POLICY RESEARCH WORKING PAPER 1843

Externalities and Bailouts

Hard and Soft Budget Constraints in Intergovernmental Fiscal Relations

David E. Wildasin

A local government's ability to extract a "bailout" from a central government depends on how big it is. Larger localities may rationally expect bailouts — and thus operate under softer budget constraints. Effective fiscal decentralization requires an institutional structure that minimizes these adverse incentives.

The World Bank
Development Research Group
November 1997

Summary findings

Subnational governments are assuming greater fiscal responsibility in many developing and transition countries. There is concern, however, that fiscal decentralization may weaken fiscal discipline — that local authorities may undertake commitments or incur debt obligations that subsequently result in massive central government support, in the form of extraordinary transfers, or "bailouts." (Recent experience in major U.S. cities shows that these problems are not restricted to developing countries.) Such bailouts could in turn cause national fiscal imbalances, excessive borrowing, and macroeconomic instability.

Some analysts recommend that central authorities maintain strict control over the fiscal behavior of lower-level governments, but others argue that such controls could undercut the goals of fiscal decentralization, including autonomy.

Wildasin shows that central authorities may have strong incentives to prop up the finances of local governments when the public services provided locally benefit the rest of society. The prospect of such interventions may in turn create incentives for localities to underprovide services that produce substantial spillover benefits, using local resources instead for purposes that may benefit local constituencies but not nonresidents.

When central fiscal interventions are big enough, and when a loss of local control over the use of fiscal resources is not too costly to local residents, local decisionmakers will act to induce central government bailouts, resulting in inefficient outcomes for the system as a whole.

This is not to say that fiscal decentralization produces perverse incentives or requires central government control over local fiscal policies. But incentives for bailouts can be especially strong when local governments are considered "too big to fail" — for example, New York, Philadelphia, and Washington, DC (in the United States) and São Paulo and Rio de Janeiro (in Brazil). In such cases, the repercussions from major breakdowns in the provision of services — or in debt servicing — can be too costly for central governments to ignore.

Problems of fiscal discipline may result not because there is too much fiscal decentralization, says Wildasin, but because there is too little. It may make sense to carry out more thorough decentralization — for example, devolving fiscal authorities to smaller jurisdictions or special-purpose functional units, or subdividing large subnational jurisdictions into many smaller units.

This paper — a product of the Development Research Group — is part of a larger effort in the group to understand fiscal decentralization. Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Cynthia Bernardo, room N10-055, telephone 202-473-7699, fax 202-522-1154, Internet address prdpe@worldbank.org. November 1997. (36 pages)

EXTERNALITIES AND BAILOUTS

Hard and Soft Budget Constraints in Intergovernmental Fiscal Relations

by

David E. Wildasin
Department of Economics
Vanderbilt University
Nashville, TN 37235
USA

It is hereby declared to be a public policy of the Commonwealth ... to foster the fiscal integrity of cities of the first class to assure that these cities provide for the health, safety, and welfare of their citizens; pay principal and interest owed on their debt obligations when due; meet financial obligations to their employees, vendors, and suppliers; and provide for proper financial planning procedures and budgeting practices. The inability of a city of the first class to provide essential services to its citizens as a result of a fiscal emergency is hereby determined to affect adversely the health, safety, and welfare not only of the citizens of that municipality but also of other citizens in this Commonwealth.*

Mayor Marion Barry ... opposed the efforts of a member of Congress to provide $42 million to the D.C. police department. He said it would be unfair to single out the police department for additional funds when all city agencies are in financial straits.†

I. Introduction

In established federations such as the United States and Canada, lower-level governments – states, provinces, or municipalities – occasionally fall into serious fiscal crises in which expenditures are drastically out of balance with current and projected revenues, resulting in loss of access to the capital market and the prospect large and sudden reductions in service provision and employment or increases in local taxes. Such crises are relatively rare, but when they do occur, higher-level governments may intervene in the fiscal management of major cities, such as New York in the 1970s, or Washington, D.C. and Philadelphia, Pennsylvania at present. These interventions often involve the establishment of boards or commissions especially appointed for the task, possibly endowed with exceptional authority to cut spending, insure the continued provision of essential services, or fire, furlough, or reassign public-sector workers (Noto and Rymarowicz [1995]).

In many developing countries or in the transition economies, subnational government fiscal crises seem be much more commonplace. A thorough discussion of the experience of these countries goes beyond the scope of the present essay, but a brief review of a few important cases will help to reveal the nature and importance of the issue throughout the world. (See Wildasin [forthcoming] for further discussion and references.) In China, fiscal decentralization has accompanied the general economic decentralization of the past decade of economic reform. The entire fiscal system, including the tax system, the assignment of expenditure functions among levels of government, and the structure of intergovernmental fiscal relations, has been in a state of flux. The central government allocates significant amounts

* Pennsylvania Intergovernmental Cooperation Authority Act for Cities of the First Class, Laws of Pennsylvania, Act 1991-6, Session of 1991, Section 102. "Cities of the first class" in Pennsylvania are those with populations in excess of 1.5 million. Philadelphia is the only city in this class; the second largest city in the state is Pittsburgh, with a population of less than .4 million.

† "As D.C. Police Struggle On, Change Pays Off in New York" by R. Castaneda (pp, A1, A8), *The Washington Post*, Mar. 30, 1996.

of resources to lower-level governments through intergovernmental grants. Competing demands by regional authorities for additional fiscal assistance from the central government have even led to use of "loans" from the central bank to lower-level governments, leading to expansion of the money supply and, in effect, to direct inflationary finance of lower-level governments (The World Bank, 1995a). In India, the other giant of the developing world, a number of state governments are in fiscal crisis, accumulating significant liabilities to the central authorities in the form of intergovernmental borrowing (The World Bank, 1995b).

Intergovernmental fiscal relations in several Latin American countries such as Argentina and Brazil are similarly characterized by problems of financial solvency for state or local governments leading to stopgap lending or grants from higher-level governments. In the Brazilian case, for example, a number of the country's major commercial banks are owned by state governments, and the states themselves have borrowed heavily from these banks to finance their public expenditures. The magnitude of this indebtedness is large and has grown rapidly (from about US $57 billion in 1991 to about US $110 billion in 1995, which may be compared to Brazil's total external debt in 1991 of about US $120 billion). In the face of higher real interest rates, some states (notably Sao Paulo, whose debt amounts to almost half of all state government debt in Brazil) have stopped servicing their debt, pushing these major banks into financial insolvency. In order to avoid a general financial crisis, the central bank of Brazil is assuming control of these banks and, in the process, may end up absorbing large amounts of state government debt. In effect, then, the fiscal crises of some state governments (particularly the larger and more prosperous states such as Sao Paulo and Rio de Janeiro) will have been resolved through special (and far from transparent) interventions by central monetary and fiscal authorities.

In Russia, as in China, economic and political reforms coupled with strong regional tensions have resulted in a major restructuring of fiscal systems, including the fiscal relations among governments. Here, too, the central government has been unsuccessful in establishing an orderly and enforceable system for allocation of fiscal resources between the central and lower-level governments; indeed, in circumstances reminiscent of the United States under the Articles of Confederation (and of the Soviet Union prior to its collapse: Bahl and Wallich [1995], McLure *et al.* [1995]), regional governments have unilaterally withheld all or part of the taxes collected in their territories, even when specified shares of these revenues are supposed to flow to the central government (The World Bank [1996a, b]).

The examples of these and other countries raise doubts about the appropriateness and viability of fiscal decentralization, especially in developing and transition economies. Lower-level governments seem to find themselves in recurring fiscal crises that necessitate, or at least elicit, *ad hoc* interventions by central governments. "Fiscal discipline" or "hard" budget constraints appear to be lacking, and inefficient, inequitable and simply chaotic public policies may result. Unanticipated transfers from central governments or central banks to lower-level governments may undermine macroeconomic and price stability. Although fiscal decentralization has played a critical role in the reform of the public sector throughout

the world in recent years, it might be desirable to reverse this trend if it contributes to dysfunctional fiscal performance. Indeed, some analysts and policy advisers (*e.g.*, Bird *et al.* [1995], Prud'homme [1995], Tanzi [1996]) recommend re-establishment of central government control of the fiscal affairs of lower-level governments through limitations on their fiscal autonomy (e.g., constraints on local borrowing authority) or through reversals of recent fiscal decentralization initiatives. As is widely appreciated by those involved in this debate, however, the stakes are high. Retrenchment of central planning and fiscal controls is, after all, a major element in the process of economic liberalization that has characterized economic and political reforms in many countries in recent years, and many would hesitate to sacrifice the potential benefits of the transition to greater economic and political decentralization of which fiscal decentralization is a part. [1]

In a policy context, it is possible and probably desirable to argue that *ad hoc* fiscal relations between governments characterized by bailouts and soft budget constraints *ought* to be avoided and that these relations *ought* to be characterized by transparency, accountability, and predictability. While the value of normative policy advocacy or "moral suasion" should not be underestimated, it is also useful to try to ascertain what institutional structures or fundamental economic conditions may or may not be conducive to desired policy outcomes. [2] Indeed, since issues of fiscal decentralization and intergovernmental fiscal relations are often settled, in significant part, through the formal mechanism of written constitutions, it seem especially appropriate to analyze them from the "constitutional" perspective of the public choice school, which views the actual realizations of policy as the equilibrium outcome of an institutional framework for public sector decisionmaking. Exhortations by policy advisers to change soft budget constraints into hard ones are ultimately likely to be of limited effectiveness if institutions are structured in such a way that policymakers find *ad hoc* fiscal interventions to be irresistibly attractive. What institutional or other factors drive policymakers to soften budget constraints in some cases and enable them to resist pressures

[1] For further discussion of recent trends toward fiscal decentralization in Europe, Latin America, China, and elsewhere, and of the policy debates surrounding these trends, see Bird (1994), the InterAmerican Development Bank (1994), Owens and Panella (1991), and Roy (1995) in addition to the sources cited above. Ter-Minasian (1996) reviews the institutions and experience of subnational borrowing for several developed and developing countries, finding that market discipline helps to establish incentives for effective borrowing in countries with well-developed financial markets but administrative controls or rules imposed by higher-level authorities may be needed to constrain local borrowing.

[2] For example, Ter-Minasian (1996, p. 16) notes: "In countries with a history of bailouts of insolvent subnational governments by the central government, a firm and sustained re- f o engage in further operations of this kind will be necessary to change expectations and behaviors of market participants." Similarly, Tanzi (1996, p. 310) writes that "decentralization will contribute to macroeconomic stability ... as long as all these possibilities [of subnational borrowing] exist, as long as there is the belief that the central government will honor subnational governments' obligations, and as long as the incentives for these governments encourage higher spending." While these remarks suggest that central governments have it within their power to avoid adverse fiscal outcomes by acting firmly with respect to subnational governments, they also suggest that central governments have often chosen not to act in that way. What, then, can be done to change the nature of the process that leads central and subnational governments to act in ways that lead to adverse outcomes?

for bailouts in other cases?

No simple economic analysis can hope to address every important aspect of this very complicated issue. When considering the fiscal crises of major American cities like New York, Washington, DC, or Philadelphia, Brazilian states like Sao Paulo, or Indian states like Orissa, it is natural to look for specific triggering events such as high unemployment rates, fluctuations in real interest rates, the flight of tax bases from central cities to suburbs, or corrupt political administrations. [3] When attempting to model the fiscal relations between central and lower-level governments, it is natural to think of "bailouts" as stages in a repeated game, and to think of central governments as facing a tradeoff between beneficial short-run interventions and long-run reputation and incentives, rather like creditors in private or international capital markets or like central bankers concerned with market expectations. These considerations are very likely to be crucial elements of the complete story of soft budget constraints for lower-level governments. [4]

The present paper, however, offers a somewhat different perspective on the problem of soft budget constraints. It may be motivated by asking why some lower-level governments fall into fiscal crises that seem to require interventions by higher-level governments while others do not. For instance, in the United States, there are many thousands of local governments; they vary widely in size, function, and administrative organization and serve regions of extremely diverse demographic and economic composition and circumstances. Typically, a large share of the expenditures of these localities is financed or subsidized by transfers from higher-level governments which may be explicit (grants) or implicit (Federal income-tax deductibility of many local taxes). However, these transfers are systematic in nature and are not triggered by the fiscal crises of individual localities. Most localities do not face problems in servicing their debt, do not receive special infusions of funds from state

[3] See Inman (1995) for a clear discussion of the Philadelphia case.

