reporting at least one intergovernmental service contract by 18 percentage points. This reflects the growing number of private service contracts, however, not an absolute decline in intergovernmental contracting.

Municipal contracting in the 1980s is dominated by two quite disparate service categories—public works and health and social welfare services. For both of these functional categories, nearly two-thirds of the municipalities in the ACIR sample reported at least one service contract. General government operations (e.g., legal, clerical, and computer services) is the third most frequently contracted service area, with 41 percent of the sample reporting one or more contracts. Municipal use of service contracts in the areas

of transportation, public safety, and education ranges between a third and a quarter of our sample. Contracting for park and recreational services is less frequent, with only 13 percent of the sample reporting any contracted services in this functional area.

Contracting for both public works and health and welfare is dominated by the private sector. Over half of the municipalities in the sample had contracted with at least one private vendor for these services. Governmental units are the dominant contract producer only for public safety services. Among private producers, for-profit vendors dominate service contracts for general government services, transportation and public works. Their nonprofit counterparts have a significant edge in service contracts for cultural activities (e.g., libraries and museums) and a slight advantage in the proportion of municipalities contracting with them for health and welfare services. Neighborhood associations, while an important new contract producer, to date have had a negligible impact on the level of private service contracting. Their greatest contribution to service production is in the area of public safety and health and welfare.

Empirical studies of garbage collection,⁴ electrical power,⁵ fire protection,⁶ police protection,⁷ and an assortment of custodial and general services⁸ have found that contracted service production nets significant sav-

ings over in-house government production. These findings suggest that contracting is frequently a more efficient arrangement for producing a number of different services provided by local government. The presence of efficiency gains from contracting is usually explained in terms of (a) economies of scale in production and (b) the benefits of greater competition on the production side of the local public economy.

The precise advantage or disadvantage of contracting, with respect to specific goods and services, can be expected to vary from one provision unit to another. Contracting is not uniformly preferred to in-house production for any given service type across all provision units in all circumstances. Small communities may choose to contract out for reasons of economies of scale that larger communities can capture in-house; both small and large communities may choose to do so for reasons of competition, but the extent of competition may vary from one local public economy to another. The organization of the *provision side* of a local public economy has substantial impact on the organization of the *production side*. This becomes clear from the ACIR analysis.

Explaining the Variation in Contracting

Table 2.1 summarizes the results of a multivariate analysis of municipal contracting.⁹ The dependent

Table 2.1
LOGIT ESTIMATES FOR DEGREE OF MUNICIPAL CONTRACTING, 1982-83

<u>Variable</u>	<u>Logit Estimate</u>	<u>Standard Error</u>
Intercept	3.72	.723
Population ²	.524-5E*	.160-5E
Population	-.000*	-.000
Number of Units of Local Government in SMSA	.001*	.000
Tax Burden	-3.71	.346
Relative Size of Municipal Workforce	-15.79*	11.030
Property Tax Limitation (length of time in effect, 0 = none)	.182*	.081
Local Expenditure Limitation (length of time in effect, 0 = none)	.103	.082
State Centralization	-.039	-.018
Ease of Local Annexation and Consolidation	-.173*	.055
City Manager Form	.328*	.124
Percent of Work Force Unionized	.004	.010

*P < .05

-2 Log Likelihood Ratio = 1,561

X² of log likelihood ratio = 75.8*

Percent cases predicted = 84.1

Gamma = .308*

Percent of cases incorrectly predicted by categories of dependent variable:

0-10% of services contracted = 9.7

10-33% of services contracted = 0

>33% of services contracted = 16.5

variable is the percentage of municipal functions in which at least one contract is reported, broken into three categories. Most of the independent variables included in the model measure some aspect of provision arrangements, or rules governing the organization of the provision side. Several of these variables are significantly related to the use of contracting, as measured here. The model correctly predicts a little more than 84% of the cases.

Among rules affecting the provision side, two types are significant: (1) state-imposed property tax limitations on local government and (2) the relative ease of municipal annexation and consolidation. More long-standing property tax limitations and more constraining procedures for municipal annexation and consolidation are both associated with more contracting. (Expenditure limitation is not significant.) One interpretation of these findings is that local public economies that place more power in the hands of citizens—to approve increases in tax rates and adjustments of municipal boundaries—tend to engage in more contracting. Tax burden is not significant in this model, although this finding is not supported by other empirical work.¹⁰

Another variable that significantly affects the degree of municipal contracting is the number of local governments located within a municipality's Standard Metropolitan Statistical Area (SMSA). This variable can be interpreted as a measure of the "extent" of the local government market for services. The number of provision units in a local public economy is equivalent to the number of potential consumers of externally produced services. At least since Adam Smith, economists have hypothesized that specialization in production varies with the extent of a market. The greater the number of consumers in a market, the more specialized production becomes. A greater degree of contracting in a local public economy is associated with greater specialization on the production side. Increasing the number of provision units within a local public economy tends to increase the reliance on contracting of any single provision unit within that setting. Differentiation on the provision side has an amplified effect on the degree of differentiation on the production side.

The relationship of municipal population size to contracting is curvilinear in this model. Both small and large municipalities tend to contract out more than mid-size municipalities.¹¹ Economies of scale may account for the greater propensity of small municipalities to engage in contracting. The same explanation is unavailable, however, to explain greater contracting by large municipalities. Indeed, a rationale for increasing the size of local governments, from a traditional perspective, is to avoid dependence on outside producers—to be self-sufficient. This effect appears to be quite limited. Consistent with an economies of scale

argument, however, is a diseconomies of scale argument. Average costs are generally understood to have a u-shaped relationship to scale of production. As production increases from zero, average costs initially decline, then level off, and finally after a point begin to increase. One important variant of diseconomies of scale in production is diseconomies of scale in management. As the size of a provision unit increases, beyond some point, scale economies attained as a technical matter of production may be offset by management difficulties that multiply as the provision unit attempts to organize more and more production in-house. From this perspective, the tendency toward increased reliance on contracting by larger municipalities would reflect an effort by public managers to simplify their internal management problems by securing outside production of selected service components.

The sample of municipalities used to generate the estimates in *Table 2.1* was limited to municipalities over 10,000 in population. If smaller provision units were included, the relationship between small size and a greater reliance on contracting would almost certainly be more pronounced. An ACIR field study of the highly differentiated local public economy found in St. Louis County, Missouri, examined contracting in three service areas—police, fire, and streets—for 90 municipalities ranging in size from less than 1,000 residents to about 55,000 residents.¹² Municipalities under 10,000—and especially those under 5,000—relied much more heavily than larger municipalities on service contracts with external producers. Indeed, for the production of basic services delivered directly to citizens—police patrol, fire fighting, street repair—municipalities tended to shift to in-house production at the minimum point where significant economies of scale can be captured. For these basic services, the size threshold for in-house production appears to be quite low—somewhere around 1,000 persons for basic street services and 2,000 persons for police patrol.

Another provision-arrangement variable that is significant in *Table 2.1* is the choice of a city-manager form of government. This finding should be treated with some caution, however, since it has not been substantiated by other analysts.¹³

The Relevance of Provision Arrangements

In order for citizen-consumers to realize a benefit from contracting, there must also exist a provision unit able to acquire information about alternative producers, choose a production mode, select a specific producer, negotiate a contract, and monitor performance. These are not idle tasks. They require agents who bring a citizen-consumer perspective to their work and who remain alert to alternative possibilities. Provisioning activities are crucial to the utility of contracting.¹⁴

Despite evidence that contracting offers substantial efficiency gains, empirical research also discloses considerable variability in those gains.¹⁵ This variation has not been explained, and further research is needed. Variation in the performance of activities on the provision side, however, may account for much of the variation in efficiency gains.

It is not difficult to imagine situations in which contracting harms, rather than helps, the interests of citizens. The possibilities range from "sweetheart" contracts where the chosen vendor is, for one reason or another, not the most efficient, to outright corruption of public officials. Corruption, however, is a problem on the provision side. The use of contracting does not relieve citizens and officials of the need to maintain appropriate provision arrangements, including accountability for the expenditure of public funds.

Provision arrangements are also important because provision determines how efficiency gains from contracting will be distributed. Who benefits from contracting? Do citizen-consumers derive a benefit, either from tax savings, increased levels of service, or both? Or do local government politicians, managers, and bureaucrats grab the lion's share of benefit by in effect distributing the efficiency gains from contracting in ways beneficial to their particular interests, especially by increasing provision levels rather than decreasing revenues?

The ACIR study contains some interesting results relevant to these questions. First, there is evidence that efficiency gains from contracting tend to reduce municipal expenditures when municipalities contract out less than 25 percent of their service responsibilities. Second, however, there is evidence that gains from contracting tend *not* to reduce expenditures when contracting moves beyond 25 percent of service responsibilities. The meaning of this finding is unclear. Robert M. Stein and Dolores T. Martin¹⁶ conjecture that bureau managers who contract out some portion of their work load reach a bargain with city managers that allows the gains from contracting to remain within the bureau. The gains then go toward increased service provision rather than decreased spending. On the other hand, the finding may mean that citizens are choosing to support increased service provision in view of the lower tax price for services made possible by contracting. If there are efficiency gains from contracting, it follows that the more a municipality is able to contract out (presumably within some limit that varies from place to place), the lower is the tax price of services provided. As the tax price decreases, the quantities demanded by citizens can be expected to increase.

The fundamental importance of the contracting option is the ability of a provision unit to *choose*. The availability of a marketplace on the production side does not necessarily mean that provision units should

always choose to enter the market, as collective consumers, rather than produce for themselves.¹⁷ A basic function of provision units is to decide how to arrange for production. For example, a study of municipal contracting in Los Angeles County, California,¹⁸ found that, in general, contracting is associated with lower expenditures. However, an analysis of municipalities that contract with the Los Angeles County Sheriff for production of police services discovered the following: communities that want a relatively low level of policing seem to do well by contracting; communities that want a higher level of policing do well producing for themselves. In other words, by contracting, communities that demand less police protection can get what they want without having to organize a police department; however, those communities that demand more police protection would find a contract arrangement with the county sheriff to be more expensive than in-house production.¹⁹ In a situation such as this one, it is the ability of provision units to choose how to arrange for production, and not the inherent superiority of one production mode over another, that becomes the key factor in determining efficiency outcomes.

"PURE PROVISION" UNITS

Are Small Local Governments Nonviable?

The issue of local government viability was addressed in some detail by the ACIR in a 1969 report on *State Aid to Local Government*. The Commission urged states to establish criteria of viability to be used to assess the worthiness of a local unit to receive state funds or to compel its dissolution. The relevant criteria were seen to include (1) measures of fiscal capacity, (2) economic diversity, and (3) minimum population and geographic size. The third criterion was assumed to be relevant to realizing economies of scale and making appropriate use of both specialized personnel and new technology. The states were commended for having used their powers over local governance to reduce greatly the number of school districts in the nation and urged to undertake the same course of action with respect to general-purpose governments.²⁰

Although the language of the report presupposed a large number of nonviable units, the ACIR did not propose specific criteria to ascertain their existence until 1981. In its comprehensive report on *State and Local Roles in the Federal System*, the Commission recommended that states use two minimum standards of viability: (1) that any local government (general or special purpose) located in the urbanized portion of an SMSA have the equivalent of at least one full-time employee, or (2) that all general-purpose units "perform at least four functions, or only two functions, provided that each of the two constitutes at least 10%

of the jurisdiction's current expenditure budget." The Commission further recommended dissolution of local governments that do not meet either of these standards. The major empirical finding cited to support the recommendation consisted of 1977 Census of Governments statistics reporting 4,424 municipalities, 8,673 townships, 17,534 special districts, and 280 school districts with no full-time equivalent employment, as well as 9,614 municipalities with a population under 1,000.²¹

Distinguishing provision and production raises a serious question about the basic logic of this recommendation, repealed by the Commission in June 1987.²² Units that "perform" few functions in the sense of both providing and producing goods and services can be quite viable—and active—as provision units. Jurisdictions that produce little or nothing for themselves are nonetheless active in shopping for services from among potential vendors, both public and private, and in raising revenue to pay for those services. These activities, usually undertaken by citizens who serve as part-time officials, give representation to the interests of citizen-consumers in those jurisdictions. Criteria pertaining to the number of full-time equivalent employees are irrelevant to the viability of a provision unit. Criteria pertaining to number of functions "performed" are similarly irrelevant, when performance is defined to include both provision and production.

These conclusions are supported by an ACIR study of St. Louis County,²³ which contained 22 municipalities with an estimated population under 1,000 in 1984. Nearly all of these small units are active as providers of the two basic services provided by municipalities in St. Louis County—police and streets. Only a single municipality fails to report provision in both service areas. Missouri law requires that municipalities with a resident population of 400 or more persons provide full-time police services. Among the nine municipalities not required to make provision by state law, only one, with a population of 31 persons, fails to make explicit provision for police services beyond that provided by county police. Six contract to receive service (two of which supplement the contract with their own part-time producers) and two maintain their own part-time departments. Street services indicate somewhat less activity. Four municipalities, including the single nonprovider of police services, reported no expenditures for street services in 1985; their average population was 196 persons. (At this scale of organization, however, there is a distinct possibility of making regular provision for street repair at intervals that exceed one year.) Still, the average level of provision among the remaining municipalities was substantial. The average street expenditure for these 18 municipalities was about \$75 per household, compared to an average street expenditure of \$104 per household for

the great majority of county municipalities (those for which comparable data were available). Information supplied by 10 out these 18 municipalities disclosed only one that produced any street services for itself. Unlike police, most contracting for street services is with private producers.

The appropriateness of the earlier ACIR recommendations urging states to establish criteria of viability, without proposing any specific criteria, are more difficult to evaluate. The adaptive capabilities of local public economies require that provision units be subject to modification, including the possibility of both dissolution and merger. The governance processes of local public economies typically contain procedures that allow these changes to occur. The basic question is one of the preferred process: what decision makers are best situated to judge the viability of a provision unit? Specifically, can local citizens make these judgments for themselves or is state intervention needed? A comparative assessment of specific institutional mechanisms must await further research.

Do "Pure Provision" Units Have Advantages on the Provision Side?

The separation of provision and production casts a much different light on the potential economic viability and usefulness of very small local governments. As is now widely accepted, inability to produce does not entail inability to provide. The potential for contracting out the production of public goods and services (with either public or private vendors) offers the possibility of maintaining "pure provision" units—local governments that produce little or nothing for themselves, but remain very active as providers. The implications extend well beyond the potential viability of small local governments. A number of municipalities in Los Angeles County, California, for example, function as pure provision units—through the so-called Lakewood Plan—although they have sufficient size to capture significant economies of scale from in-house production.

Clearly, it is incorrect to assume that local governments that produce no services are inactive. On the provision side, these units may be very active indeed. These activities include raising revenue, holding elections, deliberating on the needs of the community, choosing desired goods and services, determining levels of supply, shopping for service vendors of choice, negotiating contracts, and monitoring service flows.

In very small units, much of this work is done by ordinary citizens. Professional services, perhaps those of a planner or an engineer, a lawyer or an accountant, can also be employed as needed. *Lack of production does not indicate lack of performance.*

An additional question is whether "pure provision" has advantages on the provision side as well as on the production side. Most discussions of contract-

ing focus on production-side advantages—the ability to produce services at lower cost; however, there may be important provision-side advantages as well. Anthony Downs has laid out such an argument in an essay titled “Separating the Planning and Procurement of Public Goods from Their Production and Delivery.”²⁴ In our terms, “planning and procurement” are provision-side activities. Downs argues that organizing production extensively in-house biases provision-side decisions in favor of producers rather than citizen-consumers. The interests of producers are commingled with those of consumers, resulting in provision-side decisions that fail to represent consumer interests fully. Those who are elected to represent citizens instead acquire a production-side perspective. Production criteria come to dominate the decisions of provision units. The interests of producers and consumers are no more the same in the public sector as in the private sector.

Downs does not apply this argument specifically to local government contracting. Clearly, however, a mayor or city manager in a pure provision unit fulfills a much different role than the same officer in a provision unit that organizes most service production in-house. In a pure provision unit, a chief administrative officer (CAO) performs a procurement function, representing the interests of consumers to producers. If, on the other hand, production is organized in-house, the CAO must fulfill two roles—both representing consumers and organizing production. An incentive to represent consumers may be weakened by the CAO’s direct ties to production. Accountability would, in this case, be more easily obtained when provision units are free to represent the provision side exclusively.

It follows that contracting not only offers potential economies on the production side, but also may enable provision units to perform provision functions more effectively. This argument contrasts to the more common argument made by critics of contracting—that one of the disadvantages of contracting is a loss of direct control over production by public managers. If public managers are viewed as representing a public interest, then loss of direct control can be seen as disadvantageous. On the other hand, to the extent that public managers tend to reflect producer interests, loss of direct managerial control over production may actually enhance citizen control.

More research needs to be done to determine which view has greater accuracy. Public managers do not uniformly adopt a producer’s perspective. A distinction needs to be made, for example, between de-

partment heads and central management. The characteristics of provision units, including size, heterogeneity, and form of government, may also have effects upon the perspective of a CAO. Is it possible for a public manager to decide even-handedly between in-house production and contracting, while maintaining both options? Still, the possibility that contracting also offers provision side advantages cannot be dismissed. Pure provision units may have distinctive properties that deserve to be emulated rather than eliminated.

