

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

August, 1997

Abstract

Central government matching grants can, in principle, induce socially-efficient provision of local public goods that produce spillover benefits. Local underprovision of public goods may however elicit direct central-government provision and finance (a “bailout”) that makes local residents better off than under grant-subsidized local provision; local underprovision that induces bailouts reveals the local budget constraint to be “soft.” Simulations suggest that the ability of a locality to extract a welfare-improving bailout depends positively on its size: budget constraints are more likely to be “hard” for small localities.

* The research reported here was largely undertaken visiting the Public Economics Division of the Policy Research Department of the World Bank. Earlier versions were presented at the Universities of British Columbia and Illinois, the World Bank, and at the Stanford SITE workshop. I am grateful to seminar participants and to many colleagues at the World Bank, especially J. Ahmad and G. Eskeland, for stimulating discussions on the topic of this paper, but retain responsibility for errors.

[†] Telephone: +1 615 343 2468. Fax: +1 615 343 8495.
Internet: wildasin@ctrvax.vanderbilt.edu.

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. In many developing countries or in the transition economies, on the other hand, subnational government fiscal crises seem to be much more commonplace.¹ 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

* 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.

¹ See Appendix A for a brief overview of recent international experience and references to additional literature.

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. Indeed, fiscal decentralization may itself provide crucial institutional support for market-oriented reforms (see, e.g., Weingast (1993, 1995), Qian and Weingast (1996), McKinnon and Nechyba (1997)).

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 seems 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 for bailouts in other cases?

The literature of fiscal federalism has not addressed these questions previously, which

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 refusal to 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?

indeed cannot even arise within the standard framework. Following classic treatments by authors such as Buchanan (1950), Oates (1972), Boadway and Flatters (1982), and many others, central governments are conventionally viewed as agents that establish a structure of intergovernmental fiscal relations, for instance through an assignment of expenditure functions and revenue sources to each level of government, accompanied by a system of fiscal transfers. *Taking this structure as given*, lower-level governments and market participants are then supposed to respond to the fiscal incentives that it embodies. Within this framework, the literature has examined how local governments adjust expenditure levels in response to matching grants from a central authority, how transfers affect the allocation of population among regions, and many other specific issues. There is no room in this framework for the notion that lower-level governments can somehow violate the constraints that a higher-level government attempts to impose on them and that the higher-level government would then have to intervene in order to re-establish fiscal control. Since much of the analytical work on fiscal federalism has been motivated by policy problems arising within the North American context, and since the institutions of fiscal federalism in the US and Canada well-established and relatively stable, it is unsurprising that analysts have ignored the problems of “soft budget constraints” and “bailouts.”

The present essay develops 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. In general, 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.³

A first basic premise of this 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 is assumed to establish programs of intergovernmental transfers that function as Pigovian corrective subsidies. A second basic premise of the model, however, is that public-sector decisionmaking proceeds sequentially. The central government’s program of corrective intergovernmental transfers 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

³ For ease of reference, lower-level governments will often be referred to below as “localities,” but this term should be understood to embrace jurisdictions such as states or provinces as well.

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 be “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 shows that the hardness of local budget constraints is partly in the hands of the localities themselves. In particular, even under conditions in which localities could elicit bailouts from the center, they may not find it advantageous to do so. In these cases, 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.

Much of the formal analysis assumes 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. Proofs of major results are presented in Appendix B.

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.⁴ 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

⁴ 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.

(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.⁵ 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.⁶ 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 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. It is assumed, therefore, 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 .

All governments, both local and central, use distortionless lump-sum taxes 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. Two types of grant instruments are considered: lump-sum and matching conditional grants. Let m denote the 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 lump-sum grants may be uniform across localities, the notation distinguishes the lump-sum grant level by locality because, as will become clear, the center may use lump-sum grants to intervene specifically in the financing of local public good provision in a particular locality.)