[4] A series of recent contributions investigate how soft budget constraints rise for state-owned enterprises and how decentralization of capital-market institutions may harden their constraints; see, e.g., Qian (1994), Dewatripont and Maskin (1995), and von Thadden (1995). In a somewhat similar vein, Crémer (1995) emphasizes that restriction of information flows between principals and agents, such as might result from vertical disintegration or greater decentralization of borrower-lender relationships, may enable principals to discipline agents more credibly. What might be called "trapped creditors" play an important role in some of these analysis, that is, lenders who have already invested some capital in a project and who, in the face of these sunk costs, may decide to invest still a bit more in a bad project in order to see it through to completion, even though they would have chosen to avoid it *ex ante*. There are undoubtedly useful applications or variations of these ideas that could (and, in future research, should) be developed in the setting of intergovernmental fiscal relations. The approach taken here, however, is basically different in its emphasis on the role of interjurisdictional externalities as a contributing factor to soft budget constraints. By analogy to trapped creditors, the model developed below relies on intrinsic externalities to forge a linkage between any one lower-level jurisdiction and the rest of the economy which, in a sense, may "trap" the rest of the economy into having an interest in the decisions made within a locality. As Qian (1994) observes, although "paternalism" plays a significant role in the writings of Kornai (1986) on soft budget constraints, this concept has no obvious counterpart in pure investment models. The externalities that appear in the present analysis may capture something of the "paternalistic" argument.

legislatures and do not fall under the jurisdiction of state-appointed financial control boards. These elementary observations shows that fiscal decentralization need not, in and of itself, give rise to recurring fiscal crises, soft budget constraints, or bailouts.

Indeed, fiscal decentralization might actually make soft budget constraints *less* rather than *more* likely. Informally speaking, one can imagine a continuum of economic and political organizational forms representing varying degrees of centralization. Completely centralized government (perhaps even at the world level) would represent one extreme of this continuum. Moving successively toward greater decentralization along this continuum, one would encounter state and provincial governments, general-purpose local governments such as counties or municipalities, and special-purpose local governments such as school districts, or water authorities. Entities of still smaller geographical scope and greater functional specialization lie still further in the direction of increased decentralization, gradually crossing a rather diffuse boundary between the public and the private sector. For instance, local refuse collection may be contracted out by local governments to private companies; zoning and other land-use controls provide very limited and specific local public constraints on uses of land which remains fundamentally under the control of private agents. Public, quasi-private and private services may exist side-by-side in this part of the continuum, as often occurs with recreational and sports facilities or for-profit, not-for-profit, and public hospitals. At the complete decentralization extreme of the continuum are individual firms and households in their private spheres of activity. In a mixed economy, some economic activity is organized by entities at almost every point along this continuum, and each such entity normally faces formal constraints on its financial flows. As one moves along this continuum, is it really the case that greater decentralization leads to softer budget constraints? Central governments typically exert less direct financial control over local governments than state or provincial governments and still less control over individual households and firms, but this control does not seem essential to establish fiscal or market discipline over very decentralized activities. Is it possible that the fiscal crises observed in some countries could be the result of too *little* decentralization rather than too *much* decentralization?

The present essay attempts to shed light on these issues by developing a model of interactions between central and lower-level governments in which there is a meaningful distinction between "hard" and "soft" budget constraints and in which it is possible to identify conditions likely to result in one or the other. Of course, the mere fact that the fiscal affairs of higher- and lower- level governments are interrelated, for instance by fiscal transfers from the center to lower-level governments, does not imply that lower-level governments face "soft" budget constraints. The budget constraints facing lower-level governments may be altered by central government policies such as fiscal transfers but as long as the transfer recipients perceive these constraints to be binding it is natural to characterize them as "hard." The notion of a "soft" budget constraint is that circumstances may transpire in which the budget constraint facing a locality [5] is altered in some contingent fashion allowing

[5] For ease of reference, lower-level governments will often be referred to below as "locali-

an outcome to occur which would not have been attainable under the "normal", "initial", or "announced" constraint. The term "bailout" may be apt in this context; "receivership" is another term that seems to capture part of the flavor of the phenomenon. In general, the fiscal crises of lower-level governments seem to give rise to varying combinations of added central government control (loss of local "autonomy") and added fiscal assistance to local governments ("bailouts"), and the model presented here provides a framework within which these elements are clearly revealed.

A first basic premise of the model is that the fiscal affairs of an individual locality are not a matter of indifference to the "central government", interpreted here as an agent representing the interests of the country as a whole and thus of those who reside outside of the given locality. Formally, it is assumed that local governments provide public goods which not only benefit their own residents but that produce external benefits for residents of other localities as well. Because of these externalities, the central government, acting in the interest of the whole society, may (and, in the model, does) establish programs of intergovernmental transfers that function as Pigovian corrective subsidies. A second basic premise of the model, however, is that there is a sequential structure to decisionmaking. The center moves first in establishing a program of corrective intergovernmental transfers. This policy establishes a putative budget constraint for localities, which then choose their levels of local taxes and expenditures. The story does not end there, however. In particular, the center can move last, after observing local fiscal decisions, by taking direct control over local expenditures and by financing incremental local spending from central funds. It is this third-stage fiscal intervention by the center which corresponds to a "bailout" and which, if it occurs, reveals the local government budget constraint to have been "soft."

A major goal of the analysis is to identify the conditions under which the center does or does not intervene at the third stage. Section II presents the basic structure of the model and recapitulates some of the standard principles of intergovernmental fiscal relations in the presence of interjurisdictional externalities, Section III models the use of bailouts by the central government. It shows that the externalities generated by local public goods may make such interventions attractive from the viewpoint of the rest of society, but that there are also conditions under which the center would not intervene, in which case localities definitely face *hard* budget constraints. The analysis in Section III also indicates that the hardness of local budget constraints is partly in the hands of the localities themselves. In particular, there are conditions under which localities could take actions that would elicit bailouts from the center, but in which the payoff to local residents in the event of a bailout is sufficiently unattractive that the locality does not act in such a way as to trigger central intervention in the third stage. In this case, too, the local budget constraint is hard in the sense that a third-stage intervention would not be observed, highlighting the fact that bailouts and soft budget constraints emerge from the actions of both central and local governments.

---

ties," but this term should be understood to embrace jurisdictions such as states or provinces as well.

The model is highly stylized in order to bring out clearly some basic insights, and one aspect of this stylization is the assumption that all localities are of identical size, i.e., they contain equal numbers of residents. Given a fixed national population size, the size of local government can be taken as a parameterization of the degree of fiscal decentralization: when the national population is divided into only a few large jurisdictions, the degree of fiscal decentralization is low, while a high degree of decentralization is represented in the model by small and numerous jurisdictions. Recalling the notion of a centralization/decentralization continuum mentioned earlier, it is natural to ask how jurisdictional size affects the hardness of local budget constraints. Section IV addresses this issue by carrying out comparative statics analysis of the effect of jurisdictional size on the willingness of the central government to intervene in local fiscal affairs and on the incentives that localities have to induce such interventions. It is necessary to impose additional structure on the model in order to make this analysis tractable, but, subject to the limitations of additional simplifying assumptions, the analysis indicates that budget constraint do indeed tend to be soft for large localities and hard for small ones.

Section V discusses some of the policy implications of the analysis and identifies directions for further research.

II. Local Public Good Provision, Spillovers, and Intergovernmental Grants

A. The Basic Model

Household Preferences, Endowments, and Externalities. Suppose that the total population in the economy is N. In order to simplify the notation and in order to demonstrate that the results do not depend on asymmetries in household attributes, all households are assumed to have identical preferences and endowments. Each household directly consumes three commodities, x, z and G. The first is interpreted as an all-purpose private good which will serve as numéraire. Each household is endowed with w units of this commodity. Consumption of the second good, which may be thought of as health, water, or education, yields external benefits to other households. The third commodity, G, is a Samuelson public good whose level of provision, as described below, is determined by a central government.

The external effects associated with consumption of good z creates an interdependency among households which is crucial in the following analysis. The precise structure of these external effects could take various forms without affecting the results substantially. However, given the intended applications, it is desirable for this structure to incorporate three elements. (i) External effects should be to some degree additive and anonymous. Each household may care about the total number of vaccinations given to other households (the total number of households with access to potable water, the total basic literacy rate), etc., but may be more or less indifferent as to exactly which households obtain the vaccination; in other words, any one household's preferences exhibit a high degree of substitutability in the external benefits generated by other households' consumption of good z. (ii) The external

effects associated with any one household's consumption of good z should be characterized by diminishing returns. While there may be a large external effect associated with meeting "basic needs", such as vaccinations for communicable disease (installation of standpipes, basic literacy, etc.), the external benefits associated with treatment of injuries (residential water hookups, secondary education) may be much smaller, while cosmetic surgery (availability of water for watering of lawns, some forms of university education) may yield almost no external benefit at all. Thus, in particular, a given amount of resources spent on health care (water supply, education) will produce more external benefits if distributed relatively equally among the population. (iii) Household preferences should be characterized by a diminishing marginal rate of substitution between external benefits and consumption of other goods, i.e., a household's marginal willingness to pay for the health care (water availability, education) of others diminishes as their consumption level rises.

To capture these features, suppose that each household h has a strictly quasi-concave utility function $u(x_h, z_h, Z, G)$ where $Z \equiv \sum_{h'} \beta(z_{h'})$, with $\beta' > 0 > \beta''$. Here, (x_h, z_h) is household h's own-consumption of goods x and z; the third argument of the utility function, Z, reflects the external benefit to h of the consumption of good z by other households. [6] The concavity of the function β reflects the "basic need" aspect of the external effects (item (ii) above). The fact that the external effects are summed across households reflects the substitutability of the externalities produced by different households (item (i)). The strict quasi-concavity of $u(\cdot)$ reflects the diminishing marginal returns to aggregate externalities (item (iii)). It is also assumed that no goods are inferior.

Government structure, technology, and finance. There is a fixed total population N and a single central government. The population is partitioned into local jurisdictions which, for the sake of notational simplicity, are assumed all to contain the same number of households, n; households are immobile among jurisdictions. Since the analysis focuses on the relationship between a single jurisdiction and the central government, the assumption that all localities are of the same size is not critical, but this and other symmetry assumptions simplify the notation and exposition. [7] Each locality determines a level of good z to be provided to each of its residents, i.e., z is treated as a local public good. In this and the next section, all of the analysis is based on the assumption of an exogenously-fixed size for local jurisdictions, and the extent to which the cost of provision of good z depends on population size – i.e., the extent to which this good is subject to crowding or congestion effects – is therefore irrelevant. Section IV, however, is devoted to an analysis of the effects of jurisdictional size, and for the purposes of that analysis it is useful to abstract from economies or

[6] Since it eases the notation slightly to do so, household h's own consumption of good z is assumed to contribute a term $\beta(z_h)$ to the summation over h' in the utility function. This assumption is inconsequential for the analysis, however.