CONCLUSION

A considerable body of research suggests that contracting potentially offers production-side advantages—obtaining greater efficiency from economies of scale for small provision units, helping to simplify public management in large provision units, and offering gains from producer competition for any provision unit. The incentive of public managers to contract out is less clear, however. One line of reasoning suggests that the incentive of managers to shift production to outside vendors is limited to circumstances in which the efficiency gains from contracting can be captured for increased service production. This presupposes a producer’s bias on the part of public managers. Some limited evidence of this outcome emerges from the ACIR analysis of municipal contracting data, which shows that those municipalities engaging in greater levels of contracting do not reduce their total expenditures relative to those municipalities that engage in less contracting. This could mean that even those municipalities that contract out more than 25 percent of their functions continue to exhibit a producers’ bias. Such an interpretation, however, must be regarded as conjectural.

“Pure provision” units, on the other hand, are relatively free of a producer’s bias.²⁵ Such units offer the potential for obtaining the least distorted expression of consumer preferences possible in the public sector. In Los Angeles County, California, pure provision units include mid-size suburban municipalities. In St. Louis County, Missouri, by contrast, pure provision units are limited to those so small that they would be unable to capture sufficient economies of scale to produce services in-house. Yet nearly all of these units can be characterized as active providers. Whether their performance as provision units is superior to those units that also organize in-house production is a question that awaits further research.

NOTES

¹Robert Stein and Dolores Martin, under contract to ACIR, studied the experience of American local governments with respect to contracting in the early 1980s. For this purpose, data on municipal contracting was collected from two separate surveys of municipal governments for the period 1982-83. Both surveys were conducted by the International City Man-

agement Association. The samples for both surveys were selected from the same “universe” of communities: all cities with populations over 10,000 and a one-eighth sample by geographical region of cities with populations under 10,000. The 1982 study surveyed local governmental units about their use of private sector service contracts for 64 functional activities. The 1983 study queried a comparable sample of governments on their use of intergovernmental and joint service agreements

for the provision of 42 functional activities. Data from both surveys were merged producing a sample of 890 communities who responded to both surveys. The merged data set provides information on intergovernmental and private (profit and non-profit) service contracts for 34 functions in seven functional areas. The distribution of cases for the merged sample on measures of region location, population size, form of government and metropolitan status show them to be representative of the total "population" of all cities over 10,000 in the United States. The sample of cities under 10,000 is too small (only 17 observations) to allow any meaningful generalizations about this class of cities. The complete data analysis is contained in Stein and Martin, "Contracting for Municipal Services," ACIR Technical Paper, 1987.

²Unfortunately, the measure of contracting used here does not allow an assessment of the relative importance of contracting in dollar terms compared to municipally delivered goods and services. More detailed spending information is unavailable.

³ACIR, *Intergovernmental Service Arrangements for Delivering Local Public Services: Update 1983*, A-103, October 1985, pp. 45-53.

⁴Thomas Borcharding, Werner Pommerehne, and Frederick Schneider, "Comparing the Efficiency of Private and Public Sector Arrangements," *Journal of Law and Economics*, Supplement 2, (1982), pp. 127-56.

⁵Louis De Alessi, "An Economic Analysis of Government Ownership and Regulation: Theory and Evidence from the Electric Power Industry," *Public Choice* 19 (1974), pp. 1-42, and Robert W. Spann, "Public Versus Private Provision of Government Services," in Thomas Borcharding (ed.), *Budgets and Bureaucrats: The Sources of Government Growth* (Durham, NC: Duke University Press, 1977), pp. 71-89.

⁶Roger S. Ahlbrandt, Jr., "Efficiency in the Provision of Fire Services," *Public Choice* 18 (1973), pp. 1-15.

⁷Stephen L. Mehay, "Intergovernmental Contracting for Police Services: An Empirical Analysis," *Land Economics* 55 (1979), pp. 59-72. See also, for somewhat mixed results, Sidney Sonenblum, John J. Kirilin, and John C. Ries, *How Cities Provide Services: An Evaluation of Alternative Delivery Structures* (Cambridge, Mass.: Ballinger Publishing Co., 1977), Chapter 5.

⁸Robert T. Deacon, "State and Local Expenditure," in John E. Peterson and C.L. Spain (eds.), *Essays in Public Finance and Financial Management: State and Local Perspectives* (New York: Chatham House, 1978), pp. 22-33, and Thomas Borcharding, Werner Pommerehne, and Friedrich Schneider, *op. cit.*

⁹The analysis presented here is from the study, cited above, by Robert M. Stein and Dolores T. Martin, conducted for ACIR. See Stein and Martin, *op. cit.* The reader may want to compare the findings discussed here with an analysis by James M.

Ferris, "The Decision to Contract Out: An Empirical Analysis," *Urban Affairs Quarterly* 22 (December 1986), pp. 289-311. See also James M. Ferris and Elizabeth Grady, "Contracting Out: For What? With Whom?" *Public Administration Review* 46 (July/August 1986), pp. 332-44.

¹⁰This finding is not borne out in the analysis done by Ferris, *op. cit.*

¹¹This finding emerges only from a multivariate analysis that controls for other factors affecting the level of contracting. A large number of small governments are found in non-metropolitan areas where potential competition among vendors would be lower.

¹²See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.

¹³Ferris, *op. cit.*, finds no relationship between city manager form and extent of contracting.

¹⁴Because provision arrangements are public, the treatment of contracting as a form of "privatization" is somewhat misleading. Only the production side is subject to privatization with contracting, not the provision side.

¹⁵See Barbara Stevens, *Delivery of Municipal Services Efficiently*, U.S. Department of Housing and Urban Development (1982). In an intensive series of case studies examining eight services in 20 cities, half self-producing and half contracting, Stevens found substantial variation in efficiency gains from contracting, while accounting for variation in service quality.

¹⁶See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.

¹⁷Stevens, *op. cit.*, found that some self-producing cities performed specific functions more efficiently, taking account of service quality, than contracting cities, even though the pronounced tendency was for more efficient performance by contracting cities.

¹⁸Sidney Sonenblum, John J. Kirilin and John C. Rees, *op. cit.*

¹⁹*Ibid.*

²⁰ACIR, *State Aid to Local Government*, A-34, April 1969, pp. 25-27.

²¹ACIR, *State and Local Roles in the Federal System*, A-88, April 1982, pp. 446-48.

²²See *Recommendation One*, this volume.

²³See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.

²⁴Anthony Downs, *Urban Problems and Prospects*, 2nd ed. (Chicago: Rand McNally College Publishing Company, 1976), Chapter 12.

²⁵Officials in pure provision units can nevertheless exhibit a "provision bias," i.e., a bias in favor of providing, as opposed to not providing, services.

Differentiating Provision

INTRODUCTION

To the extent that state and local rules allow, the provision side of a local public economy will tend to be differentiated among a variety of provision units, small and large, with some units “nested” inside others (e.g., municipalities within a county). If production can be separated from provision, it follows that production criteria no longer need to determine, or even influence, the choice or design of provision units (e.g., municipalities, special districts, self-assessing residential associations). Small provision units can enter into relationships with large production units, just as large provision units can do business with small production units. Examples of both abound. Similarly, ordinary citizens, who make decisions about provision, can engage professionals to make decisions about production, just as generalist professionals who serve as agents of citizens can engage specialists. Production criteria, such as economies of scale and professionalism, need not be deciding factors in the formation and design of provision units. The extent of differentiation on the provision side is limited only by (1) the rules governing the formation of provision units and (2) expected transaction costs—specifically, the total costs of organizing and the marginal costs of operating each additional unit (as discussed in *Chapter One*).

TYPES OF PROVISION UNITS

The variety of potentially useful provision units is quite large and, very likely, not yet fully developed in practice. Traditional categories, such as general-purpose versus special-purpose local governments, are inadequate for understanding the full range of possibilities. Provision units can vary along a number of dimensions: (1) method of formation—especially whether public or private; (2) size of jurisdiction—one neighborhood, many neighborhoods, or areawide; (3) scope of provision—the range of services that can potentially be provided; (4) coercive powers, including regulatory capabilities, available to provision units; and (5) preemptive powers with respect to other units

that may be nested within a provision unit. It is important to describe provision units in accordance with such functional categories rather than simply by their formal or legal name. Sometimes the legal name—city or village—is apt and descriptive, but often not. Analyzing intergovernmental policy on the basis of the names of units rather than their functional attributes can result in misleading conclusions.

Three Basic Types

Among the traditional provision units created as a matter of public law, three types can be distinguished. The first is a unit created by local citizens at their option with boundaries largely of their determination. Most “municipalities” fit this category. The second is a unit created by state government, usually throughout a state, with boundaries determined independently of local choice. Virtually all counties and most townships are included in this category. The third is a unit usually created by local citizens at their option with boundaries largely of their determination *as a supplement* to the first two categories. This category is composed mostly of “independent special districts.”

These three types of units complement each other in important respects. The limiting factor in most municipal organizations is that one municipality cannot overlap another—certainly not one municipality of the same type as the other. Cities are not nested inside cities, nor villages inside villages. Counties and townships have the virtue of being able to overlap other types of units.¹ Thus cities and villages, as well as townships, are usually nested inside counties. The limiting factor on county and township organization is their ordinarily inflexible and arbitrary boundaries, determined usually on a statewide basis rather than by local residents. Special districts have the twin virtues of being able to overlap other units, including other special districts, and of having flexible boundaries, often determined by local citizens.

Counties and townships can usually be understood as units that exist to make provision for some minimal level of local services—that level apt to be demanded

by any community. They also exist to deliver state services on a decentralized basis. Municipalities are created to provide additional increments and new types of services in those areas where local demand is greater (or in some cases lesser).² Independent special districts can be used to further differentiate provision in two ways: (1) varying the boundaries from those of a municipality, township, or county to accommodate a particular community of interest or (2) creating a redundant decision structure (e.g., a board of education) with boundaries coterminous with a general-purpose unit in order to make separate provision for a specific set of services.

The Special Case of Special Districts

Numerous ACIR recommendations over the years have rested on a conclusion that the use of special-purpose, as compared to general-purpose, governments ought to be constrained. In a 1964 report, *The Problem of Special Districts in American Government*, the Commission concluded that, in spite of their utility in filling institutional gaps among general-purpose governments in a locality, too many special districts had been created or allowed to continue when no longer needed. Since then, the number of special districts has increased by 56 percent, from 18,323 in 1962 to 28,588 in 1982. (This number does not include school districts.) The perceived disadvantages of special district governments have included fiscal fragmentation, lack of accountability, and inefficient performance of services as compared to general-purpose governments.³

Special districts vary widely in the services provided and in their relationships with other units of government. The principal referent in this section is to "independent" special districts (i.e., those that are sufficiently autonomous from other units of government to possess their own taxing authority). "Dependent" special districts are generally created by the action of other units rather than directly by citizens. Dependent districts tend to be managed by representatives of those other units, and do not ordinarily have separate taxing authority. More restricted still are special tax districts within a provision unit that allow a separate tax rate to be levied within the district but provide for no autonomous decision making. The latter can be used to attain better fiscal equivalence, but do not supply political representation commensurate with the special taxing authority. Most school districts are independent special districts. Even those not independent in the strict sense usually possess considerable autonomy, except for taxation and final approval of a budget.

Special districts may be viewed as an effort to economize on the marginal costs of operating additional provision units. Citizens may choose to create a special district, for example, instead of organizing a municipality. If the citizen demand for services that

justifies the organization of a new unit is limited in scope to one or a few service areas—perhaps fire protection or sewerage—then organization of a municipality would unnecessarily add to the marginal costs of operating the new unit. The decision making or transaction costs of operating a provision unit, including the costs of citizen participation, can be reduced by limiting the scope of actions that can be taken. Special districts tend to be much more limited in the actions they are authorized by law to take than are most general-purpose governments.

It follows that the availability of the option of creating a special district enables citizens to obtain a more optimal set of provision units. States that severely constrain the creation of special districts may therefore limit citizens to less optimal forms of organization. Forcing citizens to choose between no additional provision unit (beyond the basic unit—often a county—created by state government) and creating a general-purpose government (such as a municipality) cannot increase efficiency on the provision side.

Special districts also offer alternatives to annexation and consolidation by general-purpose governments. Eliminating this option could only reduce provision-side efficiency. Suppose that two adjacent municipalities are unable separately to provide for a single service that affects both—perhaps public transportation. Requiring both municipalities to merge in order to provide a single service may distort the expression of citizen preferences for other services. Without the option of creating a special district, citizens are left to choose between two decidedly suboptimal alternatives.

Illinois and California lead the nation in numbers of special districts. A comprehensive study of special districts in Cook County, Illinois, where the 196 districts included 93 organized to provide park services and 45 to provide fire protection, concluded that their advantages outweigh disadvantages to citizens.⁴ A study of special districts in California found that district governments fulfilled a great variety of legitimate citizen demands that could not satisfactorily be met through general-purpose governments.⁵

To some extent, special districts may be used to circumvent restrictions on local government authority (e.g., debt restrictions). David L. Chicoine and Norman Walzer argue that this may be the case to some extent in Illinois, and cite, in support of the proposition, a finding that counties in which a larger percentage of the population lives in home-rule municipalities (with no tax and debt limits imposed by state statute) tend to have fewer taxing units per 10,000 residents.⁶ Statutory tax and debt limits that do not allow local citizens to approve increases beyond those limits can distort the provision of local public goods and, if special districts offer a route of escape, also distort the structure of the local public economy. Special districts created to avoid such restrictions do not necessarily

have other advantages that would lead to their creation. Local citizens are then being forced to suboptimize. The remedy for this situation is not to constrain the formation of special districts, but to remove overly constraining restrictions on local government authority.⁷

Citing the number of special districts, or their rate of growth, is exceptionally weak evidence of poor performance for this generic type of unit. Citizens, who often create special districts pursuant to legislative authority, presumably would not do so unless they expected the new unit to perform better than existing units for the specific purpose at hand. Independent special districts, in general, are supplementary units. They exhibit more variety than any other type of provision unit. Their variety and flexibility can allow communities of interest that do not happen to coincide with existing local government boundaries to organize in order to provide a limited set of goods and services. To be sure, all special districts are not equally worthy, and the citizens and officials who create such districts can make mistakes, but each district should be evaluated on the basis of its performance as a provision unit, especially its ability to represent citizen-consumer interests. A blanket bias against special districts does not appear to be warranted, given the limitations inherent in the design of so-called general-purpose governments. The rapid growth in the number of special districts may largely reflect those structural limitations.

Neighborhood Provision Units

Most local public economies consist of a variety of provision units—counties, perhaps townships, municipalities of various size and type, and diverse special districts. Some of these units exist side by side; some are nested inside others. Some local public economies, however, are not highly differentiated on the provision side. Large municipalities tend to preclude the organization of supplementary provision units within their boundaries. Proposals to allow for the formation of neighborhood governments would increase the potential variety of provision units.⁸ These proposals are based upon a judgment that larger municipalities—especially, but not only, central cities—may often not adequately reflect the interests of diverse neighborhoods within their boundaries.

Small units of local government—those under 1,000 population—have been characterized as “toy governments,” “postage stamp governments,” “peanut governments,” and “lilliputs.” Somehow, the term “government” is identified with greater concerns than maintaining the livability of a neighborhood that contains no more than a few hundred households. The legal nomenclature associated with local governments is often no help. Fourth class “cities” in Missouri, for example, have a maximum size of 3,000 people. (Sometimes, the legal nomenclature does help by des-

ignating such units as “villages.” But just as frequently, we ignore the distinction by lumping villages and cities together, for purposes of analysis, as “municipalities.”) If the government of a “city” of 3,000 residents attempted to function like a city of 30,000, it would not be economically viable. A residential community of 3,000 is not, however, a city—at least not in the customary functional sense. Nor is its government really a “city government,” except in name, despite the presence of a mayor and council.

A disparity between name and function does not make small municipalities insignificant. In metropolitan areas, small villages and “cities” are often functionally *neighborhood* governments. Once the legal nomenclature is set aside, neighborhood governments can be seen to exist in large numbers, mostly as small municipalities nested within metropolitan counties, even though the number of such units is often not sufficient to affect large numbers of people when compared to the populations of the largest municipalities.

Another way to think about neighborhood governments is as a public analog of private subdivisions. Residential subdivisions, when collectively organized, are able to function as provision units for their residents. While little is known about their functions, such subdivisions are the most rapidly growing type of local provision unit around the nation.⁹ Although formally and legally a private association, organized subdivisions are functionally equivalent, for homeowners, to a municipality—able to make collective provision for local public goods and services. Subdivision associations are private in the sense that they are initially organized by a developer who creates the association by attaching membership and fee requirements as deed covenants when each separate parcel is sold. Thus the subdivision exists as a matter of private rather than public law. Yet residents become members of the subdivision association upon acquiring title to property just as they become citizens of a municipality upon taking up residence within its boundaries. The difference is between property ownership as a qualification for membership and residence as a qualification for citizenship. With membership comes a legal obligation to pay dues, as well as rights to participate in decisions, either directly or through an elected board.

In St. Louis County, Missouri, the legal form is a “trust indenture,” which involves the creation of a board of trustees and the payment of a subdivision fee, most often assessed on a flat rate or frontage basis. Nonpayment becomes a lien on the property. With rare exceptions, these subdivisions are “pure provision” units. All services are contracted out, either to private or public vendors. Especially instructive in the St. Louis County experience is the successful use of subdivisions to provide a full range of residential street services as determined by residents, including street and sidewalk repair, cleaning, snow removal, lighting,

mowing, and trimming. Many subdivisions also regulate street access to their community, closing streets that would open their residential communities to arterial traffic. Street services, as a result, are the most highly differentiated aspect of the local public economy in St. Louis County.¹⁰

Neighborhood governments have been notoriously difficult to create *within* large incorporated cities. This is because, in the "rules of the game," municipal incorporation generally *preempts* any further public incorporation without the consent of the municipality. Neighborhoods cannot ordinarily incorporate as municipal governments within an already incorporated city. A municipality is viewed, not only as a general purpose government, but as a monopolist. This rule inhibits the development of a differentiated local public economy within large central cities. In the suburbs, on the other hand, a highly differentiated local public economy is able to develop because county government—the only general-purpose government outside the city—does not generally preempt municipal incorporation. Neighborhood governments are likely to emerge in large central cities only as state law (or local charter) is amended to extend neighborhood incorporation powers, analogous to municipal incorporation powers, to citizens. To create a subdivision association in an established neighborhood would probably require some relaxation of the rule of willing consent. In form, it would be more like a municipality, but it could also be limited in function to those activities deemed appropriate for neighborhood provision.