⁵ 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 $N \bmod n$.

⁶ The crucial distinction between goods x and z is that the former represents a use of local resources that does not produce externalities while the latter does. For convenience, x is referred to as a “private good,” but it should be interpreted as a composite of privately-provided goods plus any uses of local public funds for activities that do not produce spillover benefits.

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. The level of local public good consumed by each resident of locality i is thus

$$z_i = c_i / (1 - m) + g_i. \quad (1)$$

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.,

$$x_i = w - c_i - T. \quad (2)$$

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

$$x_i + (1 - m)z_i = w - T + g_i. \quad (3)$$

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

$$\sum_i n g_i + m \sum_i n z_i + G = NT \quad (4)$$

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. To define a non-cooperative equilibrium for this system, restricting attention to symmetric equilibria, 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 ⁷

$$(P) \quad \max_{x_i, z_i} \quad u(x_i, z_i, (N-n)\beta(\bar{z}) + n\beta(z_i), G) \quad \text{subject to (3)}.$$

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* if $z_i = \bar{z}$ at a solution to (P) and if the central government budget constraint is satisfied when $z_i = \bar{z}$ for all localities i .

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

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

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

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

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

where derivatives of $u(\cdot)$ are denoted by subscripts.

Optimal Grant Design. It is well known that Nash equilibria, as defined above, are generally not efficient. It is however possible to use matching grants to achieve an efficient

⁷ 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.

Nash equilibrium in which there is still some meaningful degree of fiscal decentralization. The first-order conditions for (P) imply that each locality chooses a level of local public good provision z' such that

$$\frac{u_z(x', z', Z', G^*)}{u_x(x', z', Z', G^*)} + n \frac{u_Z(x', z', Z', G^*)}{u_x(x', z', Z', G^*)} \beta'(z') = 1 - m \quad (7)$$

where x' is endowment net of local and central taxes. If the center selects the matching rate

$$m^* = (N - n) \frac{u_Z(x^*, z^*, Z^*, G^*)}{u_x(x^*, z^*, Z^*, G^*)} \beta'(z^*) \quad (8)$$

and a level of taxes $T_m^* = m^* z^* + G^*$ per capita, the equilibrium condition (7) coincides with the optimality condition (6.1) and the Nash equilibrium is efficient.

III. Credibility and 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.

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 (8) 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 .

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.⁸

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

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

subject to

$$ng'_i + G' = NT - m^* \left((N - n)z^* + n\frac{c'_i}{1 - m^*} \right). \quad (9)$$

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 (9) are already determined when deciding on the bailout g'_i to be given to locality i . Using (9) 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

$$\frac{u_Z(\cdot)}{u_G(\cdot)} \beta' \left(\frac{c'_i}{1 - m^*} + g'_i \right) \leq 1 \quad (10)$$

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,

$$D \equiv n \left(u_{ZZ}(\cdot) \beta'(\cdot)^2 - 2u_{GZ}(\cdot) \beta'(\cdot) + u_{GG}(\cdot) \right) + u_Z(\cdot) \beta''(\cdot) < 0 \quad (11)$$

⁸ No doubt other objectives for central government policy might be postulated. As further justification for the current hypothesis, note that any individual locality that deviates from the first-best policy contains less than half of the total population, and that maximization of the welfare of the residents of other jurisdictions is thus consistent with a majority-rule political process. (I am grateful to G. Besharov for this observation.)