[7] The total number of localities in the economy, N/n might not be an integer. Since the basic analysis applies even with asymmetric jurisdictions, however, any departures from an integer number of localities could be accommodated by allowing for a "remainder" locality of size N mod n.

diseconomies of scale in the provision of the local public good so that the effects of jurisdictional size in the analysis are "uncontaminated" by technological considerations. For the sake of notational economy and consistency, therefore, let us assume at the outset that z is a "quasi-private" local public good, that is, the total cost of providing a given number of units of the good to each of a jurisdiction's residents is proportional to the number of households residing in locality. For simplicity, units are chosen so that each unit of this good has a cost of 1. The total cost of providing each resident in a locality with z units of the good is thus nz. [8]

In order to abstract from the complications associated with distortionary taxation, all localities and the central government are assumed to have uniform lump-sum taxes at their disposal to finance their expenditures. Local public good provision may be paid for either from local taxes or with grants from the central to local governments. The role of grants in the analysis will be to affect the level of local provision of a good, namely good z, that generates spillover benefits for the rest of the society. As is customary in the literature, two types of grant instruments are considered: lump-sum and matching conditional grants. Let m denote share of local expenditures reimbursed by the center through matching grants; on the grounds of symmetry, the matching rate is assumed to be the same for all localities. Let g_i denote any lump-sum grant from the center to locality i, expressed in per capita terms. Although this lump-sum grant may be uniform across localities, the notation distinguishes the lump-sum grant level by locality because, as will become clear, we want to consider the possibility that the center may use lump-sum grants to intervene specifically in the financing of local public good provision in a particular locality. Let c_i be the per-capita level of local "own-contributions" to the provision of the local public good in locality i, i.e., c_i denotes the amount paid in local taxes paid by each resident of the locality. For the purposes of the discussion in the present section, it is assumed that matching and lump-sum grants are not used in conjunction with one another. The level of local public good consumed by each resident of locality i is therefore

$$z_i = c_i + g_i \qquad (1L)$$

when the central government provides lump-sum grants or

$$z_i = c_i/(1-m) \qquad (1M)$$

[8] It is certainly reasonable to ask why quasi-private goods should be provided through the public sector. One possible explanation is that for some goods and in some institutional contexts (e.g., under court decrees about educational access), it may be infeasible to exclude consumers on the basis of price. It is not the purpose of the present discussion, however, to evaluate the normative justification for public provision of quasi-private local public goods. Rather, let us merely observe that many of the goods provided by local governments – education is an obvious example – are certainly congestible and that the cost of provision is indeed roughly proportional to jurisdictional size once minimum efficient scale is achieved. The case of quasi-private local public goods is therefore empirically relevant. As noted, it is also analytically helpful since it means that the results of Section IV do not depend on the presence or absence of scale effects in local public good provision. The implications of scale economies for the issue of optimal jurisdiction size are discussed briefly in the conclusion.

under matching grants. Letting T denote the amount of central government lump-sum tax imposed on each household, the level of good x consumed by a resident of locality i is the level of endowment net of local and central government taxes, i.e.,

$$x_i = w - c_i - T. \tag{2}$$

The menu of bundles of local private and public good consumption available to residents of locality i must therefore satisfy the budget constraint

$$x_i + z_i = w - T + g_i \quad z_i \geq g_i \tag{3L}$$

when lump-sum grants are used (the inequality in (3L) reflects the fact that the grants are conditional in nature) or

$$x_i + (1-m)z_i = w - T \tag{3M}$$

if matching grants are used. The constraints (3L) and (3M) combine the household budget constraints (2) and the local government budget constraints (1L) and (1M) by elimination of the local own-contribution variable c_i. Note that lump-sum grants act like a pure income transfer to the residents of a locality, whereas matching grants lower the relative price of the local public good.

A system of grants and central government taxation is feasible only if it satisfies the central government budget constraint. Summing total fiscal transfers across all localities i, this constraint takes the form

$$\sum_i n g_i + G = NT \tag{4L}$$

under lump-sum grants and

$$m \sum_i n z_i + G = NT \tag{4M}$$

under matching grants.

B. Grants and Efficiency with Hard Budget Constraints

The standard normative theory of intergovernmental grants in the presence of inter-jurisdictional externalities applies the principles of Pigovian welfare economics to intergovernmental fiscal relations (Oates [1972]). Localities are treated like households that optimize subject to their budget constraints. If local public goods produce spillover benefits, equilibrium allocations will be inefficient in the absence of interventions by a higher-level government since each locality ignores the external benefits of its expenditures. Matching grants lower the relative price of local public expenditures and can be used as Pigovian corrective subsidies to induce efficient local spending. The model presented above has been deliberately structured to conform to the basic suppositions of this standard theory, and a concise formal restatement of the essential elements of the theory is useful as a point of reference for the analysis to follow.

Equilibrium. First, since the residents of each locality are identical, it is natural to assume that local policies are chosen to maximize the welfare of these residents, a behavioral hypothesis that it is compatible with simple majority voting or other plausible public choice mechanisms. Each locality is assumed to take as parametrically-given the fiscal instruments of the central government, a critical assumption that is reconsidered in Section III. Because of spillover effects, the most-preferred level of local public good provision in any one locality depends on the level of provision undertaken by other localities, giving rise to strategic interactions among local jurisdictions. It is customary to suppose that localities achieve a Nash non-cooperative equilibrium in levels of local public good provision. [9] Provided that grant policy treats all localities uniformly, the simplifying assumptions of the model make it natural to focus on symmetric Nash equilibria. To define such an equilibrium formally, let $\bar{z}$ denote the per capita level of z provided to the residents of all localities other than i. Locality i is assumed to solve the problem

$$(P) \qquad max_{<x_i,z_i>} \quad u(x_i, z_i, (N-n)\beta(\bar{z}) + n\beta(z_i), G)$$

subject to (3L) or (3M). (Notice that the specification in (P) reflects the fact that each resident in locality i derives an external benefit from the provision of good z to n fellow residents as well as from the provision of good z by other localities to their $(N-n)$ residents.) Given specified values of the central government fiscal instruments, a level of local public good provision $\bar{z}$ is defined to be a *Nash non-cooperative equilibrium with lump-sum (respectively, matching) grants* if $z_i = \bar{z}$ at a solution to (P), subject to (3L) (respectively, (3M)) for all localities i and if the central government budget constraint (4L) (respectively, (4M)) is satisfied when $z_i = \bar{z}$ for all localities i. [10]

Optimality. Let us now define and characterize an optimal allocation of resources for this economy in order to establish a benchmark for normative evaluation of alternative policies and the equilibrium outcomes they generate.

Since households have identical preferences and endowments and since localities are of equal size and have identical technologies, it is natural to focus on allocations of resources such that all households attain identical consumption bundles and utility levels. Given the strong symmetry and convexity assumptions of this model, it is obvious that there is a unique (x^*, z^*, G^*) that maximizes the common utility of all households subject to the fundamental resource constraint

$$\sum_h x_h + \sum_h z_h + G \leq Nw. \qquad (5)$$

[9] As discussed, e.g., in Wildasin (1991), Nash equilibria in local expenditures generally differ from Nash equilibria in local taxes. This distinction is partly obviated in the present model since localities use lump-sum taxes. In the presence of matching grants, however, Nash equilibria in levels of local expenditure, z_i, generally differ from Nash equilibria in levels of own-source expenditure, c_i. While the specification of the strategy space is in general a matter of considerable importance, it is not discussed further here.

[10] Technical issues of existence and uniqueness of Nash equilibria in this type of model have been discussed at some length in the literature but are suppressed here for brevity.

Note that since there are no economies of scale in provision of good z, jurisdictional size is irrelevant in the determination of efficient allocations. Defining $Z^* \equiv \sum_h \beta(z^*)$, this first-best allocation of resources is characterized by the Samuelson-type conditions

$$\frac{u_z(x^*, z^*, Z^*, G^*)}{u_x(x^*, z^*, Z^*, G^*)} + N\frac{u_Z(x^*, z^*, Z^*, G^*)}{u_x(x^*, z^*, Z^*, G^*)}\beta'(z^*) = 1 \tag{6.1}$$

$$N\frac{u_G(x^*, z^*, Z^*, G^*)}{u_x(x^*, z^*, Z^*, G^*)} = 1 \tag{6.2}$$

where derivatives of $u(\cdot)$ are denoted by subscripts. Note for future use that (4) implies

$$\frac{u_z(x^*, z^*, Z^*, G^*)}{u_x(x^*, z^*, Z^*, G^*)} < 1 \tag{7.1}$$

$$N\frac{u_Z(x^*, z^*, Z^*, G^*)}{u_x(x^*, z^*, Z^*, G^*)}\beta'(z^*) < 1 \tag{7.2}$$

provided that, at an optimum, the local provision of good z does generate some spillovers (i.e. provided that $u_Z(\cdot)\beta'(\cdot) > 0$) and provided that there is some purely local benefit from the provision of good z (i.e., $u_z(\cdot) > 0$).

Optimal Grant Design. It is well known that Nash equilibria, as defined above, are generally not efficient. In a somewhat degenerate sense, lump-sum grants offer one means by which central governments can induce efficient decisions by local governments. Specifically, the central government can impose a tax of $T_L^* = z^* + G^*/N$ on each household and use this tax to offer a lump-sum grant of $g^* = z^*$ per capita to each jurisdiction. Given this level of lump-sum grants, no locality wishes to impose any local taxes in order to increase its level of provision of good z. [11] In this case, all local provision of good z is financed by grants from the center, so that there is no own-source revenue collected by any locality and the fiscal system is in effect fully centralized.

More interestingly, it is possible to use matching grants to achieve an efficient Nash equilibrium in which there is still some meaningful degree of fiscal decentralization. The first-order conditions for (P) which must hold in a Nash equilibrium with matching grants imply that each locality chooses a level of local public good provision z' such that

$$\frac{u_z(x', z', Z', G^*)}{u_x(x', z', Z', G^*)} + n\frac{u_Z(x', z', Z', G^*)}{u_x(x', z', Z', G^*)}\beta'(z') = 1 - m \tag{8}$$

where x' is endowment net of local and central taxes. Following standard Pigovian principles, if the center selects the matching rate

$$m^* = (N - n)\frac{u_Z(x^*, z^*, Z^*, G^*)}{u_x(x^*, z^*, Z^*, G^*)}\beta'(z^*) \tag{9}$$

[11] In fact, each locality would prefer to reduce its expenditures on good z below $z^* = g^*$ but this is impossible by virtue of the assumed conditionality of these grants.

and a level of taxes $T_m^* = m^* z^* + G^*$ per capita, the equilibrium condition (8) coincides with the optimality condition (6.1) and the Nash equilibrium is efficient.

The essential features of this analysis are illustrated in Figure 1, which depicts the choices between x_i and z_i facing a particular locality i, taking as given the levels of local public good provision in other jurisdictions. In the absence of central government taxes and grants, the menu of attainable own-consumption bundles (x_i, z_i) is given by the 45-degree line AB, where the distance $OA = w$. With central taxes of T_m per capita and a matching rate of m, the budget constraint becomes the flattened line $A'B'$, where the distance $OA' = w - T_m$. The vertical distance between $A'B'$ and AB represents the difference between the matching grants received by locality i and the taxes that it pays to the central government. Only if locality i chooses the consumption bundle E at which the two constraints cross will the amount of grants received be equal to the amount of taxes paid. The central government budget constraint $(4M)$ requires that localities choose points like E "on average" or in aggregate; in the present model, with identical preferences and endowments for all households, an efficient Nash equilibrium with matching grants is characterized by identical levels of public good provision in all localities and thus each locality must choose a consumption bundle like E, such that it is neither a net contributor to nor a net recipient from the central government. In Figure 1, U_i^E is a projection of a level curve of locality i's utility function onto the (x_i, z_i) plane. The locality will choose point E on the budget constraint $A'B'$ if this indifference curve is tangent to $A'B'$ at E, as illustrated. If central government taxes and matching rates are at their optimal levels of T_m^* and m^*, then each locality faces a budget constraint $A'B'$ that intersects AB at a level of $z_i^e = z^*$, the efficient level of provision of good z. Moreover, given that all other localities provide z^* units of good z, the indifference curves of locality i projected onto the (x_i, z_i) plane will be such that there is a tangency between an indifference curve and $A'B'$ at point E and locality i will therefore choose $z_i^e = z^*$.

To summarize the basic lessons of the standard theory, a program of lump-sum grants produces income effects but no relative price or substitution effects and is thus a poor instrument for central governments to use in attempting to internalize the external benefits associated with the *decentralized* provision of local public goods. On the other hand, it is possible to achieve efficient outcomes with matching grants while still preserving meaningful fiscal decentralization, since local own-contributions to local public expenditures are generally positive in equilibrium. Moreover, under a system of matching grants, the central government does not directly control the level of local public good provision; rather, it alters the budget constraint facing a locality, leaving the choice of a consumption bundle along that constraint in the hands of local decisionmakers.

This standard theory has several limitations. For instance, it can be criticized for ignoring the possibility that localities might engage in direct negotiations ("Coasian bargains") to arrive at efficient outcomes or for assuming that the central government can ascertain the magnitude of interjurisdictional spillovers and thus to calculate appropriate matching grant

rate given in (9). It also abstracts from the complexity of local public-sector decisionmaking by in effect treating each locality as a single representative agent that optimizes subject to its budget constraint; this assumption is comparatively innocuous when all residents are identical, as assumed here, but it can be quite problematic in the more realistic case where residents are heterogeneous. Despite these and other limitations, however, the standard model provides a simple and familiar starting point for the present analysis.

III. Credibility and Bailouts

A. The Basic Rationale for Bailouts

One key feature of the conventional analysis of intergovernmental grants is the assumption that localities take the parameters of central government grant policies as given. In game-theoretic terms, one might describe the central government as a Stackelberg leader or first mover, with localities adapting as well as they can to the policies announced to them. In practice, however, the central government may not be able to enforce a commitment to its announced policies; indeed, a departure from announced constraints seems to be an essential feature of "bailouts" and "soft" budget constraints. A more elaborate sequential structure of actions by central and local governments is needed.