In 1967, the ACIR recommended that states authorize large cities and county governments in metropolitan areas to establish "neighborhood subunits" with limited functional and taxing powers.¹¹ The term "subunit" was carefully chosen. Cities and counties would be authorized, not required, to create these entities, and could dissolve them at will. Subunits of existing local governments were distinguished from new local units that would exacerbate perceived problems of urban "fragmentation." In 1985 the Commission renewed its recommendation, urging the use of neighborhood subunits in distressed communities, but omitting reference to dissolution and adding the possibility of creation through initiative and referendum procedures.¹² The latter would allow for creation by citizens, in a manner much like municipal incorporation.

Organized neighborhoods need to be understood as autonomous provision units. The same sort of incorporation process that gives rise to municipalities can be applied to the organization of neighborhoods. The key is a state statute (or municipal charter provision) that specifies (1) a procedure by which *citizens* can act to create a neighborhood unit on their own motion and (2) the powers and duties of such a unit once created. The latter should include, at a minimum, power to raise revenue, receive revenue by

transfer, perform services under contract, and accept functional responsibility by transfer from an overlapping jurisdiction.

Neighborhood units can be understood as analogous to municipalities or special districts, rather than counties and townships. Instead of attempting to create neighborhood subunits on a uniform basis throughout a city, a neighborhood incorporation approach envisions an incremental process of constitutional choice on the part of local citizens in discrete neighborhoods. Neighborhood units would then be formed only where citizens have calculated that the advantages outweigh the costs. Moreover, only a process of neighborhood unit formation removed from city hall can succeed in creating alternative, potentially competitive, processes of collective choice for the distinct neighborhoods of large cities. Until this happens, the advantages of a diverse local public economy are not likely to be available to the largest concentrations of urban population in America.¹³

PERFORMANCE DIFFERENCES AMONG PROVISION UNITS

Different types of provision units are appropriate for various local public goods and services. To say the same thing in reverse, no single type of provision unit is equally well suited to providing for all local public goods and services. One dimension on which provision units vary widely is *size*. The public problems to which service provision responds also appear in a variety of sizes and shapes. A well-functioning local public economy will exhibit a fairly close match between the size of provision units and the size of the problems being addressed. Problems best addressed on a small scale will be handled by relatively small provision units, while problems best addressed on a larger scale will be matched to larger units.

Evidence of such matching can be found in the more highly differentiated local public economies. In St. Louis County, Missouri, for example, street service provision tends to be divided among three types of units: (1) small municipalities and subdivisions, (2) larger municipalities, and (3) county government.¹⁴ The first group provides for residential streets with priority attention to the control of traffic flow. The second group is more heterogeneous in its scope of provision, but often leaves residential streets to subdivisions, concentrating on streets with higher traffic volume. The third unit, county government, provides for a countywide arterial street system. Only the unincorporated portion of the county, where county government usually provides for all street services (except state highways), departs substantially from this three-fold division of responsibility. In the field of public education, regular elementary and secondary schools are maintained in St. Louis County by 23 separate

school districts, while special education and vocational-technical training are the province of a countywide special district.

Size Effects

Wide variation also exists in the size of provision units used to provide the same services. Some local public economies exhibit much less differentiation than others, or have developed unevenly. This raises the possibility that some provision responsibilities may not be well matched to existing provision units. One service area that has been carefully studied in this regard is police protection. Elinor Ostrom and colleagues have conducted extensive studies to determine the effect of jurisdictional size on citizen evaluations of police, among other measures of police performance. Their work consistently demonstrates that *smaller* units tend to be *more* responsive providers of police services. The research program included the study of police services in a number of metropolitan areas, consistently focusing upon delivery of services in relatively similar neighborhoods by police departments of varying size, ranging from a few part-time officers to more than 2,000 officers. As performance measures, the studies used victimization rates, whether citizens call police when victimized, speed of response when victimized, and general citizen evaluations of police. No study has found a large police department (over 350 officers) able to perform more effectively in delivering direct services to citizens in similar neighborhoods than smaller departments. Frequently, small departments deliver better service at lower cost.¹⁵

Similarly, William A. Niskanen and Mickey Levy found school district size (beyond a minimum level) to have a consistent negative effect on various measures of student achievement in California.¹⁶ Using a sample of 144 unified school districts (those providing both elementary and secondary education) with student populations of 2,000 or more, they found that students in larger districts tended to have lower scores on standardized achievement tests, when controlling for the effects of family poverty, minority background, and community wealth as these operate through median student IQ scores and expenditures per student.¹⁷ They also found, searching for an explanation of these findings, that larger districts tend to have lower teacher turnover and, for this reason (not because of salary schedules), higher teacher salaries and larger class sizes.

Both the police studies and the school district studies, however, measure the combined effect of provision and production characteristics. By and large, the studies look at provision units that choose to organize their own production units—police departments and schools. The results may be better interpreted as showing that smaller provision units, as long as they are sufficiently large to capture economies of scale,

are better able to organize and operate their own production units. School districts that were to contract out for the production of education with various producers might show much different results,¹⁸ as might municipalities contracting with neighborhood units to produce police patrol services.

Provision functions with respect to the *same* service need not be vested in a single unit. Provision for public education, for example, is usually shared in the United States between each state and its local school districts. The economic rationale for this sort of arrangement is the existence of externalities or spill-over effects. Even if a provision unit is organized to represent the interests of immediate beneficiaries—those who directly receive services—others outside the basic provision unit may be affected by the level of provision. Grants-in-aid that transfer funds from a more inclusive provision unit to less inclusive units give fiscal expression to the interests of others outside the basic provision unit. The basic unit, such as a local school district, however, must take the responsibility of arranging for production, if not organizing its own production unit.

Other possibilities exist for dividing and sharing the distinct functions of provision. Municipalities (or other more inclusive units) sometimes contract with neighborhood units to deliver services within their own neighborhoods. This arrangement may increase the accountability of producers to the neighborhood, not because of municipal provision activity, but because the neighborhood organization is able to function, formally or informally, to represent the interests of neighborhood residents—a provision function. Such an arrangement makes economic sense when economies of scale in production are limited and when the larger unit is willing and better able to finance provision (e.g., social services).

How Much Differentiation?

The amount of differentiation among provision units—the number and variety that can economically be created—is necessarily constrained, as discussed in *Chapter One*, by the costs of organizing and operating additional units. Multi-purpose units often make sense because they enable citizens to economize on the costs of organizing and operating multiple provision units; yet, as noted above, special-purpose units can be chosen for the same reason. The most efficient local public economy, on the provision side, is one that has an optimal number and mix of provision units, taking into account these costs.

In traditional terms, differentiation is “fragmentation.” An empirical literature focusing upon the fiscal effects of both “fragmentation” and overlapping jurisdictions has developed in recent years among public finance economists. Generally, this line of research has found *lower* levels of local government expenditure to be associated with *higher* levels of fragmenta-

tion and overlap, even when controlling for the level of community demand for local public goods and services. Cross-sectional research by Richard E. Wagner and Warren E. Weber,¹⁹ Thomas Dilorenzo,²⁰ and Mark Schneider²¹ all point consistently in the same direction. Dilorenzo employed a measure of "concentration," the degree to which tax or expenditure efforts are concentrated in the four largest jurisdictions (of whatever type) in a county for specific services, and found that concentration is associated with higher expenditure levels for the county as a whole. Schneider looked at fragmentation as measured more traditionally—number of municipalities per 100,000 residents in an urban county—and explicitly controlled for the level of municipal demand for local public goods and services. The latter is important; if a variety of provision units are more demand-responsive, this can lead either to greater or lesser government expenditure. Using a demand-side model, Schneider observed a negative sign between fragmentation and growth in expenditures over a five-year period.

It should be noted that Chicoine and Walzer²² did not replicate these findings in a study limited to 101 Illinois counties, excluding Cook County. Using traditional measures of fragmentation, they found no relationship to expenditures. They also found a positive relationship to expenditures using an indicator of fragmentation based upon industrial organization theory; they did not, however, provide a replication of Dilorenzo's industrial concentration model. Chicoine and Walzer also argue, however, that special districts have been created in Illinois as a response to fiscal limits on general-purpose local governments. If this is the case, the number of governmental units may in fact be suboptimal (i.e., too large).

With reference specifically to police services, Ostrom and Parks²³ found that the number of police jurisdictions per 100,000 population is strongly and negatively related to per capita expenditures for police in a nationwide sample of SMSAs. The greater the number of jurisdictions in an SMSA, the less they tend to spend for policing. Controlling for suburban respondents' service quality ratings (the great majority of jurisdictions being suburban) strengthens this relationship, indicating that the higher expenditures in less fragmented areas do not, in the minds of citizen-consumers, simply reflect higher service quality.

It seems fair to say that the weight of evidence at this time indicates that differentiated provision (fragmentation) is more likely than not to have favorable fiscal effects in efficiency terms. Missing from this research, however, is a consideration of intervening variables—mainly institutional—that might condition the effect of jurisdictional fragmentation on public expenditures and thus explain or specify this relationship. If there are institutional constraints that tend to inhibit the separation of provision from production

(e.g., legal restraints on intergovernmental contracting), "fragmentation" can be expected to be less efficient.²⁴

A more basic problem is the one raised by Chicoine and Walzer—that the degree of differentiation or fragmentation in a local public economy may be suboptimal due to constraints on local authority imposed by state rules. The extent of differentiation depends upon the rules that govern the establishment and operation of provision units. These rules can result in a suboptimal number (or the wrong mix) of provision units either in the sense that there are too few provision units or too many. Differentiation or fragmentation is neither an unqualified good nor an unqualified bad. Future research should take into account the rules by which local public economies are organized—at the level of governance—in order to determine the extent to which different local public economies are able to attain an optimal number and mix of provision units. Efficiency is expected to be associated with optimality, not a greater or lesser number of provision units per se.

Two studies have attempted to examine citizen perceptions of service quality as a function of fragmentation. James A. Christenson and Carolyn E. Sachs used statistical scaling techniques to collapse citizen evaluations of services into a single indicator.²⁵ Then, using counties in North Carolina as units of analysis, they found that county fragmentation is negatively related to service quality. Use of the scaling technique may, however, obscure important information, as may the aggregation of data by counties. Chicoine and Walzer used similar techniques to study Illinois counties (excluding Cook County) and found similar results using a single scale of service perception.²⁶ They point out that this result does not support, as a general proposition, the idea that smaller units of government are more responsive to citizen preferences. Aggregating citizen evaluations by counties, however, does not permit a direct examination of this hypothesis. Even in counties with a very large number of units, it is possible that most citizens reside in units that are comparatively large. When analyzing discrete services, Chicoine and Walzer found mixed results. Citizen perceptions of education and parks are favorably affected by fragmentation in their findings, while highways and libraries are unfavorably affected. They conclude that different patterns of organization may be appropriate for different types of services.

CONCLUSION

A theory of local public economies suggests that differentiated provision—jurisdictional fragmentation—is neither a wholesale virtue nor an unmitigated vice. Yet, while more empirical work remains to be done, the relative efficiency of maintaining a variety of provision units in a metropolitan area has been well estab-

lished. It would be incorrect, both theoretically and empirically, to claim that more differentiation or fragmentation is always better than less. Research clearly

shows, however, that more differentiated provision can be, and frequently is, better than less.

NOTES

- ¹In Virginia, counties and municipalities have separate jurisdiction. The same is also true in special cases, such as the separation of St. Louis City from St. Louis County and Baltimore City from Baltimore County.
- ²In New England, "towns" are organized much like townships, but tend to function as municipalities.
- ³ACIR, *The Problem of Special Districts in American Government*, A-22, May 1964, pp. 73-84.
- ⁴Donald Foster Stetzer, *Special Districts in Cook County: Toward a Geography of Local Government*, Research Paper No. 169 (Chicago: The University of Chicago, Department of Geography, 1975).
- ⁵Robert B. Hawkins, Jr., *Self-Government by District: Myth and Reality* (Stanford: Hoover Institution Press, 1976), p. 116.
- ⁶David L. Chicoine and Norman Walzer, *Governmental Structure and Local Public Finance* (Boston: Oelgeschlager, Gunn & Hain, 1985), p. 80. One possible difficulty with this finding is that home-rule authority is automatically extended to municipalities larger than 25,000 population and to others only by local referendum. Thus, to some extent, a negative correlation between the percentage of citizens living in home-rule municipalities and number of governments per 10,000 citizens may be expected.
- ⁷This was apparently accomplished in Illinois when the Illinois Constitution of 1970 extended fiscal home-rule to municipalities.
- ⁸Two books have become classic statements on this subject: Milton Kotler, *Neighborhood Government: The Local Foundations of Political Life* (Indianapolis: Bobbs-Merrill, 1969) and Alan Altshuler, *Community Control: The Black Demand for Participation in Large American Cities* (Pegasus, 1970).
- ⁹See a study published by the Joint Economic Committee of Congress, "Stimulating Community Enterprise: A Response to Fiscal Strains in the Public Sector." December 31, 1984.
- ¹⁰Although an ACIR study was unable to generate a complete census of subdivisions in the county, 27 municipalities report a total of 427 street-providing subdivisions. See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.
- ¹¹ACIR, *Fiscal Balance in the American Federal System*, Vol. 2, *Metropolitan Fiscal Disparities*, A-31, October 1967, pp. 16-17.
- ¹²ACIR, *The States and Distressed Communities: The Final Report*, A-101, November 1985, p. 245.
- ¹³More research on the prospects and pitfalls of neighborhood incorporation is needed. Experience with neighborhood advisory commissions and voluntary neighborhood associations in large central cities can usefully be compared to small municipalities and organized subdivisions in metropolitan counties. The experience of neighborhood development corporations can also be examined for possible lessons in organization.
- ¹⁴See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.
- ¹⁵A useful summary of much of this research program can be found in Elinor Ostrom, "Size and Performance in a Federal System," *Publius* 6 (Spring 1976), pp. 33-73. Also see Elinor Ostrom and Gordon P. Whitaker, "Does Local Community Control of Police Make a Difference? Some Preliminary Findings," *American Journal of Political Science* 17 (February 1973), pp. 48-76; Elinor Ostrom and Roger B. Parks, "Suburban Police Departments: Too Many and Too Small?" In Louis H. Masotti and Jeffrey K. Hadden (eds.), "The Urbanization of the Suburbs," *Urban Affairs Annual Reviews* 7 (Beverly Hills: Sage Publications, 1973), pp. 303-402; and Elinor Ostrom and Gordon P. Whitaker, "Community Control and Governmental Responsiveness: The Case of Police in Black Neighborhoods," in Terry N. Clark (ed.), *Urban Policy Analysis: Directions for Future Research*, *Urban Affairs Annual Reviews* 8 (Beverly Hills: Sage Publications, 1974), pp. 303-34.
- ¹⁶William Niskanen and Mickey Levy, "Cities and Schools: A Case for Community Government in California," Working Paper No. 14, (Berkeley: Graduate School of Public Policy, University of California, 1974).
- ¹⁷The procedure used was two-stage least squares regression.
- ¹⁸Robert L. Bish and Vincent Ostrom briefly discuss performance contracting in public education in *Understanding Urban Government: Metropolitan Reform Reconsidered* (Washington, DC: American Enterprise Institute, 1973), pp. 38-39.
- ¹⁹Richard E. Wagner and Warren E. Weber, "Competition, Monopoly, and the Organization of Government in Metropolitan Areas," *Journal of Law and Economics* 18 (December 1975), pp. 661-84.
- ²⁰Thomas Dilorenzo, "Economic Competition and Political Competition: An Empirical Note," *Public Choice* 40 (1983), pp. 203-09.
- ²¹Mark Schneider, "Fragmentation and the Growth of Government," *Public Choice* 48 (1986), pp. 255-63.
- ²²David L. Chicoine and Norman Walzer, *op. cit.*
- ²³Elinor Ostrom and Roger B. Parks, *op. cit.*, p. 390.
- ²⁴Future research that may examine this relationship by specifying more elaborate multi-equation models, in contrast to the single-equation models cited in this chapter.
- ²⁵James A. Christenson and Carolyn E. Sachs, "The Impact of Government Size and Number of Administrative Units on the Quality of Public Services," *Administrative Science Quarterly* 25 (March 1980), pp. 89-101.
- ²⁶Chicoine and Walzer, *op. cit.*

Differentiating Production

INTRODUCTION

Organizing the production side of a local public economy is likely to involve a mixture of production arrangements responsive to diverse criteria having to do with the division of labor, economies (and diseconomies) of scale, and competition. Each provision unit is confronted with the question of how to arrange for production. It is impossible to know a priori, without specific time-and-place information, what set of arrangements will be optimal. As discussed in *Chapter Two*, distinguishing provision and production does not necessarily imply separation of production from provision in discrete circumstances. Most provision units, except for small neighborhood units, choose to organize some aspect of production for themselves. The differentiation of the production side of a local public economy is a result of provision units making different choices among alternative production modes.