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 , and the first-order condition (10) can be used to solve implicitly for this functional relationship, denoted by $g'_i = \gamma(c'_i)$. Note, first, the following result:

Proposition 1: There exists a critical value $\bar{c} < z^*$ of local own-contributions to local public good provision such that

$$\gamma(c'_i) = 0 \quad \text{for all } c'_i > \bar{c}.$$

According to Proposition 1, 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? The answer is yes, if externalities are “sufficiently important.” Specifically, if

$$\frac{u_Z(x^*, z^*, (N-n)\beta(z^*) + n\beta(0), G^*)}{u_G(x^*, z^*, (N-n)\beta(z^*) + n\beta(0), G^*)} \beta'(0) > 1, \quad (12)$$

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, one can show:

Proposition 2: Suppose that (12) holds. Then there exists a critical value $0 < \bar{c} < z^*$ such that

- (a) $\gamma(c'_i) > 0$ for any own contribution level $c'_i < \bar{c}$;
- (b) $\gamma(c'_i) = 0$ for any own contribution level $c'_i \geq \bar{c}$; and
- (c) for any value of $c'_i < \bar{c}$,

$$-\frac{1}{1-m^*} \leq \gamma'(c'_i) < 0;$$

in particular, if preferences are quasi-linear in G ,

$$\gamma'(c'_i) = -1/(1-m^*).$$

Intuitively, high values of $\beta'(0)$ and of $u_Z(\cdot)/u_G(\cdot)$ increase the left-hand side of (12), indicating 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 is 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.

The bailout function described in (c) creates an implicit tax on local own-contributions to the local public good. In the special case where preferences are quasi-linear in G , each unit increase in own-contributions by locality i 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 (12) 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 (12) 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) interval $(\bar{c}'_i, z^*)$* . If instead locality i chooses $c'_i < \bar{c}'_i$, it receives a bailout $\gamma(c'_i)$, and hence its residents consume the bundle

$$x'_i = w - T - c'_i \tag{13.1}$$

$$z'_i = \frac{c'_i}{1 - m^*} + \gamma(c'_i) \quad (13.2)$$

$$Z' = (N - n)\beta(z^*) + n\beta(z'_i) \quad (13.3)$$

$$G' = NT - m^* \left((N - n)z^* + n \frac{c'_i}{1 - m^*} \right) - n\gamma(c'_i). \quad (13.4)$$

yielding a utility level $u(x'_i, z'_i, Z', G')$. 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 (13) 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

$$\left(\frac{u_z(\cdot)}{u_x(\cdot)} + n \frac{u_Z(\cdot)}{u_x(\cdot)} \beta'(\cdot) \right) \left(\frac{1}{1 - m^*} + \gamma' \right) - n \frac{u_G(\cdot)}{u_x(\cdot)} \left(\frac{m^*}{1 - m^*} + \gamma' \right) \leq 1, \quad (14)$$

with strict equality if $\hat{c}'_i > 0$. *Locality i induces a bailout – its budget constraint is soft – if*

$$u(\hat{x}'_i, \hat{z}'_i, \hat{Z}', \hat{G}') - u(x^*, z^*, Z^*, G^*) > 0. \quad (15)$$

To gain some insight into the conditions under which (15) holds, consider Figure 1. The budget constraint AB shows the combinations of x and z available to a locality using only its own resources, while A^*B^* shows the budget constraint 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 (12) 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 $c'_i \in [0, \bar{c}'_i]$. The true budget constraint facing the locality i is thus STB^* .

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 be 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.

The figure illustrates that whether a locality chooses to induce a bailout depends on its own preferences, as represented by the shape of indifference curves like U^* . It depends

as well on the external effects that its provision of good z generates for other localities, since the externalities determine the bailout policy of the central government, represented in the figure by the segment ST . Roughly speaking, if own-consumption of good z is highly substitutable with consumption of other goods, the indifference curve 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 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, 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"?

While it is difficult to answer these questions in general, the present section shows that jurisdictional size does systematically affect the likelihood and size of bailouts when preferences satisfy additional restrictions. Part A provides additional results on the bailout policy of the central government when utility functions are additively separable and quasi-linear. Part B explores the bailout policies of local governments in still more specific cases, providing a series of findings based on numerical calculations. Part C discusses the implications of the analysis for the case where jurisdictions are of unequal size.