To illustrate the basic issues within the context of the model developed above, suppose that the center offers each locality a matching rate of m^* for its expenditures on good z which would support an efficient level z of per capita expenditures. Imagine, however, that some locality sets $z_i = 0$, for example by failing to provide even basic services such as inoculation of residents against communicable diseases. This could be the result of very low local tax effort or of a diversion of local public funds to other goods and services that produce no external benefits. [12] Under these circumstances, it may not be rational, and therefore not credible, for the central government simply to accept locality i's decision not to provide good z. This question cannot be answered without some theory about how the central government formulates its policies, but to anticipate the discussion below, observe that the residents of other jurisdictions are certainly not indifferent to the level of z_i. In fact, given that locality i contributes nothing toward provision of good z, the residents of other localities might well be willing to pay something to raise z_i above 0. Thus, if the central government responds to the interests of residents in *other* jurisdictions, it might intervene by directly controlling the provision of good z in locality i, for instance by using its own resources for this purpose. Such an intervention could appropriately be described as a "bailout" for locality i.

Depending on the generosity of the prospective bailout, it might or might not be attractive to locality i to induce the center to trigger it by setting $z_i = 0$. If the prospect

[12] To keep the notational burden to a minimum, no other local government goods and services have been explicitly specified in the model. Any quasi-private local public goods that do not generate benefits to non-residents may be subsumed within the all-purpose good x, however.

of a bailout is unattractive, locality i would prefer to set $z_i = z^*$ and accept matching funds from the center, as in the standard theory. In this case, the center presents locality i with a budget constraint of the form (3M) and the locality chooses a consumption bundle satisfying that constraint. We may then say that the budget constraint is *hard*. On the other hand, suppose that the central government sets a high level of provision of good z if locality i chooses not to use its own resources for this purpose. Then the welfare of locality i's residents may actually be increased by not funding the provision of good z at all, thus inducing a central government bailout. When this occurs, the locality induces the center to reveal that the budget constraint (3M) is not actually credible. In particular, (3M) indicates that zero own-contributions to the provision of good z should result in no transfers from the center, but in fact a zero level of own-contributions actually does elicit central government financial support for the locality. The constraint in this case can be described as *soft*. In game-theoretic language, one could say that the hard budget constraint equilibrium postulated in the traditional approach is not subgame perfect.

The objective of the present section is to develop a formal model which makes precise these intuitive notions of hard and soft budget constraints and which allows a determination of conditions under which bailouts do and do not occur.

B. A Formal Model of Bailouts

It is sufficient to focus on the relationship between a single locality i and the central government. Suppose that the center announces an optimal matching rate m^* as defined in (9) and that all other localities choose first-best levels of own-contributions to the provision of good z, $c^* \equiv (1 - m*)z^*$, thus achieving the optimal provision level z^*. The question is whether locality i chooses also to set $c_i = c^*$ and accept the central matching grant, in which case there is no bailout, or instead chooses some other level of own-contributions in anticipation of a central government bailout. The locality's choice obviously depends on the bailout policy of the central government, which describes what level of bailout a locality receives if it deviates from the first-best contribution level c^*. Since the central government determines whether or not to bail out locality i after observing its choice of c_i, the relationship between the locality and the central government has a sequential structure. We therefore analyze this relationship recursively, considering first the formulation of the central government's bailout policy.

Central Government Bailout Policy.

Let the level of own-contributions chosen by locality i be denoted by c_i', which may or may not be equal to c^*. Assume that the central government can intervene in locality i's provision of good z after observing c_i' by adding a conditional lump-sum grant of g_i' to locality i's own contribution, resulting in a level of provision denoted by $z_i' = c_i' + g_i'$. In order to do this, the center must have some source of funds available; for concreteness, suppose that the center can adjust its level of expenditure on good G after observing the locality's

choice of c_i'. [13]

In addition to specifying the source of central government funding for bailouts, it is necessary to describe the objectives that guide the central government's choice of bailout. There are several plausible ways to do this, but one that seems particularly natural is to suppose that the center intervenes in the fiscal affairs of locality i on behalf of those households in the economy that do not reside in the locality. The externalities generated by local public good provision provide the presumed rationale for the program of intergovernmental matching grants to begin with, and the failure of a locality to avail itself of these grants and thus to generate the external benefits that they would finance provide the presumed justification for any bailout. It therefore seems worthwhile exploring the hypothesis that the center acts to maximize the welfare of households residing outside of the locality i that is being considered for a bailout.

Given these assumptions, and noting that $N - n$ is the total number of households residing in localities other than i, the central government's "bailout" optimization problem can be written as

$$(B) \qquad max_{<g_i', G'>} u\left(x^*, z^*, (N-n)\beta(z^*) + n\beta\left(\frac{c_i'}{1-m^*} + g_i'\right), G'\right)$$

subject to

$$ng_i' + G' = NT - m^*\left((N-n)z^* + n\frac{c_i'}{1-m^*}\right). \tag{10}$$

The objective function in (B) is the utility of a representative household residing in any locality *other* than i. Note that all of the terms on the right-hand-side of the central government budget constraint (10) are already determined when deciding on the bailout g_i' to be given to locality i. Using (10) to solve for G' in terms of g_i' and substituting into the utility function in (B) reduces the problem to an unconstrained maximization with respect to g_i'. The first-order condition

$$\frac{u_Z(\cdot)}{u_G(\cdot)}\beta'\left(\frac{c_i'}{1-m^*} + g_i'\right) \leq 1 \tag{11}$$

must hold as a strict equality if $g_i' > 0$, i.e., if there is actually a bailout. Since $\beta'(\cdot) > 0$ and $u(\cdot)$ is strictly quasi-concave,

$$D \equiv n\left(u_{ZZ}(\cdot)\beta'(\cdot)^2 - 2u_{GZ}(\cdot)\beta'(\cdot) + u_{GG}(\cdot)\right) + u_Z(\cdot)\beta''(\cdot) < 0 \tag{12}$$

[13] If there is no source of funds available to the center once c_i' is selected, for instance if T, G, and m are all precommitted, or if it is impossible for the center to change its policies after c_i' is chosen, a bailout is infeasible, budget constraints are hard and the problem under analysis disappears. The assumption that funds for bailouts are obtained by reducing expenditures on centrally-provided public goods is not crucial for the analysis, and the results would go through without essential change if one assumes instead that central government tax increases are used to finance bailouts. In practice, central governments can finance bailouts by increasing the deficit, which can be interpreted as a reduction in future public good provision or an increase in future taxes (or some uncertain mix of the two).

which ensures satisfaction of the second-order condition for (B).

The values of g_i' and G' that solve the bailout problem (B) generally depend on the level of community i's own-contribution, c_i'. For instance, suppose locality i chooses its contribution at the first-best optimal level $c_i' = z^*$. Then it follows directly from (6) and (7) that (11) must hold as a strict inequality at a solution to (B), and, by continuity, this inequality must hold for any c_i' sufficiently close to z^*. In other words, *a locality that chooses a level of own-contributions to provision of good z that is optimal or sufficiently close to optimal receives no bailout from the central government.*

Can a locality ever obtain a positive bailout? If

$$\frac{u_Z(x^*, z^*, (N-n)\beta(z^*) + n\beta(0), G^*)}{u_G(x^*, z^*, (N-n)\beta(z^*) + n\beta(0), G^*)}\beta'(0) > 1, \tag{13}$$

the central government bails out locality i if it sets $c_i' = 0$ and thus makes no contribution at all to provision of good z. Indeed, by continuity, (13) implies that the center chooses a positive bailout $g_i' > 0$ for all own-contributions c_i' sufficiently close to zero. High values of $\beta'(0)$ and of $u_Z(\cdot)/u_G(\cdot)$ increase the left-hand side of (13), indicating, as is intuitive, that a bailout is relatively attractive to the center if the first units of good z provided to a locality produce large external benefits and if external benefits have a high value, at the margin, relative to provision of the national public good G, which we have assumed to be the source of funds from which bailouts are financed. Note that this relative valuation depends on $(N-n)\beta(z^*) + n\beta(0)$ and thus on $n\,(\beta(z^*) - \beta(0))$. If a large locality does not provide good z and if each household's consumption of good z produces a large external benefit, the total level of externalities in the economy will be significantly reduced and the valuation of Z relative to G will therefore be relatively high. On the other hand, if n is very small, $(N-n)\beta(z^*) + n\beta(0) \approx Z^*$ and the fact that locality i is not providing good z will leave the relative valuation of Z and G essentially unchanged from its value at the first-best optimum.

For all values of c_i' such that the bailout g_i' is positive, (11) must hold as an equality. Since $D < 0$, one can solve this equation implicitly for the bailout level g_i' as a function $g_i' = \gamma(c_i')$ of the own-contribution level c_i' such that

$$-\frac{1}{1-m^*} \leq \gamma'(c_i') < 0; \tag{14}$$

in particular, $\gamma'(c_i') = -1/(1-m^*)$ in the special case where the utility function is quasi-linear in G. [14] In this special case, each unit increase in own-contributions by locality i

[14] To derive (14), totally differentiate the first-order condition (11). Recalling (10), it follows that

$$\gamma' \equiv \frac{dg_i'}{dc_i'} = -\left(\frac{1}{1-m^*}\right) + \frac{n\,(u_{GG}(\cdot) - u_{ZG}(\cdot)\beta'(\cdot))}{D}.$$

The second term on the right-hand side lies in the interval $[0, 1)$; in particular, this term is zero in the special case where $u_G(\cdot)$ is constant.

leads to an increase in matching grants together with a reduction in the central government bailout g_i' that leaves the level of provision of the local public good in locality i unchanged, i.e., matching grants together with the bailout function $\gamma(\cdot)$ entail a 100% implicit tax at the margin on local government public expenditures from own-source revenues. More generally, if all goods are strictly normal, $-1/(1-m^*) < \gamma'(c_i') < 0$, and the bailout function entails an implicit tax on local own-contributions at a rate of less than 100%, i.e., increases in own-contributions are partially but not completely offset by reductions in net transfers from the central government.

In summary, the interval $[0, z^*]$ can be partitioned by some level $\bar{c}_i'$ of own-contributions by locality i such that there is no central government bailout whenever own-contributions fall in the interval $[\bar{c}_i', z^*]$, where $\bar{c}_i' < z^*$. The critical value $\bar{c}_i'$ will be strictly positive if the externalities associated with locality i's provision of good z are sufficiently strong that (13) holds, and in this case the bailout is a strictly positive and monotonically declining function of the own-contribution level for all values of c_i' in the interval $[0, \bar{c}_i']$. An interesting feature of this model is that *bailouts can only be triggered by discrete departures from first-best optimal levels of local public expenditure.*

Local Government Bailout Policy.

The bailout function $\gamma(\cdot)$ derived above is determined as a matter of central government policy, but there is a sense in which local governments also have "bailout policies," since they control the level of own-contributions to the local public good and thus determine whether to trigger a central government bailout. The condition (13) shows whether locality i can induce a bailout from the center, and it is therefore a *necessary* condition for locality i to face a soft budget constraint. However, it is not *sufficient* for a soft budget constraint, since the bailout may be sufficiently unattractive that a locality would always choose an own-consumption level greater than $\bar{c}_i'$. It is thus necessary to analyze the local choice of own-contributions to local public goods.

First, if the locality chooses $c_i' \geq \bar{c}_i'$, it receives no bailout from the center. In this case, it faces the budget constraint under matching grants (3M) with $m = m^*$ and its utility-maximizing choice under this constraint is to set $c_i' = z^*$. Therefore, *the locality never chooses contribution levels in the (open) internal* $(\bar{c}_i', z^*)$.