The focus of this chapter is not upon the organization of production within production units, but upon the distribution of production tasks among independent units within service areas—police, fire, education, and so forth. Differentiated production breaks services down into service components. Production criteria often vary sharply among these different components. If one component of a service is labor intensive, while another is capital intensive, the economies of scale are almost sure to be different. As a result, it is often more efficient to organize production by relying upon independent units to produce various components of a service.

SEPARABLE PRODUCTION TASKS

Service Components

Services can be usefully distinguished as “direct” and “indirect” services. Direct services are delivered directly to citizen consumers, whereas indirect services (also called auxiliary or support services) are delivered to those who produce direct services. Indirect services are frequently organized separately from direct serv-

ices. For example, in police work, indirect services, such as dispatching and forensics, are often produced separately from police patrol and investigative work in the field. The producers of direct services become consumers of indirect services. Frequently, a provision unit will organize its own production unit in-house to produce direct services while relying on outside producers to supply indirect services. Differentiation along these lines tends to occur when indirect services require specially trained personnel and/or benefit from other economies of scale.

Differentiation also occurs, however, *among* both direct services and indirect services. Routine police patrol, for example, can be separated from the investigation of serious crimes. Like some support services, investigation is more apt to require specialized personnel and to exhibit economies of scale. Different support services, moreover, also may be characterized by different scale economies and depend upon quite different areas of expertise. Police dispatching benefits from a larger scale of production than police patrol, but crime labs benefit from a still larger scale than dispatch.

Transaction Costs: A Constraint

The extent of differentiation is constrained by the transaction costs of coordinating production between and among different units. Just as the costs of organizing and operating additional provision units constrains differentiation on the provision side, the use of multiple production units is constrained by the costs of negotiating and maintaining agreements among different organizations and coordinating the working relationships of individuals who are employed by different organizations. These coordination costs necessarily limit the number and variety of production units in any local public economy.

Some service components may be easier to sort out and produce separately than others. Indirect services may be more easily differentiated from direct services, while direct-service components are less easily differentiated from other direct-service compo-

nents. It may be easier, for example, for a police department to contract out for crime-lab services than for criminal investigation. The differentiation of indirect services establishes a relationship between a producing unit and a consuming unit. The producers of different support services may not need to coordinate at all. The differentiation of direct services, by contrast, must establish a relationship between two or more producing units both of which deliver services to the same (or an overlapping) clientele. Closer working relationships may be required. Indirect services may be more easily coordinated with direct services by means of a contractual relationship, stipulating terms and conditions in advance, while the coordination of direct services may depend upon forms of cooperation that require mutual accommodation among individuals on a day-to-day basis. These different types of coordination operate as different constraints on the structure of the production side, affecting the number and variety of production units that are organized.

The structure of the provision side also affects the differentiation of production. Large, heterogeneous provision units, such as big cities incorporated as a single municipality, tend to organize direct services in-house, while contracting out selectively for indirect service components. They substitute the transaction costs internal to a bureaucratic hierarchy for those incurred in coordinating across independent units. Coordination costs within a single organization vary with the size of the organization and the heterogeneity of the tasks undertaken by the organization. Whether internal transaction costs tend to be less, or more, than the costs of interorganizational coordination undertaken on the same scale is a complex and unanswered empirical question.

Whatever the answer, the differences in transaction costs may be dominated by differences in economies of scale. Large, undifferentiated production units combine at a single scale of organization the production of different service components that may exhibit sharply variant economies of scale.¹ A local public economy characterized by a large number of relatively small provision units is more likely to develop a highly differentiated production side. Even in this case, the extent of differentiation is nevertheless constrained. Producers of direct services that find it more advantageous to produce an indirect service component in-house can be expected to do so. In the long term, small provision units that strongly prefer in-house production will merge with other units, if possible, to achieve a favorable economy of scale and avoid excessive transaction costs.

THE EXTENT OF DIFFERENTIATION AND ITS EFFECTS

Depending on specific circumstances, different production components can be produced internally,

contracted out, or produced jointly with another jurisdiction. Elinor Ostrom, Roger B. Parks, and Gordon Whitaker² have studied the organization of police production in metropolitan areas throughout the United States. In general, they find that patterns of organization differ among the components of police service in predictable ways. Small and medium-size police departments tend to be patrol specialists and to place a greater percentage of their sworn officers on the street than do larger departments. Indirect or auxiliary services tend to be produced by larger units and coordinated, in a variety of ways, with the production of direct services by smaller units. The result is a system of police production—based upon rational criteria—that emerges from the independent choices of a variety of provision and production units.

In St. Louis County, Missouri, where an ACIR study found similar patterns in a number of service areas, public education offers a prime example of differentiating the production of indirect services. Twenty-three school districts organize for themselves the production of most direct services to students, while relying upon external producers for purchasing, data processing and/or computer services, in-service training, and selected components of direct service, such as music festivals and summer programs. The organizational arrangements used are joint production units—consortia—maintained on a cooperative basis. The most highly coordinated production arrangement, however, is in the direct service area. The Special Education District of St. Louis County is a countywide provision unit established to provide and produce education for the handicapped (as well as vocational-technical education). Only students with the most serious disabilities, however, are educated in separate schools provided by the Special Education District. The majority of disabled students receive services in the school they would ordinarily attend within their regular school district. Teachers employed by the special district are assigned to regular district schools. Students with disabilities spend a portion of their time with special teachers and the remainder of their time in ordinary classrooms. Two different production units thus coordinate service delivery to the same clientele using the same facilities.

The most dramatic evidence of diverse economies of scale in different components of policing is reported by Parks.³ In a study of 76 metropolitan areas in the U.S., Parks found that areas with a greater number of independent police producers tend to have higher efficiency ratings for police patrol service, while areas subject to greater dominance by a single supplier tend to have higher efficiency ratings for homicide investigation and radio communications. A highly differentiated system for producing police services is able to capture both types of efficiency.

Police services in the St. Louis area also offer an illustration of highly differentiated production. Sixty-four full-time police departments serve the City and County of St. Louis, producing, at a minimum, local patrol and response to calls, local traffic control, and investigations of most reported crimes. Yet only 45 percent produce their own dispatch services. Smaller departments rely upon a variety of methods to obtain dispatch services, including a joint dispatch center (supplying nine departments) and contract arrangements with larger municipal departments and with county police. Within St. Louis County, county police are the principal producer of crime lab services, including crime scene investigation, supplied to municipalities without charge.⁴ A Major Case Squad drawing investigative officers from throughout the metropolitan area can be deployed upon the request of the police chief in whose jurisdiction a serious crime occurs. St. Louis City and County are jointly served by a computerized call routing (911) system. All local police officers receive entry-level training through a joint police academy. A combination of joint service production and intergovernmental contracting is used to differentiate and coordinate the production of those service components most likely to exhibit significant economies of scale.

One of the traditional concerns about multiple production units concerns the possibility of duplication of effort. Ostrom, Parks, and Whitaker diligently searched for duplication in the organization of police protection, but found very little. The ACIR study of four service areas in St. Louis also found little duplication. Specialization, not duplication, is the prevailing characteristic of production systems that rely upon multiple units. Accompanied by relatively high levels of coordination, at least as needed, specialized production by independent production units does not result in much identifiable duplication. The mutual interest in avoiding duplication among closely related production units is apparently sufficient to minimize its occurrence.

In addition to coordination, production units specialize by means of "alternation" (i.e., dividing up production responsibility on the basis of time, space, or clientele). In this manner, production units with somewhat overlapping responsibilities can avoid duplicating one another. The overlapping jurisdictions of county and municipal police in St. Louis County, for example, does not usually lead to duplicative patrol patterns. Overlap does not necessarily lead to duplication. The tendency of multiple and/or overlapping

units to avoid duplication of effort may account for the failure of consolidation efforts to result in demonstrable cost savings, as often predicted.⁵

"Duplication" is a term sometimes used more loosely to refer to any performance of identical activities by multiple units. The existence of multiple production units, such as the large number of small police departments in St. Louis County, should not be interpreted as prima facie evidence of duplication that is uneconomical, even if these departments engage in identical activities. Only those activities that can benefit from the economies of larger scale production are duplicative in the sense that the same services could be supplied by fewer producers at less cost. "Duplication" in this usage is simply another name for an economies of scale argument.

CONCLUSION

The existence of a large number of independent producers of direct services (e.g., a large number of small police departments or fire departments) in a local public economy does not necessarily imply a lack of coordination or an inability to capture limited economies of scale in the production of particular service components. Producers in such areas tend to develop an *organizational overlay* of coordinated production arrangements.

A study of Cook County, Illinois, for example, characterizes the county outside Chicago as a "tangled mass of park districts, fire protection districts, library districts, and others. . . superimposed over an expanse of zig-zag municipal boundaries." The author goes on to comment:

With such extensive fragmentation, one might wonder how Cook County operates at all. In practice, the situation is much better than the territorial arrangement might suggest. Through time the districts have developed a series of understandings, agreements, and working arrangements with other governments and with each other. Local officials believe that this horizontal integration leads to a superior type of government.⁶

This result is one to be expected where the benefits of coordination exceed the costs. The evidence suggests that coordination costs among independent production units are not so high as to preclude important elements of mutual coordination. Service production can therefore be differentiated to a considerable degree without encountering prohibitive costs of coordination.

NOTES

¹One question that arises is why these large units tend not to contract out for the delivery of direct-service components. One reason may be the absence of available contractors. Large units tend to drive out potential competitors. Small

units tend to have larger, overlapping units to draw upon as contractors, while large, exclusive jurisdictions do not have this readily available source of supply. The development of neighborhood units in large cities would have the potential for altering this constraint. When large units do contract with small units to deliver direct services, it is usually in the social-

service area, where a relatively large number of private voluntary organizations exist as potential contractors.

²Elinor Ostrom, Roger B. Parks, and Gordon P. Whitaker, *Patterns of Metropolitan Policing* (Cambridge, Mass.: Ballinger Publishing Co., 1978).

³Roger B. Parks, "Metropolitan Structure and Systematic Performance: The Case of Police Service Delivery," in Kenneth Hanf and Theo A.J. Toonen (eds.), *Policy Implementation in Federal and Unitary States* (Dordrecht, The Netherlands: Martinus Nijhoff Publishers), pp. 161-91.

⁴Recent reductions in revenues have motivated the county police to consider supplying some services on a fee-for-service basis. In effect, the county would then shift from a provision

role to a narrower production role. Such a move could favorably affect fiscal equivalence.

⁵The most detailed study to date undertook a time-series analysis of the 1967 merger of Jacksonville with Duval County, Florida, and concluded that consolidation failed to reduce either tax or expenditure levels. J. Edwin Benton and Darwin Gamble, "City/County Consolidation and Economies of Scale: Evidence from a Time-Series Analysis in Jacksonville, Florida," *Social Science Quarterly* 65 (March 1984), pp. 190-98.

⁶Donald Foster Stetzer, *Special Districts in Cook County: Toward a Geography of Local Government*, Research Paper No. 169 (Chicago: The University of Chicago, Department of Geography, 1975), p. 151.

Governance, Change, and Civil Community

INTRODUCTION

The governance of a local public economy is separable from both provision and production activity. Governance requires a capability to make and enforce rules that apply areawide. This capability is independent of the provision and production of services on an areawide basis. Metropolitan “governance” therefore does not depend upon the establishment of a “metropolitan government” that has provision and production responsibilities. The governance of a local public economy determines its basic structure—essentially, the number and variety of provision units. The necessary rules of governance pertain generally to the creation and dissolution of local governments, the powers that each unit of government may exercise, and relationships among units. In most local public economies in the United States, this framework of rules is supplied largely through state constitutions and laws, not by metropolitan or regional governments.

Somewhat like private economies, however, local public economies vary in the degree to which they are regulated by law. Some rules, analogous to property rights, are part of an essential framework of governance. These would include rules of local-government formation and boundary adjustment, which generally allocate substantial authority to local citizens. Others, more akin to price regulation, constrain the degree to which local citizens can choose to organize and tax themselves to provide local services.

When citizens first petition and then vote to incorporate a new municipality, or vote to approve an annexation or consolidation in a referendum, or give their consent at the ballot-box to a tax-rate increase, they are participating directly in the governance of their local public economy. Citizens are then making the basic structural choices that organize the provision side. In this fashion, with variations from state to

state, citizen choice is frequently the basic method of governance in American local public economies.

CHOOSING BOUNDARIES

One of the basic governance issues affecting local public economies is the question of who is best suited to determine the boundaries of provision units. Of course, the issue of how large or how small any given unit should be is closely related to the question of how many units there ought to be in a given area. Boundaries determine who is and is not included within a jurisdiction. Choosing boundaries is thus intimately tied up with identifying a community of interest and matching communities with collective organizations. Satisfying the basic criteria of provision, such as preference expression, fiscal equivalence, and accountability (discussed in *Chapter One*), depends upon having reasonably correct boundaries. But how does one know what is correct? Can the correct boundaries be objectively determined? Or is boundary choice an inherently subjective decision-making process?

One objective factor to be considered in choosing appropriate boundaries is the scale of effects associated with the consumption of a good or service—namely, the spatial or geographical extent of an area directly affected when a good or service is supplied (or not supplied). The scale of effects associated with some goods and services, such as the supply of residential (nonarterial) streets or routine police (or security) patrol, is quite small. Other goods and services, such as arterial streets and police crime solving, have direct effects that are broader in extent, affecting larger communities of interest.

Extent or scale of effects, however, is not a determining factor. At most this criterion would establish the need for a minimum size jurisdiction for specific goods and services. If individual households hire a security patrol, for example, neighbors who do not contribute will nevertheless derive a benefit. They can act

as “free-riders.” Private provision, therefore, cannot fully aggregate demand for this service, but the exact scale on which demand is best aggregated, collectively, remains indeterminate. Boundaries clearly can be too small.¹ On the other hand, boundaries larger than required by this criterion may still be optimal.

The upper limit on the appropriate size of a provision unit is much more ambiguous. The determining factors involve the variability of citizen preferences and the way that citizens with similar preferences cluster geographically. At issue is the extent to which a relatively homogeneous good or service, provided at a given level, can satisfy a geographically defined community of interest. Some “neighborhoods,” defined as a community with a distinct set of interests and preferences compared to surrounding communities, may be quite small; others may be comparatively much larger.

Fiscal equivalence and accountability may in some circumstances suggest a more circumscribed set of boundaries than suggested by preference homogeneity. A large community with homogeneous preferences may still experience difficulties in maintaining accountability to citizens if the aggregate number of citizens increases beyond some point. The same circumstance may create incentives for some neighborhoods to raid the central treasury at the expense of others, leading to problems with fiscal equivalence. In this way, the criteria of fiscal equivalence and accountability can modify considerations based solely upon homogeneity of preferences.

To make these determinations requires knowledge of the preferences (for goods and services) of individuals in the communities involved. Local citizens, therefore, can best make the tradeoffs required between preference satisfaction, on the one hand, and the transaction costs associated with organizing and operating additional provision units, on the other. These costs, as discussed in *Chapter One*, constrain the degree to which citizens choose to create additional provision units, and include the costs of citizen participation necessary to articulate preferences and secure accountability from officials.

Appropriate boundary determinations (from the standpoint of provision-side efficiency) depend upon a set of rules that allow citizens, collectively, to make those tradeoffs. Apart from knowing the preferences of individuals in discrete communities for goods and services and how these individuals would trade off preference satisfaction (or fiscal equivalence or accountability) against the costs of organizing and operating additional units, there is no fully objective method for determining the optimal size of a provision unit.

The minimum-size jurisdiction appropriate to organize provision varies substantially among different goods and services. Some types of goods and services can be provided on a “neighborhood” scale. Other

types obviously serve a broader community of interest, such as arterial streets, some types of law enforcement (e.g., investigation of major crimes), public transportation, airports, and some types of park and recreation services. This implies that a set of provision units where smaller units are nested—like “Chinese boxes”—inside larger units will tend to be optimal.² One cannot determine a priori how much “nesting”—how many boxes inside boxes—there ought to be within a particular local area. To decide in advance that a “two-tier” arrangement, for example, is appropriate does not take into account the potential variety in organization on both the provision side and the production side—the diverse scale considerations that apply to service production in addition to the potential diversity of community preferences. The amount of nesting or territorial overlap that is appropriate can vary within, as well as among, local public economies. As with the choice of boundaries, citizens are the preferred decision makers.

No one can determine the “correct” or “best” pattern of organization for a local public economy *a priori*. This includes organization of both the provision side and the production side. Instead of trying to determine an ideal structure of metropolitan or regional organization, a normative analysis of local public economies should focus on the “rules of the game” in order to help individuals and communities to order their relationships with one another in more productive ways.

Rules governing incorporation, consolidation, and annexation, as well as disincorporation and de-annexation, define the decision-making processes by which boundary determinations are made and altered.⁴ Comprehensive reform proposals, such as city-county consolidations, that would greatly diminish the number of independent provision units in a metropolitan area, however, are much more often defeated than accepted by local voters.⁴ New units are established with some frequency; marginal adjustments in the boundaries of existing units are also fairly common; but comprehensive changes that alter a number of boundaries at once are extremely rare.

One cannot explain the existing structure of a local public economy without reference to these “rules of the game.” It may seem odd, for example, that relatively homogeneous suburbs generally exhibit a greater number and variety of provision units than more diverse central cities.⁵ One explanation lies in the rules used to organize provision units. Suburban growth begins in an unincorporated area, but the rules of incorporation and annexation frequently allow diverse communities to make independent boundary determinations. Most central cities were a product of a much different decision-making process; often their boundaries were determined by state legislatures, or the rules then in place allowed annexation without

citizen consent. Also, as discussed in *Chapter Three*, once a city is incorporated, the rules tend to preclude or inhibit the development of smaller, nested provision units—neighborhood governments—within its boundaries.