A. Jurisdictional Size and Central Government Bailout Policy.

Suppose that the utility function is additively separable in each argument and quasi-

linear in G , the national public good, i.e.,

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

an assumption which is maintained throughout this section. An 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,

Proposition 3: Assume that preferences satisfy (A). If

$$n < A'(w - T^*), \quad (14')$$

any locality that receives a bailout chooses a zero level of own-contributions to local public good provision; any locality that does not receive a bailout chooses the first-best efficient level of own-contributions to local public good provision.

It is now possible to show how jurisdiction size affects the bailout policy of the central government:

Proposition 4: Assume that preferences satisfy (A). Then there exists a critical jurisdiction size $\bar{n} \in (0, N]$ such that

- (a) if localities are of size $n \leq \bar{n}$, the central government does not provide bailouts, even for localities that choose zero levels of own-contributions to local public good provision;
- (b) if localities are of size $n > \bar{n}$, the central government does provide bailouts for localities that choose zero levels of own-contributions to local public good provision;
- (c) if localities are of size $n > \bar{n}$, the size of the bailout provided by the central government to localities that make zero own-contributions to the local public good is an increasing function of jurisdiction size, i.e.,

$$\frac{d\bar{c}'_i}{dn} > 0.$$

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 $\bar{n}$ do receive positive bailouts, if they choose a zero level of own-contributions to good z . Furthermore, 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.*

B. Jurisdictional Size and Local Government Bailout Policy.

Proposition 4 clearly suggests 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), using a specification which is both computationally convenient and economically well-behaved.⁹ 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.

The numerical results confirm Proposition 4: if localities are sufficiently small, they are not offered any bailout at all, while bailouts offered by the central government are increasing in jurisdictional size. More importantly, 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*

⁹ Specifically, it is assumed that $A(x) = a \ln x$, $B(z) = b_0 z - b_1 z^2/2$, $C(Z) = c_0 Z - c_1 Z^2/2$, and $\beta(z) = \beta_0 z - \beta_1 z^2/2$, with every coefficient strictly positive. As a normalization, it is assumed that $N = 100$ and $w = 10$. Other parameters are chosen such that $x^* = 7$, $z^* = 1$, and $G^* = 2N = 200$; thus, at the first-best optimum, 70% of income is allocated to private good consumption, 20% to the national public good G , and 10% to the local public good z . The parameter values also insure that (14') is satisfied.

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 additional details of the findings and their sensitivity to critical parameters. The three columns correspond to increasing values of a parameter of the utility function (c_0) that reflects the importance of external effects, while the three rows correspond to increasing values of a parameter (b_0) 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. Thus, localities of size $n < 21$ face hard budget constraints. This critical value $\bar{n}$ depends on the importance of external effects. The first column shows that $\bar{n} = 34$ when the external-benefit parameter is low, whereas the third column shows that $\bar{n} = 7$ when it is high. Note that $\bar{n}$ does *not* depend on the own-benefit parameter (which varies across the rows) because 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.

The own-benefits from local public good provision 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 (top

row) whereas when own-benefits are relatively more important, as reflected 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. 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 is larger and now bailouts are offered, and are attractive, even to relatively small localities. When the own-benefits parameter is small (top row), localities that contain 16% or more of the population face soft local budget constraints. If own-benefits are relatively highly valued (third row), this critical size rises to 20%.

These numerical results, though obviously special, illustrate the following important features of the model: (i) Whether bailouts occur depends on the willingness of central governments to offer them, which in turn depends on the magnitude of the externalities associated with local public good provision. (ii) The occurrence of bailouts depends not only on the central government's willingness to offer them but on the decision of localities to accept them, which in turn depends on the size of the bailouts that localities can induce and on the cost to the locality of the loss of local control over local public good provision; (iii) At least in one natural class of examples, there is a clear inverse relationship between jurisdictional size and hardness of budget constraints.