Now suppose instead that locality i chooses $c_i' < \bar{c}_i'$. It then receives a bailout $\gamma(c_i')$, and hence the consumption bundle that its residents receive is defined by the relations

$$x_i' = w - T - c_i' \tag{15.1}$$

$$z_i' = \frac{c_i'}{1-m^*} + \gamma(c_i') \tag{15.2}$$

$$Z' = (N-n)\beta(z^*) + n\beta(z_i') \tag{15.3}$$

$$G' = NT - m^*\left((N-n)z^* + n\frac{c_i'}{1-m^*}\right) - n\gamma(c_i'). \tag{15.4}$$

The utility level obtained by locality i in the event of a bailout is thus $u(x_i', z_i', Z', G')$, with

each argument depending on the choice of c_i'. Let $\hat{c}_i'$ denote the value of c_i' that maximizes utility in locality i, and let $(\hat{x}_i', \hat{z}_i', \hat{Z}, \hat{G})$ denote the values derived from (15) when $c_i' = \hat{c}_i'$. The first-order condition for a maximum of $u(x_i', z_i', Z', G')$ with respect to c_i' is

$$\left(\frac{u_z(\cdot)}{u_x(\cdot)} + n\frac{u_Z(\cdot)}{u_x(\cdot)}\beta'(\cdot)\right)\left(\frac{1}{1-m^*} + \gamma'\right) - n\frac{u_G(\cdot)}{u_x(\cdot)}\left(\frac{m^*}{1-m^*} + \gamma'\right) \leq 1, \tag{16}$$

with strict equality if $\hat{c}_i' > 0$. Note that this reduces to

$$\frac{n}{u_x(\cdot)} \leq 1 \tag{16'}$$

in the special case where the utility function is quasi-linear in G so that $\gamma' = -1/(1-m^*)$. In this case, locality 1 makes positive contributions to provision of good z only because its residents would benefit more from a smaller bailout which would free resources for the central government to spend on the national public good G. In this case we might plausibly expect locality i to set $c_i' = 0$, especially if localities are small. More generally, the first-order condition (16) requires localities to choose a level of own-contributions that takes into account central matching grants, bailouts, and their impact on the center's own public good provision level. It is not necessarily the case that the locality will set own-contributions at zero, but the possibility certainly exists. In any case, the own-contribution level $\hat{c}_i'$ lies in the interval $[0, \bar{c}_i']$ and is therefore lower than the first-best optimum.

The crucial question now is whether the utility attained by locality i in the event of a bailout is higher or lower than the utility it attains if it chooses $c_i' = z^*$ and does not induce a bailout. *Locality i induces a bailout – its budget constraint is soft* – if

$$u(\hat{x}_i', \hat{z}_i', \hat{Z}', \hat{G}') - u(x^*, z^*, Z^*, G^*) > 0. \tag{17}$$

To gain some insight into the conditions under which (17) holds, consider Figure 2. The budget constraint A^*B^* shows the menu of bundles of good x and z available to locality i when the central government offers a matching grant at rate m^*. The indifference curve U^*, which is the projection of $u(x_i, z_i, (N-n)\beta(z^*) + n\beta(z_i), G^*)$ onto the (x_i, z_i) plane, shows that a locality chooses the optimal level of z, z^*, if it adheres to the budget constraint A^*B^*. Assuming that condition (13) holds, however, A^*B^* is not really the constraint facing the locality. In particular, if it sets $c_i' = 0$, it will induce a bailout from the central government such that a positive level of $z_i = \gamma(0)$ is attained, allowing the locality to consume a bundle $(x_i, z_i) = (w - T, \gamma(0))$, denoted in the figure by point S which lies to the right of point A^*. Assuming for convenience of illustration that the utility function is quasi-linear in G, so that $\gamma' = -1/(1-m^*)$, any increases in locality i's contributions above $c_i' = 0$ result in reductions in the bailout that leave z_i unchanged until the bailout is reduced to $\gamma(\bar{c}_i^*) = 0$ at the contribution level $c_i' = \bar{c}_i'$. The vertical segment ST represents the consumption bundles attainable to locality i's residents when it chooses levels of own-contributions $c_i' \epsilon [0, \bar{c}_i']$. The budget constraint that locality i really faces, then, is not A^*B^* but STB^*. [15]

[15] If all goods are normal, $\gamma' > -1/(1-m^*)$ and the segment ST is downward-sloping rather than vertical.

As drawn, the point S lies above the indifference curve U^*. This does not necessarily mean that the locality wishes to induce a bailout, since a bailout lowers the level of provision of the national public good G and this means that the residents of locality i would require a consumption bundle (x_i, z_i) lying strictly above the indifference curve U^* to be as well off as at the no-bailout equilibrium (x^*, z^*). However, if the bailout is sufficiently generous so that $\bar{c}_i'$ and thus S lie sufficiently far to the right, the utility of local residents will indeed by maximized by inducing a bailout. On the other hand, if $\bar{c}_i'$ is sufficiently small that point S lies on or under the indifference curve U^*, local residents are definitely worse off in the event of a bailout and would not choose to induce one.

It is clear from Figure 2 that whether a locality chooses to induce a bailout depends partly on its own preferences, as represented by the shape of indifference curves like U^*. Whether the locality induces a bailout also depends on the external effects that its provision of good z generates for other localities. These effects determine the location of the critical point $\bar{c}_i'$ and thus the point T at which the level of own contributions is sufficiently high to reduce the bailout to zero, and they determine as well the rest of the bailout function $\gamma(c_i')$, represented in the diagram by the segment ST. Roughly speaking, if own-consumption of good z is highly substitutable with consumption of other goods, the indifference cure U^* will be relatively flat and a bailout is more likely. By contrast, if local residents have a very inelastic demand for good z, the lower level of provision attained under a bailout plan is likely to be unattractive and they will not induce a bailout. If external effects are strong, the schedule ST will lie relatively far to the right and bailouts are thus more likely to raise the welfare of local residents, whereas the opposite is true if external effects are small. Indeed, the distance between the socially-efficient level of local public good provision and the critical value of own-contributions $\bar{c}_i'$ below which the central government intervenes with a bailout can be viewed, in a rough way, as an indicator of the relative importance of local and external benefits from the provision of the local public good. With a low level of external benefits, $\bar{c}_i'$ is small, the points S and T in Figure 2 lie close to point A^*, and the segment ST becomes small; in the extreme case where there are no external effects, $\bar{c}_i' = 0$ and the segment ST disappears. In this case, there is no possibility of a bailout and the local budget constraint is definitely hard. By contrast, if external effects are large and local benefits from local public good provision are small, $\bar{c}_i'$ lies close to z^*. In the extreme case where there are no local own-benefits and where the local population is reduced to a single household, all of the benefits are external and $\bar{c}_i' = z^*$. In this case, a bailout is unavoidable.

IV. Bailouts and Jurisdictional Size

The analysis in Section III has proceeded on the assumption that all localities are of the same size, n. How does this variable affect the central government's bailout policy and the prospect that localities might choose to induce bailouts? Does this model suggest that fiscal decentralization, which may be interpreted within the model as the organization of the public sector into numerous small jurisdictions rather than a few large ones, is likely to soften local budget constraints? Or, on the contrary, does it suggest that bailouts are more

likely when localities are large in size? Does the model provide a formal foundation for the notion that localities can be "too big to fail"?

It is difficult to answer these questions without imposing additional structure on the model. When a locality decides whether or not to induce a bailout, it must compare two discretely different consumption bundles, (x^*, z^*, Z^*, G^*) and $(\hat{x}', \hat{z}', \hat{Z}', \hat{G}')$, involving every argument of the utility function. Assessing the effect of jurisdictional size on bailouts amounts to modelling the vector $(\hat{x}', \hat{z}', \hat{Z}', \hat{G}')$ as a function of n, determining whether the $u(\hat{x}', \hat{z}', \hat{Z}', \hat{G}')$ is increasing or decreasing in n, and comparing its value with $u(x^*, z^*, Z^*, G^*)$.

In order to shed some light on this formidable analytical problem, the present section explores the implications of successively more specialized assumptions about preferences. To begin with, suppose that the utility function is additively separable in each argument and quasi-linear in G, the national public good, i.e.,

$$(A) \qquad u(x, z, Z, G) \equiv A(x) + B(z) + C(Z) + G.$$

One important implication of quasi-linearity in G has already been noted, namely, that $\gamma'(c_i') = -1/(1-m^*)$ whenever $\gamma(c_i') > 0$. Hence, $c_i'/(1-m^*) + g_i' = c_i'/(1-m^*) + \gamma(c_i') = \bar{c}_i'$ for all values of own-contributions c_i' such that the center is willing to bail out the locality. Under these conditions, localities are likely to choose either $c_i' = z^*$ or $c_i' = 0$, since small positive own-contributions to the provision of good z are completely offset by reductions in bailouts from the center. More precisely, if household preferences satisfy (A), we know from (16′) that they will choose either $c_i' = 0$ or $c_i' = z^*$ if

$$n < A'(w - T^*). \tag{16″}$$

When (16″) holds, as will be assumed for the remainder of this section, the analysis simplifies considerably. [16] In particular, if a bailout occurs, (16″) insures that $c_i' = 0$, that the level of bailout is therefore the unique value $\gamma(0) = \bar{c}_i'$, and that the bailout is precisely the level of consumption of the local public good z that will be consumed by residents in the locality receiving the bailout, i.e., $\hat{z}_i' = \bar{c}_i' = \gamma(0)$.

Assuming that preferences satisfy (A), the unique bailout level $\bar{c}_i'$ that the center provides is determined by the first-order condition (11), which now takes the form

$$C'(Z)\beta'(\bar{c}_i') \leq 1 \tag{11′}$$

where

$$Z = (N-n)\beta(z^*) + n\beta(\bar{c}_i') = Z^* - n(\beta(z^*) - \beta(\bar{c}_i')). \tag{18}$$

[16] Specific examples where (16″) holds are provided below.

This condition is satisfied as an equality if the bailout $\bar{c}_i'$ is strictly positive, which will be true if and only if (13) holds, i.e.,

$$C'\left(Z^* - n\left(\beta(z^*) - \beta(0)\right)\right)\beta'(0) > 1. \tag{13'}$$

For a given utility function satisfying (A) and for a given external benefit function $\beta(\cdot)$, whether or not condition (13') holds depends only on the value of n. In particular, since

$$\frac{dZ}{dn} = -\left(\beta(z^*) - \beta(0)\right) < 0, \tag{19}$$

and since $C(\cdot)$ is strictly concave in Z, it follows that there is a critical value of jurisdiction size, say $\bar{n}\epsilon[0, N]$ such that (13') does not hold for all $n \leq \bar{n}$ while it does hold for all $n > \bar{n}$. That is, *there is a critical jurisdiction size $\bar{n}$ such that localities no larger than $\bar{n}$ receive no bailout from the center, whereas localities larger than n do receive positive bailouts*, if they choose a zero level of own-contributions to good z. Furthermore, using (11') to solve implicitly for $\bar{c}_i'$ as a function of n, we have

$$\frac{d\bar{c}_i'}{dn} = \frac{nC''(Z)\beta'(\bar{c}_i')}{nC''(Z)\beta'(\bar{c}_i')^2 + C'(Z)\beta''(\bar{c}_i')}\left(\beta(z^*) - \beta(\bar{c}_i')\right) > 0 \qquad \text{for} \quad n > \bar{n}, \tag{20}$$

i.e., *the size of the bailout offered by the center to a locality that makes zero own-contributions to local provision of good z is higher, the larger the locality.* In short, given additively separate and quasi-linear preferences as specified in (A), *larger localities can extract larger bailouts from the central government than smaller ones, and small localities may not be able to extract any bailout from the center at all.*

These considerations clearly suggest that larger localities are more likely to face soft budget constraints than smaller ones. However, the analysis so far is still incomplete since, as noted in Section III, whether a locality faces a hard or soft budget constraint depends not only on the size of the bailout it receives from the center, if any, but on the utility payoff of the consumption bundle it attains under a bailout as compared to that attainable with first-best optimal own-contributions. It is true that localities of size $n \leq \bar{n}$ always choose a first-best optimal level of local contributions to good z, $c_i' = z^*$, since they have no chance of obtaining a bailout. These jurisdictions therefore certainly face hard budget constraints. More generally, however, the residents of a jurisdiction that fails to finance an optimal level of provision of good z may obtain a bailout from the center, and if so, they (a) consume more than x^* units of private goods by escaping some or all of the burden of local taxes to finance good z, (b) consume a level of good z, financed in part or in total by the central government bailout, which is less than z^*, (c) enjoy a level of external benefits less than Z^*, since their fellow residents consume less than z^* units of good z, and (d) consume a level of the national public good G that is augmented by the fact that the locality receives a lower amount of matching grants from the center but that is reduced because the locality receives a bailout. Taking all of these factors into account, it is hardly obvious whether bailouts raise or lower utility for a locality's residents, or whether bailouts are more or less attractive to large localities compared with small ones.

In the face of these intricacies, it is useful to resort to illustrative numerical calculations. The calculations reported here are based on a further specialization of the assumptions embodied in (A). Specifically, assume that

$$
\begin{aligned}
A(x) &= a \ln x \qquad a > 0 \\
B(z) &= b_0 z - \frac{b_1}{2} z^2 \qquad b_i > 0, \quad i = 0, 1 \\
C(Z) &= c_0 Z - \frac{c_1}{2} Z^2 \qquad c_i > 0, \quad i = 0, 1 \\
\beta(z) &= \beta_0 z - \frac{\beta_1}{2} z^2 \qquad \beta_i > 0, \quad i = 0, 1.
\end{aligned}
$$

These functional specifications are selected both because they are computationally convenient and because they are economically well-behaved.