DYNAMICS AND ADAPTIVE CAPABILITIES

Once provision units are in place, residents can take the characteristics of different communities into account in making relocation decisions. The ability of residents to sort themselves among diverse communities on the basis of individual preferences is, to some extent, an important dynamic in a local public economy. Charles M. Tiebout demonstrated that residential mobility has the potential, subject to various constraints, to achieve an optimal allocation of resources in the provision of local public goods and services.⁶ Historical choices of boundaries therefore tend to be self-perpetuating. Once a set of jurisdictional boundaries has been determined, individuals make locational decisions accordingly, and tend over time to increase the efficiency of existing boundaries (i.e., the ability of a set of boundaries to satisfy a diverse set of individual preferences). The greater the number of choices available, the greater the effect of residential sorting. Yet existing boundaries can also become obsolete, as preferences become either more homogeneous or more heterogeneous or as organizational/operating costs either increase or decrease, and the resulting tradeoffs change.

To maintain an efficient local public economy over time requires structural flexibility and the continued availability of alternative arrangements in both provision and production. A local public economy is not static. The sources of change include (1) changing citizen preferences for goods and services, (2) population growth (or loss), and (3) developing technology. The latter includes, not only "hardware," but also the softer kind of technology associated with "know-how." Economic change in local public economies is a vector of these forces. Efficiency in both provision and production depends upon institutional adaptation to the complex change vector.

Adaptation depends upon the availability of alternatives and the development of new alternatives. On the production side, the availability of alternatives increases producer competition. In a local public economy, however, the competition is not simply among private vendors, but between public and private vendors, as well as among public vendors of services. If competition among private suppliers is constrained (which can be the case if there are substantial economies of scale in production), it may be important to maintain the option of public production even if private production seems to be a better alternative at a given point in time. Maintaining the public option may

mean, in turn, choosing not to contract out everything that could be contracted out. Maintaining a competitive environment could also mean, for a large provision unit, choosing to divide up the production of some service among different contractors rather than contracting with a single vendor. Where there are a number of small provision units, on the other hand, competition on the production side tends to be self-generating.

The development of new production arrangements is a key to adaptation and productivity improvement—also, in turn, a source of change. Development of this sort depends upon entrepreneurship,⁷ both public and private. In either case, a necessary condition of entrepreneurship is *initiative*. Entrepreneurial potential is measured by the number of possible *sources* of initiative in a local public economy. Counting the number of police chiefs, fire chiefs, directors of public works, city administrators or managers, and school superintendents will yield a crude measure of the potential for public entrepreneurship in each of these service areas. The experience of St. Louis County, Missouri, offers numerous illustrations. Entrepreneurial activity in St. Louis County, given the high potential, is on-going, and has resulted in many successful joint production efforts, including most recently an educational consortium for the application of computer technology in schools and a number of drug enforcement units maintained jointly by several groups of municipal police departments.⁸

The continued availability and development of alternatives must extend also to the provision side. Provision alternatives are sustained in several ways. One way consists of creating "nested" provision units with somewhat overlapping authority to act. The nesting of municipalities within a county is a common example. As general-purpose governments, municipalities and counties generally have overlapping authority to act even while sorting out their responsibilities on an informal basis. Provision units, like all human creations, are imperfect. If one unit fails to respond to constituents, a jurisdiction with overlapping authority is available. A form of political competition exists between officials, such as mayors and county commissioners, whose jurisdictions overlap, allowing "voter sovereignty" (analogous to consumer sovereignty) to exercise a choice between them.

Provision alternatives can also be maintained by means of special purpose governments (discussed in *Chapter Three*). The formation of a special district frequently is an alternative to enlarging the boundaries of an existing provision unit, merging two or more provision units into one, or organizing a general-purpose government. The availability of procedures for creating special districts thus significantly enlarges the range of provision alternatives. Moreover, it is always possi-

ble to seek the enactment of new rules that allow the creation of new types of units.

Some local public economies, or more likely some portions of a local public economy, do not have a wide range of alternative arrangements for provision. The struggle to create neighborhood governments in large incorporated cities proceeds from a lack of readily available institutional alternatives for disaggregating some of the collective choices made in many large urban centers. In other areas, however, available alternatives are not used, perhaps because of a fairly large scale homogeneous demand for a given service level. Frequently this may be a minimal service level supplied through a county government.⁹

In general, the greater the number of available provision units, either in place or available to be created as an adjunct to existing arrangements, the more likely citizens will be able to obtain satisfaction of their preferences. This is true for two reasons. One is the ability of residents to sort themselves among organized provision units on the basis of individual preferences for local public goods and services. The other reason depends on the ability of citizens to sort and resort responsibilities among both existing and potential units—to transfer functions. Functional transfer is an important and flexible tool of local governance, but its use depends upon the potential variety of provision units that citizens can create and empower. Municipal incorporation or special district formation not only creates a new unit, but transfers a limited set of functions from existing units to the newly created municipality. The same ability to incorporate neighborhoods within large municipalities, and to transfer some set of functions from municipality to neighborhood, is generally not available.

Recent developments in Los Angeles, Milwaukee, and Boston have underscored the difficulties that affect efforts to differentiate provision within large central cities. Residents of a high-crime area just south of downtown Los Angeles considered the creation of a special taxing district in 1987, to enhance police protection. The proposal was defeated by voters after city hall agreed to increase the number of police officers deployed in the area.¹⁰ Citizens in Milwaukee are actively considering the creation of a separate school district, centered upon a single high school, that would serve a portion of the city's predominantly black population. Such a change would require new state legislation.¹¹ In 1986, voters in Boston defeated a proposal by residents of a predominantly black, 12.5 square-mile area to secede in order to form a separate municipality of Mandela, a name selected in honor of the jailed black South African leader Nelson Mandela.¹² To allow citizens of Mandela to incorporate without securing approval from voters in the rest of Boston would also require new state legislation. Solving basic problems of service provision in the nation's

major urban centers may require renewed attention to the capabilities of citizens to make boundary changes and choose their local government units.

THE LIMITS OF INTERLOCAL COMPETITION

In addition to the governance supplied by rules, local public economies are governed in part by competitive dynamics among provision and production units. Distinguishing provision and production is essential to understanding the effect of competition on local governments. Much of the advantage derived from *separating* production from provision is the competition generated among alternative *producers*. This is especially true if private sector alternatives are included. Competition among producers, however, should be distinguished from the much more limited competition that may exist among provision units.

The ability of residents to choose among a number of municipalities does not necessarily or automatically increase competitive pressures upon local government officials as producers of services. This is because the taxable property of ordinary residents—mostly real estate—is not highly mobile. Residents can exit from a local jurisdiction, but they cannot ordinarily carry their house and other improvements with them. This severely limits the immediate impact of exit upon local governments. Commercial and industrial residents, on the other hand, have taxable property that is much more mobile. A local firm that threatens to move out of a community has leverage supplied by the mobility of its capital. Similarly, local jurisdictions that have a great deal of undeveloped land and seek to grow may compete for new residents on the basis of their tax-service packages.

In general, competition among provision units may increase the leverage of commercial and industrial firms with local governments, but not significantly enhance the power of individual residents. Competition thus limits the extent to which local provision units can rely upon commercial and industrial property for local revenue. This is a significant constraint insofar as commercial and industrial property is a major source of financial support for local government. Renewed interest in the encouragement of economic development has perhaps heightened this sort of competitive pressure.

For ordinary citizens, however, what Albert O. Hirschman has called "exit"—the ability to "vote with your feet"—is not a substitute for political participation, what Hirschman calls "voice."¹³ The governance of a local public economy depends upon rules that assign important collective decision-making roles to citizens. Increasingly, these citizen roles include approval of tax-rate increases by popular referendum. Competition among local governments cannot be expected to constrain the tax-prices of services to ordinary resi-

dents apart from the ability of citizens to impose a *political* constraint on officials.

This conclusion does not imply that “voting with your feet” is an insignificant option. As discussed above in this chapter, the process whereby citizens sort themselves over time among provision units increases efficiency by grouping citizens with like preferences together. This increases the ability of collective decision making to reflect the diverse preferences of individuals. This sorting process, however, does not automatically constrain local officials, as producers of services, to be more efficient.

LOCAL GOVERNMENT CONSTITUTIONS

The rules used to organize the governance of a local public economy should be viewed as a single configuration—what might be called a “local government constitution.” Two levels of constitutional choice are entailed: (1) an *enabling* level that determines a set of rules that local citizens can use to create and modify local governments and (2) a *chartering* level that determines the specific charter of a unit of local government as citizens act to bring it into existence in accordance with enabling rules. The enabling rules can be sorted into four types: (1) rules of association—those that establish processes, such as municipal incorporation, that enable local citizens to create municipalities or other units endowed with certain governmental powers; (2) boundary adjustment rules—to enable local citizens and officials to alter the boundaries of existing units; (3) fiscal rules—those that determine the revenue raising authority of various local units; and (4) contracting rules—to enable local units to enter into a variety of mutually agreeable relationships with one another and with private firms.¹⁴

One of the basic issues pertaining to local government constitutions is the distribution of authority between the “enabling” level and the “chartering” level. The term “home rule” is traditionally applied to local government arrangements that reserve a good deal of authority to the chartering level. The ability of local residents to frame their own charter when choosing to constitute a local government unit is fundamental to home rule,¹⁵ but the degree to which states permit an autonomous determination of the powers of local government by local citizens varies widely.¹⁶ Of special interest, usually, is the issue of *fiscal* home rule, namely, the degree to which local taxing authority is derived from local charters rather than state law. Some state restrictions on local officials, however, constitute an empowerment of local citizens, such as those requiring citizen consent to tax-rate increases in local referenda. Home rule is most commonly applied to municipalities (often of a minimum size), less frequently to counties and townships, and seldom to special districts. The latter, when created by citizens, may

exercise only those powers defined as a matter of state law.¹⁷

The *full configuration* of enabling rules—not simply the rules for each type of local unit taken separately—affects the pattern of provision and production units that citizens and officials choose to establish. For example, as discussed in *Chapter Three*, the creation of special districts in some states may derive in part from restrictive tax and debt ceilings imposed by state law on local general-purpose governments. If this results, perversely, in a greater number of special districts than citizens would otherwise choose to create, the effect is inefficient. This result, however, is a consequence of the fiscal rules that pertain to general-purpose governments, not the rules of association that pertain to the formation of special districts.

Incorporation and annexation rules, to take another example, are closely related. If state law allows unilateral annexation by municipalities, without obtaining separate concurrence from voters in the area to be annexed, this may encourage “defensive” incorporation, to avoid annexation, by communities that otherwise would not choose to incorporate. On the other hand, some communities confronted by a more encompassing incorporation effort, may seek annexation by an adjacent municipality as a preferred alternative. This strategy could indicate that the collective decision-rule for incorporation is too low, allowing some parts of a proposed municipality to impose unwanted incorporation on other communities within the proposed set of boundaries. Incorporation rules vary (both among states and among types of municipalities within states) from a simple plurality of those voting to extraordinary majorities.

Consolidation rules also vary. ACIR recommends that concurrent simple majorities be required of two (or more) municipalities seeking to merge.¹⁸ The rules may be framed differently, however, so that only a single majority throughout the proposed jurisdiction is required. Such a rule could, in effect, allow a larger municipality to annex a smaller one without its consent. In Missouri, site of a recent ACIR study in St. Louis County, consolidation rules vary between municipalities and school districts.¹⁹ While concurrent majorities are required for municipalities, school districts can be reorganized on the basis of a single vote across the proposed district. School district reorganization can also be proposed by county officials, rather than initiated by citizen petition. The obvious intent of such rules is to facilitate school district consolidation, which is by far, nationwide, the most successful consolidation movement undertaken to date.

Beginning in 1961, ACIR has encouraged states to enact rules that would limit the number of new incorporations.²⁰ Eventually, this prescription was broadened to include the establishment of new local governments of all types, including both general-purpose and

special-purpose governments, as well as the dissolution of local governments considered to be nonviable. The principal institutional mechanism recommended to achieve this end was a *boundary review agency*, established on either a statewide or local areawide basis. The broad purpose of such an agency is to control incorporations and annexations both in order to assure an orderly pattern of service delivery and to promote a better match between fiscal capacity and service needs. The Commission also proposed that the same agency be used to enjoin intergovernmental contracting among local jurisdictions when this practice is seen to split the tax base without compensating advantages.²¹

Ten states, all located in the midwestern or western United States, established some sort of state or local boundary review agency over a ten-year period beginning in 1959.²² A 1968 review concluded that all operating boundary review agencies had successfully reduced the "proliferation" of local governments.²³ Neither growth nor reduction in the number of governments, however, provides evidence that any given number is suboptimal. An Illinois study cautions against "universal application of policies, such as boundary commissions, to generally limit or restrict use of single-purpose special districts."²⁴

Fully objective criteria for determining the boundaries of a provision unit cannot be established. Instead, it is necessary to consult the preferences of citizen-consumers in the area in question. An appropriate set of boundaries depends critically on how citizen preferences are spatially distributed. Frequently, the establishment of a new unit is an alternative to expanding the boundaries of an existing unit. If so, officials in existing units may wish to constrain incorporations in order to reduce competitive pressures on their own governments. Findings from an empirical study of Local Agency Formation Commissions (LAFCOs) in California support this interpretation.²⁵ Boundary review agencies may come to represent the interests of existing local government officials instead of the preferences of citizen-consumers seeking to organize themselves to provide public services. In 1987, the ACIR therefore recommended that states exercise caution in the creation and use of boundary review agencies that have the capability of vetoing incorporations and annexations that are, or would be, approved by local citizens in popular referenda.

Among the states that have established boundary review agencies, however, there is considerable variety in their design. No comprehensive study of these agencies, comparing experience across a variety of institutional settings, has been done. Boundary review agencies that increase the amount of relevant information available to local citizens, or aid in the resolution of conflicts, may perform a useful role. ACIR intends

to conduct additional research in order to assess the use of this institutional mechanism.

FRAGMENTATION AND CIVIL COMMUNITY IN METROPOLITAN AREAS

Although the term ordinarily used to describe a differentiated local public economy is "fragmentation," a highly differentiated local public economy need not "fragment" a metropolitan community. "Fragmentation" is a term that mixes description with evaluation. It is one thing to say that a metropolitan area contains a large number of provision units; it is another to say that the multiplicity of provision units "fragments" the metropolitan community. Degree of fragmentation is usually measured, both by opponents and proponents of metropolitan government, as the ratio of number of jurisdictions to population. Such a measurement, however, tells us nothing about the *fragmenting effect* of multiple jurisdictions.

The relevant question is the degree to which a more highly differentiated local public economy subtracts from the coherence of a metropolitan community. A coherent political community is one that is able to act in relation to communitywide concerns; a metropolitan political community is one that is able to act in relation to metropolitanwide concerns. Can a highly differentiated local public economy also act as a coherent community in relation to common concerns?

Daniel J. Elazar has argued that a complex set of local governments can function as a "civil community," constituted by citizens on the basis of intergovernmental and interorganizational relationships.²⁶ One mark of a civil community would be an ability to tend simultaneously to common and diverse interests. In 1972, Elazar noted that while "the civil community has no formal status in law, it has carved a place for itself in the constitutional system that is just now coming to be recognized and has yet to be precisely defined."²⁷

An ACIR study found a civil community in St. Louis County, Missouri—nearly a million people—of immense vitality.²⁸ This community finds diverse expression in 90 municipalities, a vigorous county government, 23 school districts, and 25 fire protection districts, plus countless organized subdivisions. It also finds common expression, not only in the county government, but in organizations of municipalities, fire chiefs, and police chiefs, the Cooperating School Districts of St. Louis County, a special district for special education, and, most especially, in the county delegation to the state legislature and occasional countywide referenda.

The county delegation—31 representatives and seven senators, all elected from districts—become, in effect, "constitutional" decision makers for the civil

community. They become the “keepers of the rules” that authorize and maintain citizen governance in the civil community. Special state legislation for St. Louis County, together with traditional legislative deference on local bills, gives to the civil community a significant constitutional capability. The “rules of the game” that establish the basic parameters within which citizens can govern a local public economy are maintained by the civil community acting through the local delegation to the state legislature.

The civil community thus is able to maintain a form of metropolitan (or regional) governance without having to create a “metropolitan government.” The presumed ideal of metropolitan government would create a single provision unit for an entire metropolitan community. A local public economy, on the other hand, generally consists of a variety of provision units—small and large. A single provision unit would, almost certainly, be suboptimal. Instead of thinking of metropolitan governance in terms of large general-purpose governments capable of both provision and production, it is possible to think in terms of citizens who maintain a civil community on the basis of rules. These rules, usually embodied formally in state law or the state constitution, become a kind of “local government constitution” (i.e., a set of legal instruments for governing local governments). As Elazar points out, a civil community is, at least potentially, much more autonomous than any single unit of local government can or should be.²⁹ This is because the civil community is not simply a provision unit. It is, in effect, a constitutional unit providing, not goods and services, but a framework of rules—an institutional framework within which local citizens are able to constitute the provision units that become the building blocks of a local public economy.

One of the functions of political community deemed important at least since Aristotle is an ability

to balance the interests of diverse groups and classes. A civil community composed of numerous independent local jurisdictions provides diverse groups and classes with their own official spokesmen and spokeswomen, whose voices can be heard in the broader political forums throughout the area. To be represented by one’s “mayor” in these deliberations is no small advantage. It is the elaborate infrastructure of political organization that often gives civil communities in metropolitan areas their distinctive qualities.