C. Hard and Soft Budget Constraints and Asymmetric Jurisdictional Sizes.

The formal analysis so far has been restricted to the case where all jurisdictions are of equal size. One obvious implication of this symmetry assumption 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 analysis presented above does

not do this, and it cannot therefore be interpreted as a 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 “equilibrium breakdown” but not a model of “breakdown equilibrium.”

While an analysis of equilibrium with soft budget constraints goes beyond the scope of the present paper, it is at least possible to show how the assumption of symmetric jurisdictions is not in itself critical for the essential results. To begin with, if there are jurisdictions of different sizes, the first-best allocation cannot be supported with identical matching grants for all localities. As suggested by (8) and (9), matching rates in this model should be higher for smaller jurisdictions. It is straightforward to show, however, that with appropriately-differentiated rates, the first-best optimum would be achieved if local budget constraints are hard.

Now consider a system with identical jurisdictions satisfying the conditions of Proposition 4, and suppose that every jurisdiction is of a size $n < \bar{n}$, so that the first-best allocation with matching grants is a hard budget constraint equilibrium. Consider now a repartitioning of the population into jurisdictions of different sizes, but always keeping jurisdictional size below the critical value $\bar{n}$. One can then show that with appropriately size-adjusted matching rates, it is still the case that the central government would not offer any bailouts to any locality, and that the first-best allocation can thus still be supported with hard budget constraints:

Proposition 5: Assume that preferences satisfy (A). Then there exists a critical jurisdiction size $\bar{n} \in [0, N]$ such that if localities are of size $n \leq \bar{n}$, the central government does not provide bailouts for localities that choose zero levels of own-contributions to local public good provision, even if localities are of unequal size.

Next, suppose that there is one locality, locality 0, of size $n_0 > \bar{n}$, and that all other localities are of sizes less than $\bar{n}$. Assume that matching grants are set at their optimal size-adjusted rates, so that the first-best allocation can be supported if budget constraints are hard. Clearly, if locality 0 does not induce a bailout, none of the small jurisdictions would do so, either. However, if every small locality makes its first-best own-contribution to local public good provision, the large jurisdiction 0 may choose to induce a bailout. For example, given the parameters corresponding to the middle cell of Table 1, if $n_0 > 41$ and if every other locality contains less than 21% of the population, the large jurisdiction would find it advantageous to deviate from its first-best policy and induce a bailout; thus, the hard budget constraint outcome is not a Nash non-cooperative equilibrium in this asymmetric case.

These observations confirm that results such as those presented in Proposition 4 or in the numerical analysis of part B do not depend in any essential way on the convenient sim-

plifying assumption that all localities are of equal size. Indeed, precisely because the model assumes that all households are identical and that there are no technological economies or diseconomies of scale in jurisdiction size, it highlights the fact that jurisdictional size alone can affect behavioral incentives. Although further detailed investigation is clearly necessary, the analysis shows how certain types of institutional reforms – consolidation of several small jurisdictions into a large one, for example – might change the incentive properties facing lower-level governments in a system of intergovernmental fiscal relations in an adverse way. It also suggests that interventions by higher-level government authorities in the fiscal affairs of large lower-level governments need not destroy the credibility of the higher-level authorities vis-a-vis small localities. Indeed, it is easy to see from Proposition 5 that if the central authorities succeed in making large localities follow socially-optimal policies – for instance, through detailed oversight and regulation – a transparent system of intergovernmental transfers can be sufficient to induce socially-efficient behavior by the remaining small jurisdictions in the system. Thus, the fact that central government interventions may be required in some instances does not, in itself, imply that a fiscally-decentralized system is inherently flawed.