The functional forms and parameters are not chosen to simulate a real economy but rather to shed light on the working of the model. It is desirable, however, to choose values so that endogenous variables take on values in intermediate ranges where boundary properties of the model do not affect the results decisively. As a matter of normalization, let $N = 100$ and $w = 10$. Other parameters are chosen such that $x^* = 7, z^* = 1$, and $G^* = 2N = 200$; at the first-best optimum, then, 70% of income is allocated to private good consumption, 20% to the national public good G, and 10% to the local public good z. In addition, parameter values are selected to insure that (16″) is satisfied. This simplifies the analysis by insuring that own-contributions are either 0 or the first-best optimal amount, $z^* = 1$. The issues of primary interest concern the effect of jurisdictional size n (a) on the bailouts offered by the central government and (b) on the utility payoff to a locality if it accepts a bailout compared with the utility level at a first-best optimum. To investigate these effects, n is allowed to take on integer values from 1 to 50, representing jurisdictional sizes from 1% to 50% of the national population.

What do the numerical results show? To begin with a concise overview, it may be noted first that they of course confirm (20), that is, the bailout offered by the central government to a locality is larger, the larger the locality. Moreover, if localities are sufficiently small, they are not offered any bailout at all. Furthermore, the calculations reveal that the utility payoff in the event of a bailout is larger for larger localities, and that in some cases large localities find that utility is higher under a bailout than at the first best optimum. In these cases, *the local government budget constraint is soft* in the sense defined above, that is, (a) the central government does offer resources to the locality to finance provision of good z if the locality fails to do so (i.e., there is a non-zero bailout) and (b) the locality prefers this outcome to that attainable if it makes the first-best optimal level of own-contributions to provision of good z (i.e., the locality prefers the bailout to the first-best outcome).

Table 1 reports some additional details of the findings and their sensitivity to critical parameters. The three columns correspond to increasing values of the parameter c_0, a parameter that reflects the importance of external effects in the utility function, the three

rows correspond to increasing values of b_0, a parameter that reflects the magnitude of the own- benefit from local public good provision. The *first* number in each cell is $\bar{n}$, that is, the minimum size of jurisdiction that receives a positive bailout from the center if its own-level of contributions to provision of good z is zero; thus, in the second column, we see that localities that constitute at least 21% of the national population receive positive bailouts. If smaller localities choose not to provide the local public good, they receive no bailout at all. We immediately see that localities of size $n < 21$ thus face hard budget constraints. This critical value $\bar{n}$ depends on the importance of external effects; with a smaller value of c_0, the first column shows that $\bar{n} = 34$, whereas the third column shows that $\bar{n} = 7$ when c_0 takes on a larger value. Note that $\bar{n}$ does *not* depend on the parameter b_0, which reflects the own-benefit from local public good provision. The central government acts as the agent of *other* localities in deciding whether or not (and by how much) to bail out a given locality, and does not concern itself with the direct benefits of its bailout to the residents of the locality. Variations in b_0 do not affect the size of the external benefit from local public good provision and thus do not affect $\bar{n}$.

The own-benefits from local public good provision, as measured by b_0, certainly do affect the desirability to a locality of inducing a bailout from the center. Since the center ignores local benefits when choosing the level of bailout, the level of local public good provision $\bar{c}_i'$ is smaller than the first-best level z^* in a locality that receives a bailout. If local own- benefits are important, then the drop in local public good provision from z^* to $\bar{c}_i'$ is costly to local residents and a bailout is less appealing on this account, whereas if local own-benefits are relatively insignificant the reduction in the consumption of good z under a bailout does not matter very much to local residents. The *second* entry in each cell of Table 1 shows the minimum size of jurisdiction at which utility under a bailout exceeds the utility obtained at a first-best optimum. Thus, the second row of the second column shows that localities that constitute 41% or more of national population are better off under bailouts. These jurisdictions therefore face *soft* budget constraints. Under the same parameter values, localities of sizes ranging between 21% and 41% of the national population would indeed be bailed out by the center if they chose not to provide good z locally, but the terms of the bailout are sufficiently unattractive that they prefer not to induce bailouts. Localities of these sizes, then, face *hard* budget constraints. This is in addition to localities that contain less than 21% of the national population and which, as noted above, never obtain bailouts. Looking down rows in the middle column, we see that smaller localities ($n \geq 34$) face soft budget constraints when own-benefits from the local public good are relatively less important, as reflected in the low value of b_0 used in the top row, whereas when own-benefits are relatively more important, as reflected in the high value of b_0 used in the third row, even a locality containing half of the national population would not wish to induce a bailout ($n > 50$).

The first column shows that even very large localities ($n = 50$) do not wish to induce bailouts when external effects are weak; this is true for all three values of the own-benefit

parameter b_0. Weak external effects imply not only that the *number* of localities that can obtain non-zero bailouts from the center is small, but also that the *size* of the bailouts that they can obtain is relatively small, making bailouts unattractive. (To save space, bailout levels themselves are not reported.) In the first column, although localities containing 34% or more of the population can obtain positive bailouts, the bailout level is sufficiently small that all localities (up to and including those of size $n = 50$) are worse off under bailouts than at the first-best optimum. Thus, local government budget constraints are always hard for the parameter values shown in this column. In the third column, the external effect parameter c_0 is larger and now bailouts are offered, and are attractive, even to relatively small localities. When the own-benefits parameter b_0 is small (the top row), localities that contain 16% or more of the population face soft local budget constraints. If own-benefits are relatively highly valued (the third row), this critical size rises to 20%.

These numerical results are not of course particularly general. However, they do provide concrete illustrations of several important features of the model. For instance, they show that whether bailouts occur depend partly on the willingness of central governments to offer them, and that this willingness depends in part on the magnitude of the externalities associated with local public good provision. They also show, however, that the occurrence of bailouts is not only a matter of the central government's willingness to offer them but on the decision of localities to accept them. Furthermore, the calculations provide at least one class of examples where there is a clear inverse relationship between jurisdictional size and hardness of budget constraints.

V. Conclusions and Directions for Future Research

The preceding analysis provides a number of insights into the problem of soft budget constraints and bailouts for lower-level governments. The analysis begins with a recognition that interjurisdictional externalities can create the fundamental incentive for a central government to intervene in the fiscal affairs of lower-level governments. The classical solution to the efficiency problem created by interjurisdictional externalities is for a higher-level government to offer subsidies to localities that internalize these externalities. Normally, these subsidies entail higher levels of transfers to jurisdictions that carry out higher levels of externality-producing expenditures, but one implication of such subsidy programs is that subsidies achieve their minimal values, which are zero in the case of linear subsidy schemes, when local expenditures on externality-generating activities are zero. The problem is that a central government, acting in the interest of the rest of the society and facing a locality that has chosen a very low level of local public expenditures, may have an incentive to intervene after all to support a base level of provision of local public goods that produce externalities. A locality, anticipating this response, may choose to underprovide local public goods even when offered a system of corrective transfers that would otherwise appear to provide exactly the right incentive to make socially-efficient public expenditure decisions.

One implication of the analysis is that a locality that finds it advantageous to under-

provide local public goods in order to induce a bailout will not just reduce its public good provision by a small amount but rather will deviate discretely from the optimal level. This discrete deviation is attributable to the hypothesized loss of local control that accompanies a bailout, represented in the model by the assumption that the central government, when bailing out a locality, only takes into account the benefits and costs of local public good provision that accrue to the rest of society. This shift of control over local expenditures to an outside agency makes a bailout a discrete event, and the loss of local control can only be advantageous to a locality if it receives a discretely higher level of transfers from the center than it would at the socially-efficient level of local spending. Depending on the precise form of preferences, the best course of action for a locality that intends to induce a bailout may be to let local expenditures on externality-producing activities collapse entirely.

The conditions under which bailouts occur are intuitively clear: if the level of local expenditures selected by a central government under a bailout is "close" to the level that the locality would choose at a social optimum, then there is not much cost imposed on the locality from the loss of control over local spending levels relative to the gain from having the central government finance the totality of local public good provision. Specifying these conditions formally is difficult because they involve the interplay of preferences between local and non-local beneficiaries of local public goods. With mild restrictions on preferences, however, it can be shown that bailouts definitely do not occur when localities are sufficiently small, whereas they may – and, as calculations show, do – occur when localities are relatively large. Thus, the analysis suggests that rather than contributing to soft local budget constraints, fiscal decentralization actually can make budget constraints harder. If localities are found to pursue fiscal policies that induce bailouts, the analysis suggests that the problem is not that the public sector is too decentralized, but rather that it is too *centralized*. At least within the context of the model, institutional reforms that move the organization of the public sector in the direction of greater centralization may worsen rather ameliorate the problem of soft local government budget constraints.

Indeed, as the opening quotations suggest, there do seem to be real-world examples of cases where large localities are viewed as "too big" to fail and where they pursue fiscal policies that do in fact induce interventions by higher-level governments. The cases of New York City and of the states of Sao Paulo and Rio de Janeiro are also consistent with the basic message of the model. [17] The case of Washington, DC, provides an example where not only size, per

[17] Hillbrecht (1995) identifies state-owned banks in Brazil as an obstacle to hard state budget constraints and to monetary and fiscal discipline, developing a theoretical model in which individual states have the ability to add to the monetary base and thus to increase the economy-wide rate of inflation. (This specification is a condensed representation of a complex sequence of policies, discussed briefly in Section I, whereby the central bank ends up absorbing and monetizing the bad debts of state banks whose major assets are the non-performing debt of the state governments that own them.) In this model, the inflation rate is a kind of common-property resource: each state bears the costs of inflation engendered by the expansionary policies of other states. All states expand the money supply in order to capture seignorage, resulting in socially-excessive inflation. While this model provides a convincing explanation for how decentralized monetary policy (multiple money-creating banks) could

se, but the externalities that the city generates as the seat of the central government are likely to result in bailouts and controls. A recent study by Ades and Glaeser (1995) argues that the "urban giants" of the developing countries are likely to attract disproportionate levels of resources from central governments because of political factors; while that study does not specifically address the organizational form of local governance or the nature of intergovernmental fiscal relations, its findings are also broadly consistent with the results developed above. In research on the issues of monetary unification in the European Union, McKinnon (1995, 1996) has drawn attention to excessive borrowing by national governments as an obstacle to monetary union. In this context, he has compared the borrowing of US state and local governments with borrowing by provincial governments in Canada, noting that the amount of outstanding debt per capita in the latter is considerably higher than in the former. It is relevant to note here that Ontario and Quebec, which each account for about 25% of the population and income of Canada, have particularly high levels of indebtedness per capita. McKinnon argues that since most state and local debt in the US is domestically held (owing to special Federal income tax provisions), it is easier for the U.S. federal government to ignore bankruptcies in any one state or locality, i.e., no bailout, because such a bankruptcy does not impair the country's international credit rating and, possibly, the credit standing of the other states" (McKinnon [1996, p. 10]). He argues that the Canadian federal government would be more likley to intervene in case of a major provincial default because "the provinces borrow heavily overseas in foreign currencies as well as from a variety of domestic financial institutions" and because "contagion effects among foreign lenders ... could impair Canada's international credit standing." This is an interesting example of an externality argument associated with international capital markets. [18]

---

result in excessive inflation, it also suggests that "big" states – those containing a large fraction of population or of economic activity – would be less likely to engage in excessive monetary growth than small ones, since, by virtue of their size, they are not able to "free ride" as easily as small states. (The social costs of inflation generated by monetary growth in small states fall more heavily on outsiders than is the case for large states.) Indeed, it implies that the equilibrium rate of inflation rises as the number of states rises, i.e., as the extent of decentralization increases. While understandable within the context of the model, this finding fails to reflect the fact that it is the *large* Brazilian states, not the small ones, that face the softest budget constraints; as Hillbrecht (1995, p. 29) puts it, "São Paulo, Rio de Janeiro, Minas Gerais, and Rio Grande do Sul, ... the most important states, both politically and economically, ... can transfer their debts to the federal government more easily than small, less influential states." The foregoing analysis may help to explain why large (or "more important") subnational units face soft budget constraints while small ones do not.