CONCLUSION

The study of local government constitutions in America is in its infancy. Although scholars have for many years described and classified state rules pertaining to local governments, the effort to understand each state’s rules as a constitutional configuration and estimate its impact on patterns of provision and production, with attendant efficiency effects, is just beginning. Counting the number of units of various types does not indicate whether, or by how much, a local public economy is structurally distorted. Only a research program of comparative institutional analysis that examines the “rules of the game” in relation to the strategies chosen by participants can explain the patterns of local government that have been created and generate an empirical basis for evaluation.

One of the most pervasive features of American local governance, however, is the use of rules that rely upon local *citizens* to make basic structural choices. Within the institutional framework supplied by general enabling legislation, citizens tend to occupy a central role in the *governance* of local public economies. The unity and coherence of a local public economy is supplied by the civil community that citizens and their elected officials maintain across jurisdictional boundaries.

NOTES

¹Technically, boundaries that are too small fail to internalize the potential externalities associated with service provision; some benefits spill over to others. The spillover problem is never completely resolved. Sometimes provision units attempt to convert symbolic boundaries into physical barriers, e.g., fences, walls, or deadend streets. As long as there are local government boundaries, nevertheless, there will be spillovers. The size of the benefit spillover relative to total provision diminishes, however, with the increasing size of a provision unit. At some point, spillovers become economically insignificant.

²The “Chinese box” imagery is used by Robert A. Dahl, *After the Revolution: Authority in a Good Society* (New Haven: Yale University Press, 1970).

³Little empirical work has been done on the effects of different rules. Martin and Wagner, in a study of the effects of boundary commissions in California, concluded that the introduction of greater restrictions on processes of provision unit formation tend to increase local government expenditures. See Dolores T. Martin and Richard E. Wagner, “The Institu-

tional Framework for Municipal Incorporation: An Economic Analysis of Local Agency Formation Commissions in California,” *Journal of Law and Economics* 21 (October 1978). Similarly, Dilorenzo found that restrictions in the growth of special districts in California and Oregon resulted in higher expenditures in selected service areas as compared to states experiencing a rapid growth in special districts. See Thomas Dilorenzo, “The Expenditure Effects of Restricting Competition in Local Public Sector Industries,” *Public Choice* 37 (1981), pp. 569-78.

⁴See ACIR, *State and Local Roles in the Federal System*, A-88, April 1982, p. 396. Only some 20% of known city-county consolidation proposals have been approved by voters since 1921, with the success rate declining.

⁵It should not be inferred that suburbs necessarily lack diversity. St. Louis County, for example, is highly diverse with respect to street patterns, housing type, lot sizes, median incomes, commercial development—all of the variation that gives a city its interesting geographic texture.

⁶Charles M. Tiebout, “A Pure Theory of Local Expenditure,” *Journal of Political Economy* 44 (October 1956), pp. 416-24.

- ⁷See Barbara H. Moore (ed.), *The Entrepreneur in Local Government*, (Washington, DC: International City Management Association, 1983).
- ⁸See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.
- ⁹Large sections of unincorporated St. Louis County have been developed for a number of years without an incorporation movement arising. One has to assume satisfaction with the homogeneous, but fairly low, service levels county government provides in these areas. On the other hand, incorporation and municipal annexation movements have arisen (and achieved mixed success) in other parts of the county. County government may find it difficult to vary its level of service provision from one part of the county to another. But particular communities that become dissatisfied have alternatives available. See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.
- ¹⁰*City and State*, July 1987, p. 6.
- ¹¹Salim Muwakkil, "Breaking Away: Growing Debate on Black Autonomy," *In These Times*, November 11-17, 1987, p. 7.
- ¹²*Ibid.*
- ¹³Albert O. Hirschman, *Exit, Voice, and Loyalty* (Cambridge: Harvard University Press, 1970).
- ¹⁴Local governments are frequently characterized as "creatures of the state." This is literally true, however, only in the case of local governments that have special-act charters. General enabling legislation supplies a set of rules that allows local communities to create local governments.
- ¹⁵Joseph F. Zimmerman reported in 1983 that only Alabama, Kentucky, and Virginia did not permit municipalities to frame their own charters. See Zimmerman, *State-Local Relations: A Partnership Approach* (New York: Praeger, 1983), p. 26.
- ¹⁶See ACIR, *Measuring Local Discretionary Authority*, M-131, November, 1981.
- ¹⁷Dillon's Rule, a rule of judicial interpretation frequently cited in relation to local governments, requires that all local authority be strictly construed and that doubts be resolved against the local government claiming authority to act. To the extent that home rule allows for local determination of authority through a local charter, the restrictive effect of Dillon's Rule is ameliorated.
- ¹⁸"The simple but concurrent majority approval requirement may make the consolidation more difficult to achieve than a referendum with approval by a majority of the combined electorate in the jurisdictions affected. Yet the Commission believes that concurrent majority approval gives added assurance to the residents of each city, town, or township that they will not be unwillingly included in a consolidated government and provides a salutary political basis for launching a new municipality." ACIR, *State and Local Roles in the Federal System*, A-88, April 1982, p. 449.
- ¹⁹See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.
- ²⁰ACIR, *Governmental Structure, Organization, and Planning in Metropolitan Areas*, A-5, July 1961, pp. 39-41.
- ²¹The most recent Commission statement is found in ACIR, *State and Local Roles in the Federal System*, A-88, April 1982, p. 445. An earlier statement was included in ACIR, *Fiscal Balance in the American Federal System*, Vol. 2, *Metropolitan Fiscal Disparities*, A-31, October 1967, p. 14.
- ²²The ten states and years of adoption are as follows: Minnesota, Wisconsin, Alaska—1959; California, Colorado, New Mexico—1965; Washington, Nevada—1967; Michigan, 1968; and Oregon, 1969. This information was reported by Joseph F. Zimmerman in 1983. See Zimmerman, *op. cit.*, pp. 110-20.
- ²³Ronald C. Cease, *A Report on State and Provincial Boundary Review Boards* (Portland, Oregon: Portland State College, 1968). Cited in Zimmerman, *op. cit.*, p. 119.
- ²⁴David L. Chicoine and Norman Walzer, *Governmental Structure and Local Public Finance* (Boston: Oelgeschlager, Gunn & Hain, 1985), p. 215.
- ²⁵Dolores T. Martin, *Institutional Barriers in the Local Government Market: Effects on Efficiency and Monopoly Power*, Ph.D. Dissertation, nd. See also Martin and Wagner, *op. cit.*
- ²⁶Daniel J. Elazar, *Cities of the Prairie: The Metropolitan Frontier and American Politics* (New York: Basic Books, 1970) and *American Federalism: A View from the States*, 2nd. ed. (New York: Thomas Y. Crowell, 1972), pp. 183-92.
- ²⁷Elazar (1972), *op. cit.*, p. 185.
- ²⁸See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.
- ²⁹Elazar (1972), *op. cit.*, p. 185.

INTRODUCTION

One difficulty posed by a large number and variety of provision units in a local public economy is the emergence of disparities in provision. Some communities will have a greater ability than others to raise own-source revenues to support public services. In fact, provision-side efficiency necessarily implies a degree of disparity in spending. If different communities have different preferences, those differences will be reflected in levels of spending from own-source revenues. At the same time, however, principles of equity in a democratic society suggest that there ought to be limits to the permissible range of disparity—although no objective definition of those limits is possible.

The traditional approach to reducing fiscal disparity has emphasized the development of balanced communities, namely, municipalities whose residents are heterogeneous with respect to income or wealth. Yet such an approach is inconsistent with efforts to attain provision-side efficiency in terms of such criteria as fiscal equivalence and preference revelation. An alternative method of redistribution is to permit disparities to develop among multiple provision units, such as municipalities, but to correct inequities by means of grants from overlapping jurisdictions, principally, state and federal governments. A theory of local public economies suggests that intergovernmental grants may actually be more effective instruments of redistribution than governmental consolidation. This is because interjurisdictional transfers allow recipient communities to be organized as separate provision units, which are also able to make their own production arrangements.

ALTERNATIVE APPROACHES TO FISCAL DISPARITY PROBLEMS¹

One difficulty found in discussions of “inequity” is the often implied standard of equality or homogeneity. Yet provision-side efficiency necessarily implies a degree of inequality or heterogeneity of provision as long as the economic demand for public goods and services

is variable. One widely used principle of equity, however, does not imply equality, and, in fact, assumes inequality. This principle evaluates a process or change by its treatment of the “least advantaged” groups in a society. The question then is, “How do alternative governance arrangements in a metropolitan area affect the least advantaged members of a society?” Various benchmarks can be used for comparison, including the treatment of median-advantaged and most advantaged groups. Such an approach allows us to examine the distributional effects of alternative provision arrangements, from an equity standpoint, without having to use absolute equality as a standard.

It is axiomatic that the greater the number of non-overlapping provision units in a local public economy, the greater will be the variation in average household income among those units. By consolidating provision units and thus averaging over greater numbers of households, the variation among units is reduced. The variation among households aggregated over all units remains exactly the same. These statistical differences do not imply that a greater number of local jurisdictions causes greater economic “segregation” among residential communities. Residents sort themselves among neighborhoods according to income to some degree, whether or not neighborhoods are organized as separate jurisdictions.

Nevertheless, the existence of fiscal disparities among multiple provision units does imply a need to rely on intergovernmental transfers to relieve inequities that may develop. Historically, the ACIR has favored an approach that emphasized the development of well-balanced, heterogeneous communities as local government jurisdictions able to pool resources from different income classes and provide services equitably, according to need.² This approach sought to avoid reliance on redistributive transfers primarily on the ground that it is politically more difficult to achieve interjurisdictional equity than it is to achieve intrajurisdictional equity in a fiscally balanced community. Yet it has proven to be much more difficult, politically, to consolidate local jurisdictions than to un-

dertake fiscal transfers from overlapping jurisdictions. Local communities have found that it is not necessary to give up their jurisdictional autonomy in order to benefit from redistribution or to support other communities that are less well off.

It is also difficult to wrap a local boundary around the disparity problem. Creating a metropolitan jurisdiction may address disparities within a metropolitan area, but it does nothing for differences between metropolitan and nonmetropolitan areas, or between different metropolitan areas. How far must local boundaries be expanded? Local jurisdictions, even if highly consolidated, are likely to fall somewhat short of being fully able to address the problem of fiscal disparities.

The basic economic difficulty with an intrajurisdictional approach to redistribution, however, is its incompatibility with efforts to attain provision-side efficiency. The criterion of fiscal equivalence (discussed in *Chapter One*) is based upon distributing services to communities in proportion to the revenues raised from those communities. The redistributive purpose of developing a balanced, heterogeneous community is just the opposite: to use revenues raised from one community to provide services to another. This juxtaposition of approaches frequently leads to a conclusion that the objectives of efficiency and equity are fundamentally in conflict, requiring a hard choice between competing values.

There is, however, an alternative possibility, suggested by a theory of local public economies. This approach first allows provision units (e.g., municipalities) to form on the basis of citizen preferences, thus allowing citizens to make efficient provision arrangements—those that can best satisfy their preferences. This permits disparities to develop. Then, however, more inclusive, overlapping jurisdictions (e.g., states) can correct for unacceptable disparities through intergovernmental grants. Two advantages follow from this approach. First, provision-side efficiency can be obtained, within the limits of equity. Standards of equity are used to establish limits, within which local citizens are free to maximize their preference satisfaction, but are not used to override citizen choices by consolidating local jurisdictions. Second, the values and resources of an entire state or nation—communities much more inclusive and heterogeneous than any local jurisdiction—can then be used to guide and support redistribution on equity grounds.

When comparing fiscal disparities in a highly differentiated (or fragmented) local public economy with those of a more consolidated set of jurisdictions, the appropriate base of comparison is not the fiscal disparity that exists as a matter of own-source revenues, but the disparity that remains after intergovernmental transfers. Fragmentation *without* overlap would be unable to resolve problems of fiscal disparity; fragmentation *with* overlap—the core of federalism—is a jurisdic-

tional arrangement with greater ability to address efficiency and equity concerns simultaneously.

This conclusion is underscored by a recognition that equity problems are not confined to *fiscal* disparities across jurisdictions; equity problems also occur as *service* disparities within jurisdictions. The expenditure side of local government is at least as relevant to equity as the tax side. What is more, the efficiency with which money is spent and the responsiveness of service provision to community preferences intervene between taxes and expenditures, on the one side, and equity, on the other. *Equity is an attribute of service outcomes, as well as tax and expenditure outcomes.* Fiscal disparity among local governments is at most an intermediate indicator of equity; its relationship to service outcomes, and thus finally to equity, depends upon a host of other factors.

The attainment of equity in local service provision is, unfortunately, a complex problem. If it were possible to achieve equity simply by reducing disparities in revenue potential, then any pattern of organization that tended to increase those disparities would earn a negative rating on equity in its overall scorecard. Matters are not, however, so simple.

EMPIRICAL RESEARCH

Several unanswered empirical questions are at issue. How do differences *among* jurisdictions in a highly differentiated local public economy compare to differences among communities or neighborhoods *within* local jurisdictions, especially in large cities? How do fiscal disparities among jurisdictions compare to service disparities within jurisdictions? Moreover, how is this comparison affected by intergovernmental fiscal transfers? Which pattern of organization—fragmentation or consolidation—provides for better “trusteeship” of intergovernmental revenues? Existing studies, while not conclusive, begin to shed light on these questions.

If one examines the own-source revenue potential of local jurisdictions in the more highly differentiated local public economies found in the suburban regions of metropolitan areas, considerable variation in “fiscal capacity” is apparent. In the most extensive and revealing study to date, Mark Schneider and John R. Logan,³ who analyzed 1970-72 data for 1,139 municipalities in 31 SMSAs, report the very rich to be the most highly segregated group (dominant in 201 municipalities), followed by the very poor (dominant in 192 municipalities). The dispersion of the very poor is also somewhat greater than the very rich throughout other municipalities. A total of 746 municipalities in this study, however, are dominated by neither the very rich nor the very poor. Of this number, only 66 municipalities are dominated by any particular income class, leaving *the vast majority of suburban municipalities with a diverse population by income.* A picture

of suburbia as highly segregated by income class appears to be inaccurate, except for enclaves of the very rich and the very poor.⁴

Own-source revenue potential, however, represents only an *initial* distribution of resources. When intergovernmental transfers are taken into account, the picture that emerges is still different. Schneider and Logan report that, in 1972, the *total* revenue and expenditures of the very poor jurisdictions in their sample *exceeded* that of middle income municipalities and even, slightly, that of the very rich.⁵ Substantially, however, this finding is due to greater expenditures on social services, including housing and hospitals, in poorer communities.⁶ In 1972, federal aid as a percentage of cities' own-source revenues was only 10.8 percent. This figure increased to a high of 25.8 percent in 1978, but had fallen to 13.1 percent by 1985.⁷ One of the disadvantages of relying on federal grants-in-aid to help distressed communities is perhaps the relative instability of assistance.

The ACIR study of St. Louis County, Missouri, used two different approaches in an effort to ascertain the presence of possible inequities among citizens residing in 90 municipalities.⁸ First, the analysis included a search for "patterned inequalities," namely, variations in total revenue per municipality correlated with indicators of individual advantage or disadvantage in society. Three indicators were used: (1) median household income, (2) percent of population below poverty, and (3) percent of population nonwhite. None of the coefficients were significant. Revenue disparities among St. Louis County municipalities (after intergovernmental transfers) are not strongly patterned along income or racial lines. Second, however, the analysis also identified a small set of municipalities characterized by high tax burden (residential source revenues per household as a percentage of median household income) and low, total per capita revenue, suggesting possible inequity. Overall, tax burden is related both to race and to age—citizens who are nonwhite and in their senior years tend to bear greater tax burdens, after adjusting for income differences and other community characteristics, than do other citizens. This analysis suggests that some scope still exists for redistribution to correct inequitable tax burdens in St. Louis County, even though existing disparities in revenues are not strongly patterned along lines of race or income.

Studies of resource distribution within large centralized urban jurisdictions have generated mixed results. Various studies have found that large central-city jurisdictions tend to spend (a) more on police,⁹ (b) less on streets,¹⁰ and (c) less on libraries¹¹ in poor neighborhoods than in other neighborhoods. One well-known study of Oakland, California, found that the distribution of resources within the public school system there followed a u-shaped curve. More re-

sources were allocated to schools for the very poor and the very rich, than for "middle class" schools.¹²

The distribution of resources, whether among or within political jurisdictions, however, does not settle the question of equity. The existence of a single, large jurisdiction potentially capable of moving resources from the very rich to the very poor, in expenditure terms, does not necessarily imply an actual distribution of benefits to communities in need. Numerous studies have found that the quality of urban public education, for instance, is considerably lower in poor neighborhoods than in other neighborhoods.¹³ Research on the efficiency and responsiveness of service provision units (discussed in *Chapter Three*) at least suggests the possibility that smaller jurisdictions in "fragmented" areas may be better trustees of fiscal resources. If this is the case, then interjurisdictional fiscal transfers may be preferred to intrajurisdictional transfers as a means of addressing equity concerns. It is possible that expanding the boundaries of local government jurisdictions may not promote equity as effectively as reducing fiscal disparities through intergovernmental sharing.

This conclusion is fundamentally at odds with a widespread belief about metropolitan relationships, namely, that suburban autonomy is a major cause of central-city fiscal distress. The argument made is that wealthier residents fled to the suburbs precisely to avoid having to pay for services to poorer residents, allowing them to tax themselves exclusively for the purpose of providing their own high-quality services. Wealthy suburban residents then continue to derive benefits from their proximity to a major city without having to support it financially. To overcome this equity problem, analysts frequently propose that central cities be given unilateral annexation powers and that municipal incorporation in suburban areas be restricted.