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 underprovide 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 Brazilian states of Sao Paulo and Rio de Janeiro (discussed further in Appendix A) are also consistent with the basic message of the model. The example of Washington, DC, provides an example where not only size, per 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 likely 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.¹⁰

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 implica-

¹⁰ 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; following 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., Eppler and Schipper [1981], Inman [1981]).

tions 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. 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 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 accountable 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.¹¹ 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. In such cases, central government interventions through bailouts, regulatory oversight, or by other means,

¹¹ Note, as these examples suggest, that the analysis above does not imply that bailouts occur just because *total* local government spending is low. Recall from n. 6 above that the variable x_i , which is called “private good consumption,” should be interpreted as an aggregate of local spending on all goods and services that do not generate spillover benefits to the rest of society. Insufficient provision of “essential” services, not low government spending per se, is what triggers bailouts.

may be an unavoidable consequence of fiscal decentralization. Indeed, public goods for which this is the case may be ones which are or should be provided by central governments, since it may 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.¹² *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 generated by local failures may be a principal determinant of the willingness of potential donors to intervene with special transfers.

¹² See Eaton *et al.* 1986 and Obstfeld and Rogoff (1996) for discussions of sovereign borrowing and country risk, and for references to a large literature on this subject. A focus of many analyses of sovereign debt concerns the mechanism of contract enforcement; the present analysis, by emphasizing external effects, certainly 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.

REFERENCES

- Ades, A.F. and E.L. Glaeser (1995), "Trade and Circuses: Explaining Urban Giants," *Quarterly Journal of Economics* 110, 195–227.
- 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.
- Boadway, Robin W., and Flatters, Frank R. (1982) 'Efficiency and Equalization Payments in a Federal System of Government: A Synthesis and Extension of Recent Results,' *Canadian Journal of Economics* 15, 613–633.
- Bruce, N. (1995) "A Fiscal Federalism Analysis of Debt Policies by Sovereign Regional Governments," *Canadian Journal of Economics* 28, S195– S206.
- Buchanan, J.M. (1950), "Federalism and fiscal equity," *American Economic Review* 40, 583–599.
- 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.
- 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.
- Epplé, 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.
- Inman, R.P. (1981), " 'Municipal Pension Funding: A Theory and Some Evidence': A Comment" *Public Choice* 37, 179–187.
- 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.

- McKinnon, R. and T. Nechyba (1997), "Competition in Federal Systems: The Role of Political and Financial Constraints," in J. Ferejohn and B.R. Weingast (eds.), *The New Federalism: Can the States Be Trusted?* (Stanford: Hoover Institution Press), 3–61.
- Oates, W.E. (1972), *Fiscal Federalism* (New York: Harcourt Brace Jovanovich).
- Obstfeld, M. and K. Rogoff (1996), *Foundations of International Macroeconomics* (Cambridge: MIT Press) .
- 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. and B.R. Weingast (1996), "China's Transition to Markets: Market-Preserving Federalism, Chinese Style," *Policy Reform* 1, 149–185.
- 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.
- Weingast, B.R. (1993), "Constitutions as Governance Structures: The Political Foundations of Secure Markets," *Journal of Institutional and Theoretical Economics* 149, 286–311.
- Weingast, B.R. (1995), "The Economic Role of Political Institutions: Market-Preserving Federalism and Economic Development," *Journal of Law, Economics, and Organization* 11, 1–31.

Appendix A

The model developed in Sections II – IV attempts to capture important aspects of the recent experience of intergovernmental fiscal relations in a number of countries, albeit in an idealized and simplified fashion. Since the problems that the paper addresses have not been the subject of systematic formal investigation in the literature of fiscal federalism, there is no obvious precedent to which one can appeal for justification of the specification utilized in Sections II – IV. Thus, although a thorough discussion of the experience of intergovernmental fiscal relations in different parts of the world goes well beyond the scope of the present essay, a brief (and necessarily highly selective) review should help to motivate and justify some of the features of the formal model as well as to illustrate the practical importance of the issues at stake. Wildasin (forthcoming) provides some further discussion and references. This appendix also briefly discusses some related formal literature that deals with soft budget constraints for state-owned enterprises in transition economies.