[18] Of course, state constitutional prohibitions on borrowing and other institutional constraints may account for relatively low levels of state/local indebtedness in the US, as discussed, for example, by von Hagen (1991) and Poterba (1994). This observation is certainly consistent, however, with the notion that externalities and size contribute to soft budget constraints; folowing McKinnon's observations, one could argue that in the US case, the residents of US states have chosen to create institutions (such as constitutional rules) that limit borrowing because they anticipate adverse consequences associated with it, whereas this is not so in the Canadian case. Bruce (1995) presents a recent analysis of borrowing by lower-level governments, suggesting that household mobility – and thus, the prospect of escaping future tax burdens – may be partly responsible for excessive provincial borrowing in Canada, an argument reminiscent of analyses of municipal pension underfunding in the US (e.g., Epple and Schipper [1981], Inman [1981]).

Of course, one must proceed with caution in attempting to use a deliberately stylized theoretical model as a guide to empirical or policy analysis. Some of the simplifying assumptions exploited in the foregoing analysis have been introduced mainly for expositional convenience and could be relaxed without changing the major results, whereas others are more critical. The strong symmetry assumptions of the model, including the assumption that each locality is inhabited by identical individuals, greatly simplify the notation and formal analysis but the basic insights do not depend on them in a fundamental way. Of course, one implication of symmetry is that if it is in the interest of any one locality to act in a way that induces a bailout, it is in the interest of all. Interpreted literally, this would mean that either all localities underprovide the local public good and become subject to central government intervention or that none of them do. In the former case, the foregoing analysis does not correctly predict the true equilibrium of the system since it postulates that each locality assumes that all others adhere to the socially-optimal policies that they would follow if they could not be bailed out. This assumption is appropriate for characterizing an equilibrium when budget constraints are in fact hard, since then no jurisdiction finds it in its interest to deviate from its hard-budget equilibrium policies. If instead any one locality does choose to induce a bailout, then others would also choose to do so; the characterization of a Nash equilibrium would then entail a simultaneous determination of best replies in the presence of bailouts. The present analysis does not do this, and it cannot therefore be properly interpreted as an model of equilibrium under soft budget constraints. It can, however, be used to ascertain when the first-best allocation under hard budget constraints is in fact an equilibrium and when that allocation will break down; one might say that it provides model of the breakdown of equilibrium but not an equilibrium model of breakdown.

Actually, the problem of simultaneous deviation from hard budget constraints by all localities is more of an artifact of the strong symmetry assumptions of the model than a problem of intrinsic interest. Suppose that the economy contains many localities of different sizes and that there are localities in each size class with preferences, endowments, or technologies that make them more or less likely to act in such a way as to induce bailouts. Although a complete formal analysis of such a model would be more involved in its details than that presented above, this would not change the essential incentives that would or would not give rise to soft budget constraints for any *one* size/preference/endowment/technology type of locality. One can thus visualize a model in which only a small number of localities, corresponding to particular types (*e.g.*, the largest ones, or those with preferences and other attributes that make departures from local control of public good provision relatively painless for local residents) end up being bailed out while no bailouts occur in the rest of the system. In such an economy, the problem of modelling the simultaneous breakdown of hard budget constraints for all localities would disappear and the analysis would be rather close to that developed above, which could then be interpreted approximately as a description of what happens to the localities *within* any *one* class.

The analysis has deliberately abstracted from technological considerations that might

favor larger or smaller jurisdiction sizes. Economies of scale are of obvious importance for the efficient provision of local public goods and must be taken into account in determining the optimal degree of fiscal decentralization. (A closely related question, which may be posed either in normative or positive terms, concerns the assignment of functions to different levels of government.) One interpretation of the analysis is that there may be organizational diseconomies of scale, in the form of soft budget constraints and the inefficient incentives that they create, that must in practice be balanced against technological economies of scale in local public good provision. Theoretical and empirical study of this tradeoff presents an interesting question for further research. For instance, one might suppose that fiscal decentralization that appears to be excessive on purely technological grounds might still be attractive if it leads to harder budget constraints and the accompanying sharpening of incentives that they entail. By the same token, if technological economies make it impossible to decentralize specific public sector functions sufficiently to avoid soft budget constraints, then perhaps central government control would be preferable to an imperfect and poorly functioning partial decentralization; indeed, since bailouts in the present model involve central government financing and control of local public good provision, they may be taken, loosely, to represent (re)centralization of government functions.

As indicated in the introduction, many factors are simultaneously at work in real local fiscal crises, and any attempt to explain empirical reality or to formulate policy implications in specific cases would have to incorporate considerations that go beyond the simple analysis presented here. For example, a static model does not capture those aspects of intergovernmental fiscal that reflect the ongoing relationships between higher- and lower-level governments. The prospect of repeated failures by a lower-level government to adhere to its budget constraint, accompanied by repeated interventions by a higher-level government, is one that both parties might view rather differently than the one-shot failures described here. An analysis that explicitly incorporates dynamic interactions between governments need not, however, overturn the basic conclusions derived here; rather, it would complement the present analysis by shedding light on rather different aspects of intergovernmental fiscal relations. Similarly, one could develop a model in which localities are subject to random shocks of various kinds, such that only those localities with particular shocks (e.g., poor endowment realizations) end up choosing to induce bailouts. As discussed above in relation to the symmetry assumptions of the model, allowing for such diversity among localities is not likely to overturn the basic conclusions of the analysis though it may complicate them in detail. In brief, extensions of the analysis to incorporate uncertainty or explicit dynamics would complement it in valuable ways but would not negate its principal findings.

In practice, one way to avoid problems of bailouts is to make it less attractive for localities to induce them. This can be done by making "receivership" less attractive, for example by defining standards of fiscal responsibility and holding local officials personally liable for failure to meet them. (For example, a mayor who diverts public funds to contracts for influential friends could be jailed.) These standards are difficult to define, however. (One

may not want to impose criminal penalties on a mayor who provides generous pension benefits to workers who are members in influential unions, for instance, even though excessive compensation for politically-powerful municipal employees could be viewed as a diversion of funds from local uses that generate external benefits.) Alternatively, the center could offer matching grants on more favorable terms, raising the matching rate above the first-best efficient level of m^*. Although this may create a distortion by inducing over-provision of local public goods, it also makes bailouts less attractive to localities. Since equilibria with bailouts are also inefficient, excessively generous matching grants may be welfare-superior to matching grants set at first-best rates. This presents a subtle second-best problem that may warrant further study. A related issue concerns the normative and positive economics of the "assignment problem," that is, the problem of deciding which levels of government in a multi-level public sector should take responsibility for various public-sector functions. If there are economies of scale that necessitate provision of certain public goods by large jurisdictions, there may be no effective mechanism for a central government to harden the budget constraints of necessarily large lower-level governments, so that bailouts, and the welfare losses associated with them, are an unavoidable consequence of fiscal decentralization. Public goods for which this is the case may ones which are or should be provided by central governments, since it would be impossible to create the right incentives for efficient decentralized provision. Depending on economies of scale and the nature of benefit spillovers, other public sector functions could be efficiently provided by lower-level governments supported by appropriate corrective intergovernmental grants, and therefore would or should be provided in a decentralized fashion.

Finally, as noted in Section I, problems of fiscal discipline and soft budget constraints are not unique to subnational governments; the concept originated in the work of Kornai on state-owned enterprises in socialist economies. Quite outside the context of the planned economies, governments occasionally bail out particular industries or even individual firms (e.g., debt restructurings worked out with major banks, the Chrysler bail out). On the international level, foreign-aid donors, acting unilaterally or perhaps through multilateral agencies such as the IMF or World Bank, sometimes make extraordinary interventions in order to bail out recipient governments. [19] *Within* the local public sector, governments must allocate fiscal resources across competing functional units and there are sometimes crises in particular departments or agencies (the police department, the water company, etc.) that attract unanticipated budgetary flows, perhaps in the process upsetting the overall management of the budgetary mechanism (see, e.g., Carlsen [1996] for a typical case). In these and other cases, it is interesting to ask why some recipients are selected for bailouts while others are not. The foregoing analysis suggests that the magnitude of the externalities

[19] See Eaton *et al.* 1986 and Obstfeld and Rogoff (1996) for discussions of sovreign borrowing and country risk, and for references to a the large literature on this subject. A focus of many analyses of sovreign debt concerns the mechanism of contract enforcement; the present analysis, by emphasizing external effects, does not contradict but may complement some of the findings in that literature, in particular by helping to explain why some countries may be the objects of sustained assistance by donor/lender countries while others are not.

generated by local failures may be a principal determinant of the willingness of potential donors to intervene with special transfers.

REFERENCES

Ades, A.F. and E.L. Glaeser (1995), "Trade and Circuses: Explaining Urban Giants," *Quarterly Journal of Economics 110*, 195–227.

Bahl, R. and C. Wallich (1995), "Intergovernmental Fiscal Relations in the Russian Federation," in R.M. Bird, R.D. Ebel, and C.I. Wallich (eds.), *Decentralization of the Socialist State: Intergovernmental Finance in Transition Economies* (Washington, DC: The World Bank), 321–378.

Bird, R.M. (1994). "A Comparative Perspective on Federal Finance." In K.G. Banting, D.M. Brown, and T.J. Courchene (ed.), *The Future of Fiscal Federalism*. Kingston, Ontario: School of Policy Studies, Queen's University, 293–322.

Bird, R.M., R.D. Ebel, and C.I. Wallich (1995), "Fiscal Decentralization: From Command to Market," in R.M. Bird, R.D. Ebel, and C.I. Wallich (eds.), *Decentralization of the Socialist State: Intergovernmental Finance in Transition Economies* (Washington, DC: The World Bank), 1–68.

Bruce, N. (1995) "A Fiscal Federalism Analysis of Debt Policies by Sovreign Regional Governments," *Canadian Journal of Economics 28*, S195– S206.

Carlsen, F. (1995), "Why Is Central Regulation of Local Spending Decisions So Pervasive? Evidence from a Case Study," *Public Budgeting and Finance 15*, 43–57.

Crémer, J. (1995). "Arm's Length Relationships," *Quarterly Journal of Economics 110*, 275–296.

Dewatripont, M. and E. Maskin (1995), "Credit and Efficiency in Centralized and Decentralized Economies," *Review of Economic Studies 62*, 541–555.

Eaton, J. and M. Gersovitz (1989), "Country Risk and the Organization of International Capital Transfer," in G. Calvo *et al.* (eds.), *Debt, Stabilization, and Development* (Oxford: Basil Blackwell), 109–129.

Eaton, J., M. Gersovitz, and J.E. Stiglitz (1986), "The Pure Theory of Country Risk," *European Economic Review 30*, 481–513.

Epple, D. and K. Schipper (1981), "Municipal Pension Funding: A Theory and Some Evidence," *Public Choice 37*, 141–178.

von Hagen, J. (1991), "A Note on the Empirical Effectiveness of Formal Fiscal Restraints," *Journal of Public Economics 44*, 199-210.

Hillbrecht, R.O. (1995), "Monetary Policy and Seignorage in Brazil: Open Access and

the Role of State Banks," unpublished Ph.D. thesis, University of Illinois at Urbana-Champaign.

Inman, R.P. (1981), " 'Municipal Pension Funding: A Theory and Some Evidence': A Comment" *Public Choice 37*, 179–187.

Inman, R.P. (1995), "How To Have a Fiscal Crisis: Lessons From Philadelphia," *American Economic Review 85*, 378–383.

Inter-American Development Bank (1994), *Economic and Social Progress in Latin America: 1994 Report* (Washington, DC: Inter-American Development Bank).

Kornai, J. (1986), "The Soft Budget Constraint," *Kyklos 39*, 3–30.

McKinnon, R.I. (1995), "Intergovernmental Competition in Europe with and without a Common Currency," *Journal of Policy Modeling 17*, 463–478.

McKinnon, R.I. (1996), "Monetary Regimes, Government Borrowing Constraints, and Market-Preserving Federalism: Implications for EMU," unpublished.

McLure, C.M., Jr., C.I. Wallich, and J. Litvack (1995), "Special Issues in Russian Federal Finance: Ethnic Separatism and Natural Resources." In R.M. Bird, R.D. Ebel, and C.I. Wallich (eds.), *Decentralization of the Socialist State: Intergovernmental Finance in Transition Economies.* Washington, DC: The World Bank, 379–404.

Noto, N.A. and L. Rymarowicz (1995), "Financial Control Boards for Cities in Distress," Report for Congress (Library of Congress: Congressional Research Service).

Oates, W.E. (1972), *Fiscal Federalism* (New York: Harcourt Brace Jovanovich).

Obstfeld, M. and K. Rogoff (1996), *Foundations of International Macroeconomics* (Cambridge: MIT Press) .