Differing interpretations can be placed, however, on the so-called "flight to the suburbs." The wealthy have not been alone in this movement. Middle and lower income residents have joined the exodus from cities. The attraction of the suburbs, rather than an ability to escape from taxes used to support redistributive expenditures, may be due, in part, to the greater ability to control what happens in one's community. Citizens move to the suburbs, from this perspective, to obtain a more effective voice in community governance. Redistributive taxation cannot be avoided in any case as long as more inclusive jurisdictions support redistribution. Suburban municipalities do tend to allow greater citizen control over discrete communities, an advantage not confined to a few enclaves of the very wealthy. The fiscal imbalance between suburbs and central cities may, then, be a function of the cities' governance arrangements, leading to a loss of citizen efficacy, followed by exit. Expanding

central-city boundaries could tend in the long run to perpetuate this problem, and to exacerbate its underlying causes.

POLITICAL CHOICE FOR DISADVANTAGED COMMUNITIES

The central problem with boundary expansion as a means of reducing inequities, whether it be an expanding central city or the consolidation of suburban municipalities, is the tendency to deprive already disadvantaged communities of an ability to make their own collective choices; that is, consolidation can make economically disadvantaged communities into politically disadvantaged communities as well. The function of the provision side of a local public economy is not simply to raise revenue. It is also to enable residents to choose how to spend public revenue and, most importantly, to choose *how to arrange for production*. Poor communities without this choice may see their scarce funds eaten up by urban bureaucracies. Relying upon intrajurisdictional transfers deprives recipient communities of the ability to make these "provision side" choices. It may also disable poor communities from seeking more equitable treatment in the context of a larger, more inclusive community, through their own elected officials.

Simply because poor communities may not be able to raise all of their own revenue is no reason to deprive them still further of the ability to function as a provision unit. Receiving public funds by interjurisdictional transfer is a function that requires a provision unit. The choice of a provision unit for this purpose obviously does not have to be driven by the ability to raise revenue.

In suburban areas, provision units that represent the interests of disadvantaged communities already exist to some extent. In many central cities, however,

it may be necessary to create new provision units—or rather, to allow for their creation by citizens—in part for the purpose of receiving public funds, so that the structure of a local public economy might be more conducive to the use of an interjurisdictional approach to help disadvantaged communities. ACIR recommendations on the formation of neighborhood governments address the potential for this pattern of organization.¹⁴

The mitigation of local fiscal disparities need not be a responsibility reserved exclusively to state and national governments. Metropolitan-area resource sharing is also possible—the prime example being a partial sharing of the commercial and industrial property tax base pioneered by the Minneapolis-St. Paul metropolitan community.¹⁵ By partially sharing the revenue growth from economic development, the emergence of fiscal disparities among local jurisdictions is constrained. Other metropolitan communities should be studied to ascertain the extent to which fiscal disparities can be mitigated by the policies of local overlapping jurisdictions and by adjustments in the fiscal rules of local governance, including local revenue diversification.¹⁶

Future research should also study both the instruments of fiscal transfer used by overlapping jurisdictions, including state and federal grant-in-aid formulas, and the performance of provision units (including big cities and small cities) that receive funds. At issue is the ability of both granting and receiving jurisdictions to focus assistance on those communities in greatest need. Historically, the ACIR has closely monitored metropolitan fiscal disparities.¹⁷ The challenge, now, is to expand the scope of inquiry to include neighborhood disparities within jurisdictions in order to render a comparative assessment and formulate effective intergovernmental policies for addressing both urban and suburban equity problems.

NOTES

¹Excellent overview discussions of equity issues in both an interjurisdictional and an intrajurisdictional context are found in Astrid E. Merget and Renee A. Berger, "Equity as a Decision Rule in Local Services," in Richard C. Rich (ed.), *Analyzing Urban-Service Distributions* (Lexington, Mass.: Lexington Books, 1982), pp. 21-44 and Elinor Ostrom, "The Social Stratification-Government Inequality Thesis Explored," *Social Science Quarterly* 19 (September 1983), pp. 91-112.

²See especially ACIR, *Fiscal Balance in the American Federal System*. Vol. 2, *Metropolitan Fiscal Disparities*, A-31, October 1967.

³Mark Schneider and John R. Logan, "Fiscal Implications of Class Segregation: Inequalities in the Distribution of Public Goods and Services in Suburban Municipalities," *Urban Affairs Quarterly* 17 (September 1981), pp. 23-37.

⁴Income is not, of course, entirely adequate as a measure of fiscal capacity because of the tax resources represented by commercial and industrial property. The effect of introducing this additional variable, however, would probably not tend to

enhance the revenue potential of the very rich relative to the very poor.

⁵In fact, a u-shaped curve was observed relating income class to revenues and expenditures, as both the very poor and the very rich out-spent middle income municipalities. Breaking down expenditures by function, the very poor spent a great deal more on social services than either middle income municipalities or the very rich; but the very poor also spent somewhat more on "common functions" than did middle income communities, though not as much as the very rich. (Seven "common functions" are used in this study: general control, general government, highways, police, fire, parks and recreation, and sanitation.) In terms of total revenues and expenditures, very poor municipalities tended to be brought up to the standard of the very rich, or better, by intergovernmental transfers. In terms of spending on common municipal services, the very poor were brought up to the standard of the "middle class," or better. It should be recognized that even these levels of expenditure might not meet a standard of *compensatory* spending designed to bring standards of living in very poor communities up to some minimum. No data is available using such a standard. The intergovernmental trans-

fers taken into account do not, however, include federal and state transfer payments to individuals and families living in these communities.

⁶Schools are included only when they are a municipal responsibility. Usually schools are funded separately.

⁷ACIR, *Significant Features of Fiscal Federalism, 1987 Edition*, M-151, June 1987, p. 57.

⁸See ACIR, *Metropolitan Organization: The St. Louis Case*, forthcoming 1987.

⁹J.C. Weicher, "The Allocation of Police Protection by Income Class," *Urban Studies* 8 (October 1971), pp. 207-20; Robert L. Lineberry, "Equality" and *Urban Policy: The Distribution of Municipal Public Services* (Beverly Hills: Sage Publications, 1977); Kenneth Mladenka, "Serving the Public: The Provision of Municipal Goods and Services," Ph.D. Dissertation, Rice University (1974) and "The Distribution of Urban Police Services," *Journal of Politics* 40 (February 1978); P.F. Nardulli and J.M. Stonecash, *Politics, Professionalism, and Urban Services: The Police* (Cambridge, Mass.: Oelgeschlager, Gunn & Hain 1981).

¹⁰G.E. Antunes and J.P. Plumlee, "The Distribution of an Urban Public Service: Ethnicity, Socioeconomic Status and Bureaucracy as Determinants of the Quality of Neighborhood Streets," *Urban Affairs Quarterly* 13 (March 1977), pp. 313-32.

¹¹Frank S. Levy, Arnold J. Meltsner, and Aaron Wildavsky, *Urban Outcomes* (Berkeley: University of California Press, 1974).

¹²*Ibid.*

¹³See, for example, M.T. Katzman, *The Quality of Municipal Services, Central City Decline, and Middle-Class Flight*, Research Report R78-1 (Cambridge, Mass.: Harvard University, Department of City and Regional Planning 1978); R.A.

Berk and A. Hartman, *Race and District Differences in Per Pupil Staffing Expenditures in Chicago Elementary Schools, 1970-1971* (Evanston, Illinois: Northwestern University, Center for Urban Affairs, 1971).

¹⁴See ACIR, *The States and Distressed Communities: The Final Report*, A-101, November 1985.

¹⁵The Minneapolis-St. Paul arrangement, adopted by Minnesota statute in 1971, provided that municipalities in the metropolitan area would share 40% of the area's future growth in commercial-industrial property assessments. ACIR recommended in 1985 that other states consider the adoption of similar tax-base sharing arrangements in "densely settled" areas. See ACIR, A-101, *op cit*, November 1985, pp. 251-53.

¹⁶ACIR has long advocated diversification of the local revenue base to include sales and income taxes, as well as property taxes and user charges. See ACIR, *Local Revenue Diversification: Income, Sales Taxes and User Charges*, A-47, October 1974. A common assumption has been, however, that diversification does nothing to alleviate—and may even exacerbate—fiscal inequities. Limited evidence is now being generated to the contrary. A recent study of the seven-county Minneapolis-St. Paul area finds only a slight positive correlation between per capita taxable sales and per capita assessed valuation among cities. At the same time, "Stronger but still moderate correlations were found between sales tax bases and both per capita property tax levies and mill rates." This suggests that the use of a city sales tax would tend to mitigate interjurisdictional fiscal disparities, even though there is a wide variation in sales tax base among cities in a metropolitan area. See J. H. Fonkert, "Local Revenue Diversification in Minnesota," *Minnesota Tax Journal* 2 (winter 1986).

¹⁷The most recent report is ACIR, *Fiscal Disparities: Central Cities and Suburbs, 1981*, An Information Report, August 1984.

Conclusion: Research and Policy Directions

INTRODUCTION

The elementary distinction between the “provision” and “production” of public goods and services is fundamental to a rethinking of how citizens ought to organize the process of local government in a changing world. Provision and production are separable activities that can be linked in a variety of ways. Thus, one of the basic implications of distinguishing provision and production is the use of contractual relationships to link providers (e.g., municipalities) with producers (e.g., private enterprise, other governments). The implications, however, go well beyond contracting *per se*. Efficiency criteria, which are concerned with maximizing human welfare, yield quite different conclusions when applied to provision than when applied to production. Provision-side efficiency is concerned with how best to satisfy individual preferences; production-side efficiency is concerned with how to transform human and material resources into preferred outputs at least cost. The appropriate scale for organizing provision often differs from the appropriate scale for organizing production. When this is the case, the use of a single organization both to provide and to produce a service may be less than efficient.

A local public economy can, therefore, be viewed as having a provision side and a production side, each of which can be organized differently, according to different criteria. The provision side can be expected, optimally, to consist of some array of provision units, each able to act collectively on behalf of a specific community of interest. This is equivalent to suggesting, in traditional terms, that local government in metropolitan areas ought to be “fragmented” among a variety of units. Some provision units will be nested inside others, allowing jurisdictions to overlap. The powers and responsibilities of provision units in a well-functioning public economy will also vary in ways that tend to be complementary—as often reflected in the different capabilities and limitations of municipalities, spe-

cial districts, counties, and townships. The production side, on the other hand, will consist of both public and private units, linked to provision units through a variety of arrangements, including in-house organization of production, contracting, and the organization of joint production units.

METROPOLITAN GOVERNANCE

A great deal of intellectual energy, spanning several decades, has gone into efforts to determine the correct pattern of organization for metropolitan areas. Conclusions have included proposals for “metropolitan government” (a single unit to provide all and produce most local public goods and services for a metropolitan area) and for various two-tier (or multi-tier) arrangements, with functions carefully sorted among the tiers. Such proposals have been motivated by various considerations—economies of scale in production, service coordination, division of labor, spillover effects, and redistribution or fiscal equalization. Whatever the particular pattern of organization being proposed, a common goal of metropolitan reform has been to replace a “crazy-quilt” pattern of jurisdictions with a simpler pattern organized according to uniform criteria applied consistently throughout a metropolitan area. This approach was thought to be the only way to achieve a logical or coherent system of organization.

If provision-side criteria are based upon individual preference satisfaction, however, then there is no fully objective way of determining an appropriate set of provision units apart from the expressed preferences of local citizens. Historically, local citizens have usually resisted most metropolitan reform efforts involving substantial jurisdictional consolidation. It is important to try to understand why. Objective criteria, related to the scale of effect associated with the provision (or non-provision) of a public good or service, can be used to establish a minimum-size provision unit. For many local public goods and services, however, this lower limit is quite small, consisting of what is ordinar-

ily called a "neighborhood." The outer limit on local boundaries derives from individual preferences and the way in which these preferences cluster geographically. The ease with which a single provision unit can satisfy individual preferences increases with the preference homogeneity of the community. This proposition implies a further proposition, namely, that preference satisfaction increases with an increase in the number of provision units in a local public economy, up to some point. At the same time, the creation of provision units is constrained by the expected transaction costs associated with organizing and operating an additional unit, including the costs of citizen participation. The nature of the choice involved is one of a trade-off. The choice is between greater preference satisfaction, obtained by creating an additional unit, and lesser transaction costs.

From this line of reasoning, a "crazy-quilt" can be viewed as a logical pattern of organization; it is a pattern that conforms to a certain logic of community formation. This does not mean that any pattern is optimal. In traditional terms, "fragmentation" is neither a wholesale "bad" nor an unqualified "good." The key question is *who decides*, incrementally, whether to add (or subtract) a provision unit. In many local public economies in the United States, especially the more highly differentiated metropolitan areas, *citizens decide*, both directly and indirectly. The pattern of organization is one that emerges from the incremental choices of local citizens, acting through processes of petition and referendum. When allowed to choose, citizens in different communities tend to make different choices. Although particular individuals, such as local officials, private developers, or interested citizens, may take entrepreneurial initiative in proposing a new provision or production unit, the final decision usually rests with voters within the proposed new local boundaries. As a result, the pattern of organization that emerges from citizen choice is not uniform throughout a metropolitan area. Different communities make different trade-offs. By contrast, the establishment of a uniform pattern throughout a local public economy (perhaps a two-tier system) implies that the same trade-off is appropriate for every citizen in every community.

There is no *a priori* reason to suppose, however, that citizen choices will allow local governments simply to *proliferate*—to increase in number without limit. Instead, citizens can be expected to take into account both the potential benefits—in terms of preference satisfaction—and the potential costs—in terms of transactions—of adding a unit. Simple growth in the number of units—whether municipalities or special districts or some other type of unit—is not, from the perspective of citizens, evidence of suboptimal growth.

Much the same logic applies to the organization of the production side. Production units can be added so

long as the marginal costs—mostly transaction costs—of operating additional units are less than the benefits to be derived from more closely matching the size of production units to diverse economies of scale. The key question, again, is who decides. The answer, in this case, is generally local officials, who tend to modify production arrangements incrementally.

The *governance* of many local public economies is thus primarily in the hands of local citizens—those who make the choices that determine the basic structure of the provision side. Governance is separable from both provision and production. Metropolitan "governance" therefore does not require a metropolitan "government." Governance depends upon a capacity to make rules, but not upon a capacity to provide services. The ability of citizens to govern a local public economy also depends upon a set of rules, supplied for the most part by means of state legislation and state constitutions. These enabling rules establish the basic governance structure, or "constitution," of a local public economy.

Enabling rules include rules of municipal incorporation, annexation, consolidation, and disincorporation, plus similar rules pertaining to the formation of special-purpose governments, school districts, and other local units. In any local public economy, these rules, though derived from a variety of legal sources, should be understood as a single configuration—a local government constitution. This configuration of rules both allows citizens to govern and constrains citizens in the formation of units. Much greater attention should be paid to how these rules inter-relate and jointly affect citizen choices. Whether citizens are able to develop an optimal number of provision units depends upon how the local government constitution structures and constrains their choices.

Local public economies are potentially subject to serious structural distortion when the rules contained in a local government constitution either lead citizens to create and maintain too many units or preclude citizens from creating a sufficient number of units to satisfy preferences economically. There is no magic number that is just right for every citizen in every community. The essential consideration with respect to governance is therefore constitutional in nature—how to design a rule structure that allows citizens to make optimal choices. When issues of metropolitan governance are to be addressed, this is the crucial level of analysis and inquiry to examine.

IMPLICATIONS FOR FURTHER RESEARCH

Basic research in the governance of local public economies must be able to address this question—how can we know when provision arrangements (and production arrangements) are optimal? The answer depends upon an analysis of the relationships between

particular governance structures and the choices that are made within those governance structures. Different configurations of rules establish different incentives, and therefore, to some extent, an analysis of rules can anticipate a pattern of choice. Such inferences, however, must always be checked against empirical inquiry, examining the patterns of choice that actually occur within different structures.

Local government constitutions tend, in the United States, to be endogenous expressions of a local "civil community," not an imposition from outside authorities. The rules used to govern local public economies tend to be generated from within areawide communities, even if formally adopted by state legislatures. Unsatisfactory arrangements can be expected to generate local efforts to change the "rules of the game." A careful distinction must be made, however, between reform proposals and the problems that give rise to them. Actual reform proposals are as much a product of the theory that individuals bring to their experience as of the experience itself. Lack of conformity with an *a priori* theory may also generate efforts at reform. Local government constitutions also tend to be subject to fairly continuous modification—tinkering at the margins of organizational arrangements. This process can hold a number of clues as to how governance structures relate to choice and performance in a local public economy.

A range of research efforts is needed. One major thrust should consist of comparative research on local government constitutions across the fifty states or some sample of states. This research should identify particular configurations of rules that are of interest and examine variations in those configurations. Such a research program would move beyond previous work that has tended to examine variation in particular rules (e.g., annexation rules or home-rule provisions) without sufficient attention to how related rules are configured. One dimension of this work should be historical—tracing the evolution of rule configurations in specific states. Another major thrust should pursue comparative metropolitan-area research, building perhaps to some extent on the ACIR work in the St. Louis area. Again, an historical dimension is important, as well as cross-sectional analysis. A central focus of this work should be on *process*. The question is to what extent governance rules frame a process that enables citizens to make optimal choices of provision arrangements. The study of process is essential to be able to understand how rules structure choices. Only in this way can we learn how local public economies are governed, identify structural distortions when present, and point the way to remedies.

A research effort is also needed to focus on equity. Efficiency and equity, however, are not unrelated. Provision-side efficiency may be a first-order condition for attaining equity as a second-order condi-

tion. If particular communities of interest are deprived of the organizational means to make provision for their collective wants, effective redistribution may not be able to occur. To shed light on this important issue, comparative research is needed between disadvantaged communities that are (1) located within large heterogeneous jurisdictions, so that provision agents are not accountable to specific communities and (2) separately organized as municipalities or some other autonomous unit. This research should compare such differently governed communities with respect to responsiveness of service provision, self-help efforts, the efficiency of production arrangements, and the use of intergovernmental revenues.

Applied research is important as well. As cities, counties, metropolitan areas or non-metropolitan regions, and states experience problems, those problems become a useful focus of inquiry, for the sake of both policy development and the generation of knowledge. The framework set forth in this report can be used to guide a process of applied research directed at solving particular problems. By distinguishing between the provision and production aspects of problems, and thus clarifying the criteria and information appropriate to decisions, and by sorting out the governance issues that are pertinent, and thus clarifying the level of analysis and action appropriate to those issues, policy analysts working with this framework can contribute to the solution of problems. The better we understand how local public economies work, the better we can make them work.

IMMEDIATE POLICY IMPLICATIONS

A number of policy issues arise directly from this report. One important issue concerns the use of boundary review agencies, some of which are able to veto incorporations and annexations without regard to local citizen choice. Boundary review agencies were created on the assumption that local-government "proliferation" needed to be constrained. Yet there is no *a priori* reason to suppose that citizens create units of local government willy-nilly, or that a multiplicity of units is undesirable. To the contrary, the choices of citizens in creating, modifying, or abolishing units of local government may be a necessary condition of maintaining an optimally structured local public economy. The ACIR has therefore urged caution in the use of boundary review agencies.

Another issue of growing interest is the relationship between local governments and the increasing number of residential community associations (RCAs) in metropolitan areas. RCAs are formally private, but fulfill many public purposes. How should local governments within which RCAs are located respond to them? Should their formation be encouraged or discouraged? Once formed, should RCAs be given incentives to provide their own services or to depend on a

larger community (municipality or county) for service provision? Do RCAs represent a structural adaptation to changing conditions in local public economies (and if so, what conditions?), or are they a structural distortion that ought to be corrected through legal reform?

ACIR recommendations that large central cities establish neighborhood sub-units of government have not been widely implemented. While generating considerable interest in intellectual circles, neighborhood government is a movement that has gone practically nowhere fast. Is this because citizens are satisfied with existing arrangements? Or is it because the basic rules of governance in large municipalities do not generally allow citizens sufficient opportunity to make basic structural choices affecting the organization of their local public economies? This is a fundamental policy issue that must await further research.

CONCLUSION

Distinguishing between provision and production

establishes a basis for understanding the efficient organization of local public economies. Variety can be expected on both the provision side and the production side. Yet the potential variety is subject to constraint—largely due to the expected costs of transactions. Citizen choice is an essential element in the governance of local public economies; it is the basic source of variety. If allowed to choose, different groups of citizens tend to choose differently. Yet the resulting variety in organization conforms to a logic of economic choice and is not necessarily inefficient. Efforts to understand and improve the structure of local public economies should focus on rules of governance and how those rules either allow citizens to make optimal choices or constrain them to make suboptimal choices. Such an approach, as opposed to one that would deny citizens the opportunity to choose, can contribute to more productive local self-government in the American federal system.

Commission Findings and Conclusions

The Commission finds that the nation needs as its local foundation the most efficient and responsive—dynamic and creative—local public economies possible. It is a necessary foundation, not only for an efficient and responsive national public economy, but also for promoting greater equity through intergovernmental sharing. The efficient and responsive performance of local public economies—the complex of public and private organizations that supply local public goods and services in American communities—has depended historically upon rules that allow a variety of local units to emerge from a process of citizen choice.

The Commission concludes that the following general principles are useful in guiding the organization of local self-government in the American federal system:

1. VARIETY AND DIVERSITY

A strong case can be made for local governments that range in size from small neighborhood units to areawide units and, in functional scope, from single-purpose to multi-purpose units. To obtain the benefits of large-scale organization, it is unnecessary and unwise to drive out small-scale organization. Large and small, as well as single-purpose and multi-purpose, units of local government, when used concurrently, reflect complementary expressions of public preference, not contradictory principles of organization.

2. DISTINGUISHING PROVISION AND PRODUCTION

Local governments can usefully be viewed as “provision units,” organizations that make provision—both through tax and spending decisions and regulations—for local public goods and services. As provision units, local governments can choose how to carry out “production,” whether by organizing their own production units through government departments, or by contracting out to private firms or other govern-

mental agencies. The ability to make this choice is a basic principle of local governance.

3. OPPORTUNITIES FOR INITIATIVE, CREATIVITY, AND COOPERATION

The efficiency and productivity of local public economies depends upon opportunities for initiative, creativity, cooperation, and public entrepreneurship. These opportunities are associated with both the potential variety in provision arrangements available to local citizens and the development of alternative ways of relating provision to production, including both government production and private contracting. Distinguishing provision and production allows for greater variety in both. A significant reduction in the number and variety of local government units, in order to unite provision and production, could actually reduce the opportunities available to local citizens and officials to increase efficiency and productivity.

4. CITIZEN CHOICE AND CONSENT IN LOCAL GOVERNANCE

The governance of local public economies—“metropolitan” governance in an urban/suburban context—is best viewed as a problem in defining and altering the basic “rules of the game.” These rules—usually in the formal province of state legislatures—concern the formation of local government units, boundary changes (such as annexation), revenue raising and regulatory capabilities, and interjurisdictional relationships (cooperation and contracting). The unity of a metropolitan or rural community depends upon the coherence of these basic rules of local governance, and upon an underlying consensus, not upon the creation of a single dominant unit of local government or upon any detailed blueprint for local government reorganization. The particular pattern of local government units in a community emerges, over time, from the choices that citizens and officials make within the basic rules of local governance.

Rules that allow citizens to make the basic structural and fiscal decisions pertaining to local government are more likely to yield results consistent with citizen preferences. Local citizens, historically, are the creators and sustainers of American local government. Citizens are also “coproducers” of public services alongside of government officials. They are not simply the consumers of local services supplied in jurisdictions created by others.

5. EQUITY AND COMMUNITY CHOICE

Standards of equity in the distribution of local public goods and services are a fundamental consideration along with standards of efficiency. The quest for greater equity, however, cannot effectively be pursued independently of the search for efficiency, productivity, and responsiveness. To seek equity in ways that subtract significantly from the efficiency and productivity of a local public economy is to redistribute a shrinking economic pie. To seek equity in ways that subtract significantly from the responsiveness of service provision is to undermine citizen choice and consent.

Equity considerations alone may be insufficient to justify a reduction in the number and variety of provision units in a local public economy. At the same time, equity concerns can be, and ought to be, vigorously addressed by overlapping jurisdictions, metropolitan as well as state and national. Interjurisdictional fiscal transfers and local tax-base sharing may be more effective instruments for addressing equity concerns in the context of continued efficiency and responsiveness than expanded local government boundaries intended solely to create the possibility of intrajurisdictional transfers.

In addressing equity concerns, citizen choice and consent remain important principles. A highly efficient and responsive local public economy is better able to address the problems of distressed communities and is likely to offer better stewardship of funds received from both state and national governments. Greater equity depends upon both self-reliance and a willingness to share. Citizens in distressed communities may be especially needful of autonomous community organization, so as to be able to receive intergovernmental funding and use it to local community advantage.

6. STATE LEGISLATURES: TRUSTEES OF LOCAL GOVERNANCE

State legislatures are the trustees of the basic rules of local governance in America. The laws and constitutions of each state are the basic legal instruments of local governance. Legislatures must, by distinguishing provision and production, understand the need for variety and diversity. They should also understand that responsive local governance depends upon opportunities for citizen choice and a need to obtain citizen consent. Where there is little variety on the provision side—as in many large central cities—legislatures have the authority to create new potential for variety by enabling neighborhoods to incorporate, and should do so. Where there is little citizen choice and little separation of provision from production, state legislatures should closely examine the basic rules of local governance in order to open up new possibilities, and allow a more creative and efficient local public economy to emerge from the choices of local citizens and their elected officials.

Commission Recommendations*

Recommendation 1: **NONVIALE GOVERNMENTS**

The Commission rescinds its 1981 recommendation on the dissolution of "nonviable governments."

The 1981 recommendation is quoted in full below:

The Commission recommends that states, through a local government boundary commission, other state agency, or the state legislature, establish or supplement standards for local government viability (a) by requiring any local government, general or special-purpose, in the urbanized portion of a standard metropolitan statistical area (SMSA), to have the equivalent of at least one full-time employee, or, (b) by requiring general-purpose units to perform at least four functions, or only two functions, provided that each of the two constitutes at least 10% of the jurisdiction's current expenditure budget. If either of these standards is not met, the state, after offering adequate opportunity for a hearing for the affected local government(s), shall consider dissolving the local government and providing for the transfer to and performance of its functions by (an) appropriate unit(s) of general local government. [*State and Local Roles in the Federal System*, A-88, April 1982, pp. 446-47]

The Commission concludes that, while the viability of units of local government is an important issue, criteria pertaining to the number of full-time equivalent employees and the number of functions produced rather than provided are inappropriate standards for determining local government viability. Small units of government that make use of

part-time citizen-officials and that function as "pure provision" units, contracting for the production of most or even all services provided, may serve very useful purposes for their citizens. Lack of viability as a production unit does not necessarily indicate lack of viability as a provision unit.

Different criteria apply to organizing the *provision* of local public goods and services as distinguished from their *production*. The provision side involves financing, specifying performance standards, and representing the interests of citizen-consumers. The production side is concerned with making a product or delivering a service. Very small units of local government (under 1,000 in population) can serve useful purposes on the provision side, even when they do not function as production units. The ACIR study of St. Louis County, where 21 municipalities under 1,000 population are located, demonstrates the viability of very small units as organizations for the provision of services, while production is contracted to a variety of public agencies and private firms.

Small suburban municipalities, despite their legal status, tend to function effectively as "neighborhood governments," a concept endorsed by the Commission as recently as 1985. No evidence exists to support the view that small units necessarily impede the performance of areawide functions where those functions are assigned to an overlapping jurisdiction. To dissolve these local units of government without the consent of their citizens would violate both the democratic process and the historic tradition of American local self-governance.

Recommendation 2: **SPECIAL-PURPOSE GOVERNMENTS**

The Commission finds that special-purpose governments may be a useful and efficient form of organization for local citizens. The option of organ-

*Adopted unanimously by the Commission, June 5, 1987, in San Francisco, California.

izing a special-purpose government complements the organization of general-purpose local governments. The Commission finds no a priori reason to reduce the number of special-purpose governments or to restrict their growth arbitrarily. The utility of special-purpose governments is best judged on a case-by-case basis.

The Commission recommends, therefore, that states place no extraordinary limits or procedural burdens upon the creation of special-purpose governments by local citizens.

The utility of special-purpose governments derives from the inherent limitations of general-purpose governments. Municipalities tend to be mutually exclusive jurisdictions; their boundaries generally cannot overlap. Counties and townships may overlap municipalities, but have relatively fixed boundaries. Special-purpose governments have flexible boundaries and may overlap other units, including municipalities. The use of special-purpose governments enables local communities to obtain a closer match between the geographical scope of problems and the jurisdictional boundaries of a unit of local government organized to address those problems. Special-purpose governments also address the needs of citizens who desire only a limited range of services.

Recommendation 3:

BOUNDARY REVIEW AGENCIES

The Commission finds that many of the reasons

that once supported the idea that states ought to create either state or local boundary review agencies have been brought into serious question, both by events and by new research. The Commission recommends, therefore, that caution be exercised in the establishment and use of boundary review agencies, especially those empowered to veto the creation of local units of government by affected citizens.

Boundary review agencies are created for the purpose of regulating the establishment of new units of local government, as well as controlling boundary adjustments through annexation and consolidation. The premise often is that local citizens are not competent to make appropriate judgments about local government organization. Distinguishing provision units from production units, and recognizing that local governments are fundamentally provision units, suggests that local citizens may in fact be "the experts" on the number and variety of local government jurisdictions that ought to exist in a local area. Boundary review agencies ought not to be used by local government officials to stifle citizen initiative and forestall competitive pressures for efficient local government performance. The ability to create new units of local government, and give their consent to an expansion in the boundaries of existing units, are important citizen capabilities in the organization of local public economies.

Recent ACIR Publications

The 1986 Federal Tax Reform Act: Its Effect on Both Federal and State Personal Income Tax Liabilities , SR-7, forthcoming 12/87	\$5.00
The Organization of Local Public Economies , A-109, 64 pp.	\$5.00
Measuring State Fiscal Capacity, 1987 Edition , M-156, 12/87, 152 pp.	\$10.00
Significant Features of Fiscal Federalism, 1988 Edition, Vol 1 , M-155, 12/87, 128 pp.	\$10.00
Is Constitutional Reform Necessary to Reinvigorate Federalism? A Roundtable Discussion , M-154, 11/87, 39 pp.	\$5.00
Local Revenue Diversification: User Charges , SR-6, 10/87, 64 pp.	\$5.00
The Transformation in American Politics: Implications for Federalism , B-9R, 10/87, 88 pp.	\$5.00
Changing Public Attitudes on Governments and Taxes , S-16, 9/87, 64 pp.	\$10.00
Devolving Selected Federal-Aid Highway Programs and Revenue Bases: A Critical Appraisal , A-108, 9/87, 56 pp.	\$10.00
Estimates of Revenue Potential from State Taxation of Out-of-State Mail Order Sales , SR-5, 9/87, 10 pp.	\$3.00
A Catalog of Federal Grant-in-Aid Programs to State and Local Governments: Grants Funded FY 1987 , M-153, 8/87, 36 pp.	\$10.00
Fiscal Discipline in the Federal System: National Reform and the Experience of the States , A-107, 8/87, 58 pp.	\$10.00
Federalism and the Constitution: A Symposium on <i>Garcia</i> , M-152, 7/87, 88 pp.	\$10.00
Local Perspectives on State-Local Highway Consultation and Cooperation , SR-4, 7/87, 48 pp.	\$5.00
Summary of Welfare Reform Hearings—1986 , SR-3, 6/87, 31 pp.	\$5.00
Significant Features of Fiscal Federalism, 1987 Edition , M-151, 6/87, 304 pp.	\$15.00
Preliminary Estimates of the Effect of the 1986 Federal Tax Reform Act on State Personal Income Tax Liabilities , SR-2, 12/86, 16 pp.	\$5.00
Measuring State Fiscal Capacity: Alternative Methods and Their Uses, 1986 Edition , M-150, 9/86, 181 pp.	\$10.00
The Transformation in American Politics: Implications for Federalism , A-106, 8/86, 400 pp.	\$10.00
The Agricultural Recession: Its Impact on the Finances of State and Local Governments , SR-1, 6/86, 60 pp.	\$5.00
State and Local Taxation of Out-of-State Mail Order Sales , A-105, 4/86, 160 pp.	\$5.00
A Framework for Studying the Controversy Concerning the Federal Courts and Federalism , M-149, 4/86, 75 pp.	\$5.00
Devolving Federal Program Responsibilities and Revenue Sources to State and Local Governments , A-104, 3/86, 88 pp.	\$5.00

The reports of the Advisory Commission on Intergovernmental Relations are released in five series: the "A" series denotes reports containing Commission recommendations; the "M" series contains Commission information reports; the "S" series identifies reports based on public opinion surveys; the "B" series reports are abbreviated summaries of full reports; and the "SR" series are staff information reports. Reports may be obtained from ACIR, 1111-20th Street, NW, Washington, DC 20575.

What is ACIR?



The Advisory Commission on Intergovernmental Relations (ACIR) was created by the Congress in 1959 to monitor the operation of the American federal system and to recommend improvements. ACIR is a permanent national bipartisan body representing the executive and legislative branches of federal, state, and local government and the public.

The Commission is composed of 26 members—nine representing the federal government, 14 representing state and local government, and three representing the public. The President appoints 20—three private citizens and three federal executive officials directly and four governors, three state legislators, four mayors, and three elected county officials from slates nominated by the National Governors' Conference, the Council of State Governments, the National League of Cities/U.S. Conference of Mayors, and the National Association of Counties. The three Senators are chosen by the President of the Senate and the three Representatives by the Speaker of the House.

Each Commission member serves a two-year term and may be reappointed.

As a continuing body, the Commission approaches its work by addressing itself to specific issues and problems, the resolution of which would produce improved coopera-

tion among the levels of government and more effective functioning of the federal system. In addition to dealing with the all-important functional and structural relationships among the various governments, the Commission has also extensively studied critical stresses currently be placed on traditional governmental taxing practices. One of the long-range efforts of the Commission has been to seek ways to improve federal, state, and local governmental taxing practices and policies to achieve equitable allocation of resources, increased efficiency in collection and administration and reduced compliance burdens upon the taxpayers.

Studies undertaken by the Commission have dealt with subjects as diverse as transportation and as specific as state taxation of out-of-state depositories; as wide ranging as substate regionalism to the more specialized issue of local revenue diversification. In selecting items for the work program, the Commission considers the relative importance and urgency of the problem, its manageability from the point of view of finances and staff available to ACIR and the extent to which the Commission can make a fruitful contribution toward the solution of the problem.

After selecting specific intergovernmental issues for investigation, ACIR follows a multistep procedure that assures review and comment by representatives of all points of view, all affected levels of government, technical experts, and interested groups. The Commission then debates each issue and formulates its policy position. Commission findings and recommendations are published and draft bills and executive orders developed to assist in implementing ACIR policies