Owens, J. and G. Panella (eds.) (1991), *Local Government Finance: An International Perspective* (Amsterdam: North-Holland).

Poterba, J. (1994), "State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Politics," *Journal of Political Economy 102*, 799–821.

Prud'homme, R. (1995), "The Dangers of Decentralization," *The World Bank Research Observer 10*, 201–220.

Qian, Y. (1994), "A Theory of Shortage in Socialist Economies Based on the 'Soft Budget Constraint', " *American Economic Review 84*, 145–156.

Roy, J. (ed.) (1995), *Macroeconomic Management and Fiscal Decentralization* (Washington, DC: The Economic Development Institute of the World Bank).

Shah, A. (1994), "The Reform of Intergovernmental Fiscal Relations in Developing and Emerging Market Economies," World Bank Policy and Research Series No. 23.

Tanzi, V. (1996), "Fiscal Federalism and Decentralization: A Review of Some Efficiency and Macroeconomics Aspects," in M. Bruno and B. Pleskovic (eds.), *Annual World Bank Conference on Development Economics* (Washington DC: The World Bank), 295–316.

Ter-Minasian, T. (1996), "Borrowing by Subnational Governments: Issues and Selected International Experiences," International Monetary Fund, Fiscal Affairs Department, unpublished.

von Thadden, E.-L. (1995), "Long-Term Contracts, Short-Term Investment, and Monitoring," *Review of Economic Studies 62*, 557-575.

Wildasin, D.E. (forthcoming), "Fiscal Aspects of Evolving Federations: Issues for Policy and Research," in D.E. Wildasin (ed.) *Fiscal Aspects of Evolving Federations* (Cambridge: Cambridge University Press).

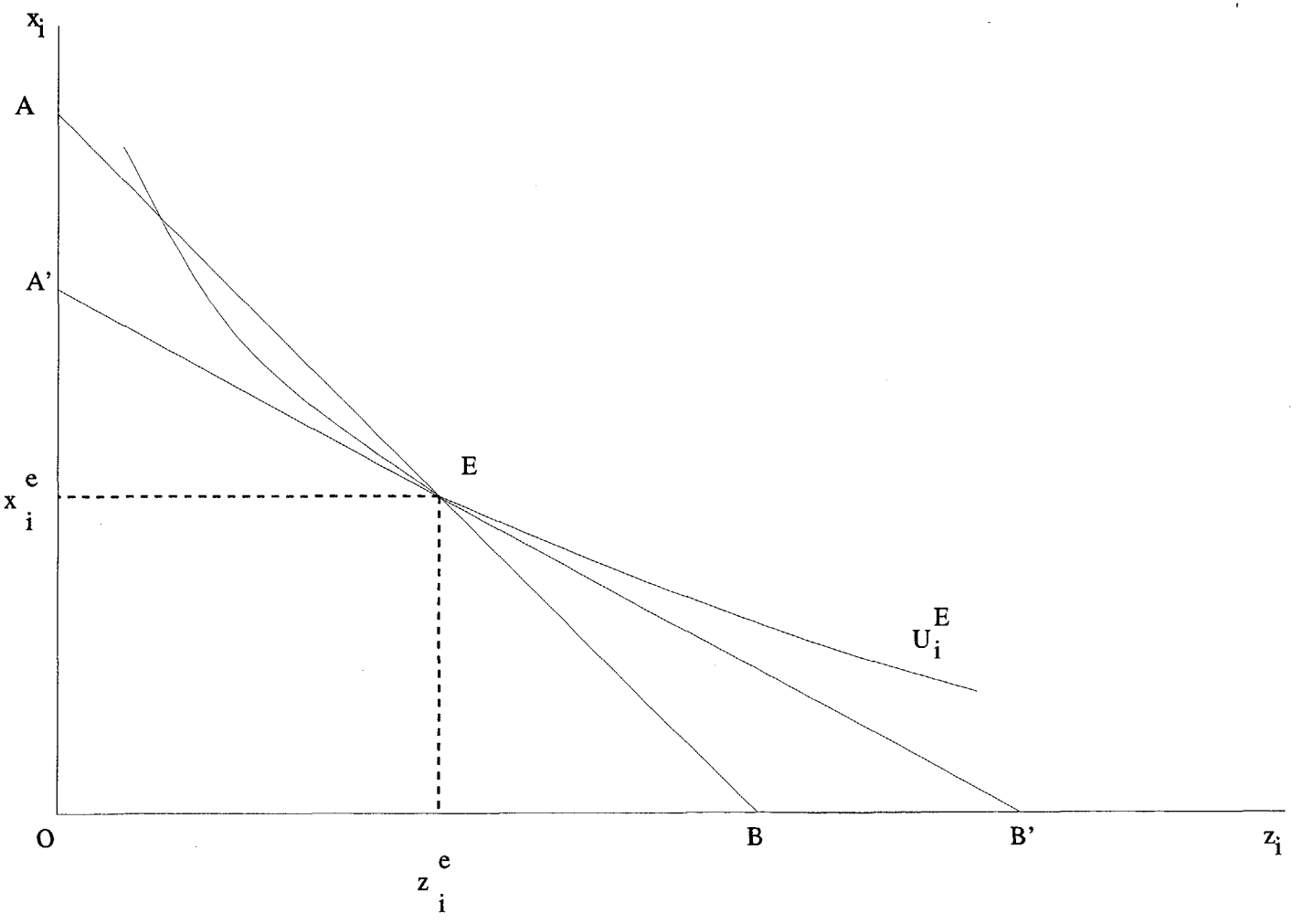
The World Bank (1995a), *China: Macroeconomic Stability in a Decentralized Economy* (Washington, DC: The World Bank).

The World Bank (1995b), *Economic Developments in India: Achievements and Challenges* (Washington, DC: The World Bank).

The World Bank (1995c), *Brazil State Debt: Crisis and Reform* (Washington, DC: The World Bank).

The World Bank (1996a), *Fiscal Management in Russia* (Washington, DC: The World Bank).

The World Bank (1996b), *Russian Federation: Toward Medium-Term Viability* (Washington, DC: The World Bank).


x_i
A
A'
x_i^e
E
U_i^E
O
B
B'
z_i
z_i^e

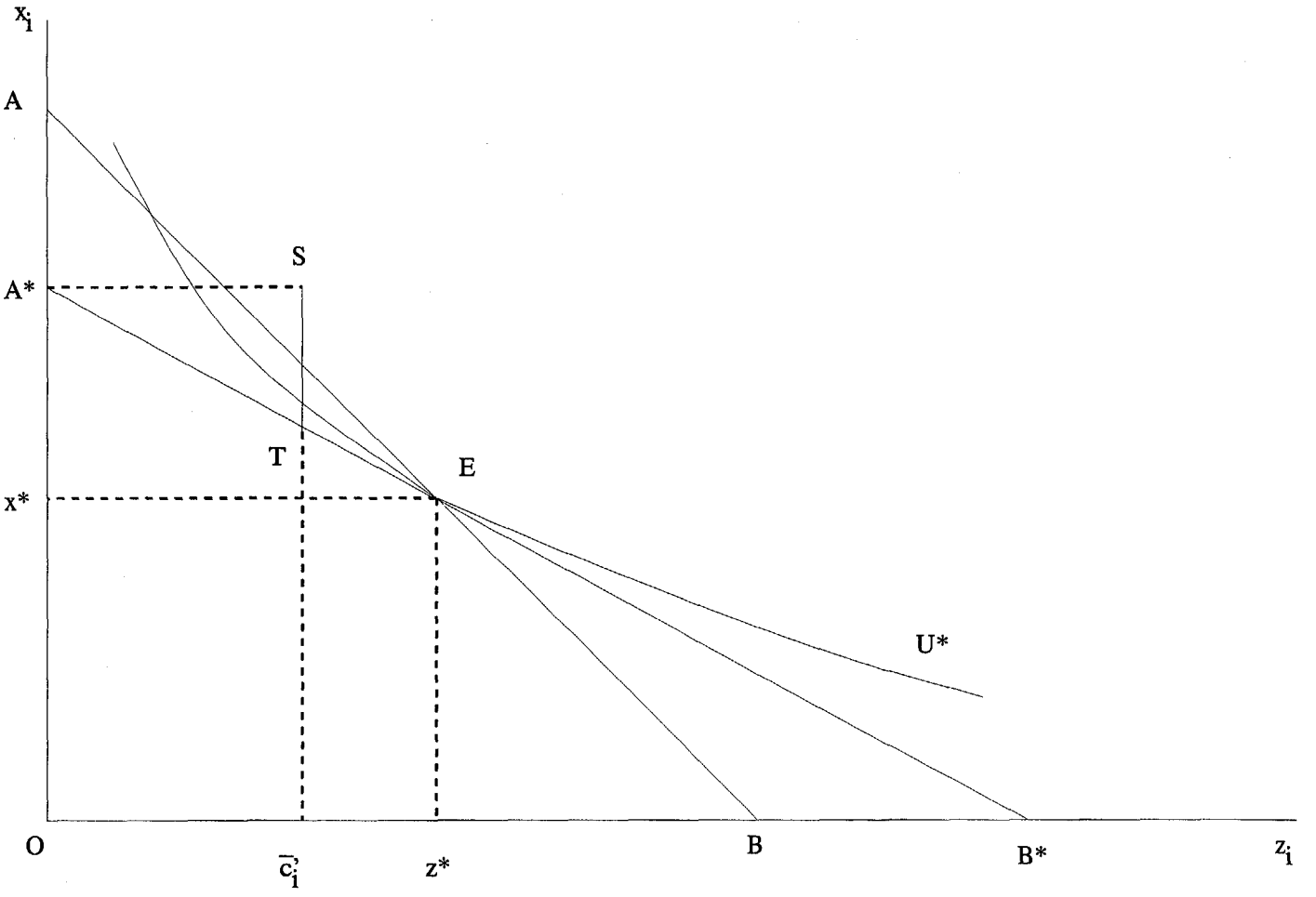


FIGURE 2

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS1828	The Determinants of Banking Crises: Evidence from Developed and Developing Countries	Asli Demirgüç-Kunt Enrica Detragiache	September 1997	P. Sintim-Aboagye 38526
WPS1829	Economic Reform and progress in Latin America and the Caribbean	Norman Loayza Luisa Palacios	September 1997	E. Khine 37471
WPS1830	Private Ownership and Corporate Performance: Some Lessons from Transition Economies	Roman Frydman Cheryl W. Gray Marek Hessel Andrzej Rapaczynski	September 1997	B. Moore 38526
WPS1831	How Trade Patterns and Technology Flows Affect Productivity Growth	Wolfgang Keller	September 1997	J. Ngaine 37947
WPS1832	Pension Reform in Bolivia: Innovative Solutions to Common Problems	Hermann von Gersdorff	September 1997	C. Pavlak 82099
WPS1833	Cost Recovery and Pricing of Payment Services	David B. Humphrey Robert H. Keppler Fernando Montes-Negret	October 1997	T. Ishibe 38968
WPS1834	The Comparative Advantage of Government: A Review	Pedro Belli	October 1997	L. Schunk 31779
WPS1835	Cost-Benefit Analysis of the Global Dracunculiasis Eradication Campaign	Aehyung Kim Ajay Tandon Ernesto Ruiz-Tiben	October 1997	A. Kim 35029
WPS1836	Health and Labor Productivity: The Economic Impact of Onchocercal Skin Disease	Aehyung Kim Ajay Tandon Asrat Hairu and Others	October 1997	A. Kim 35029
WPS1837	How Estonia's Economic Transition Affected Employment and Wages (1989–95)	Rivo Noorkoiv Peter F. Orazem Allan Puur Milan Vodopivec	October 1997	S. Fallon 38009
WPS1838	The Limits and Merits of Participation	Paulo Vieira da Cunha Maria Valeria Junho Pena	October 1997	E. Thomas 31019
WPS1839	The Effects of Financial Liberalization and New Bank Entry on Market Structure and Competition in Turkey	Cevdet Denizer	November 1997	E. Khine 37471
WPS1840	Efficient Public Sector Downsizing	Martin Rama	November 1997	S. Fallon 38009
WPS1841	Patterns of Metropolitan Development: What Have We Learned?	Gregory K. Ingram	November 1997	J. Ponchamni 31052

Policy Research Working Paper Series

	Title	Author	Date	*Contact* for paper
WPS1842	Motorization and the Provision of Roads in Countries and Cities	Gregory K. Ingram Zhi Liu	November 1997	J. Ponchamni 31052