As one might anticipate, intergovernmental fiscal relations and fiscal decentralization are of critical importance in the two largest countries in the world, China and India. The economic, legal, political, and other conditions of these countries are dissimilar in many respects, but fiscal decentralization and macroeconomic stability are topics of great current interest in both. In China, fiscal decentralization has accompanied the general economic decentralization of the past decade of economic reform.¹³ 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 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).

Since independence, India has implemented an elaborate system of fiscal assistance from the central to the state governments through the Planning Commission and Finance Commission. In addition to outright grants, the center has provided implicit transfers to state governments in the form of loans at concessionary rates. Prompted in part by an effort to control the fiscal deficit at the central government level, these transfers have been curtailed in recent years, with loans from the center being offered at less favorable rates than previously. While lending to states at higher interest rates improves the fiscal balances of the central government, it simultaneously increases the fiscal burden on the states, a number of which are now in fiscal crisis (The World Bank, 1995b).

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

¹³ See Qian and Weingast (1996) for a more detailed discussion of fiscal/political decentralization in China and its interactions with the overall process of economic reform.

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]).

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.¹⁴ The examples of these and other countries raise doubts about the appropriateness and viability of fiscal decentralization, especially in developing and transition economies, giving rise to an important policy debate outlined in the Introduction.

No simple economic analysis can hope to address every important aspect of this very complicated issue, which indeed warrants much more attention than it has received to date. The approach taken here 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 govern-

¹⁴ The Brazilian case is discussed further below. In addition to the sources cited above, recent trends toward fiscal decentralization in Europe, Latin America, China, and elsewhere, and the policy debates surrounding these trends, are treated, *e.g.*, in Bird (1994), the InterAmerican Development Bank (1994), Owens and Panella (1991), Roy (1995), and Shah (1994). 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. By contrast, authors such as McKinnon and Nechyba (1997) argue that fiscal decentralization may enhance rather than weaken fiscal discipline and that central governments are more likely to face soft budget constraints and to engage in excessive borrowing.

ments; 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 legislatures and do not fall under the jurisdiction of state-appointed financial control boards.¹⁵ 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, one should consider whether fiscal decentralization might actually make soft budget constraints *less* rather than *more* likely (Wildasin [1996]). 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?¹⁶

¹⁵ The US experience with local fiscal crises is reviewed in Noto and Rymarowicz (1995), a study undertaken in order to develop a strategy for dealing with the fiscal crisis of Washington, DC.

¹⁶ An interesting recent study by Hillbrecht (1995) identifies state-owned banks in Brazil as an obstacle to hard state budget constraints and to monetary and fiscal discipline. (I

There is a wide range of approaches to the modeling of local fiscal crises and soft budget constraints, and different approaches are more or less useful for different purposes. For example, 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.¹⁷ Such special precipitating events would presumably play a critical role in explaining the *timing* of bailouts. They play no role in the present analysis, however, which focuses entirely on a structural explanation for soft budget constraints. Metaphorically, one can think of stochastic shocks as pushing some fiscally-distressed lower-level governments “over the brink”; the present analysis is concerned rather with why some localities, but perhaps not others, may stray *close* to the brink in the first place.

It is also natural to think of the fiscal relations between central and lower-level governments in terms of a repeated game, where “bailouts” may occur at certain stages of the game, and to suppose that central governments face 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 also play no role in the present analysis, although they are surely important elements of the complete story of soft budget constraints for lower-level governments. Indeed, a series of recent contributions investigates how soft budget constraints arise for state-owned enterprises and how decentralization of capital-market institutions may harden

am grateful to L. Allston for bringing this study to my attention.) Hillbrecht develops 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 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 analysis in the present paper may help to explain why large (or “more important”) subnational units face soft budget constraints while small ones do not.

¹⁷ For example, see Inman (1995) for a clear discussion of the Philadelphia case.

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 in the present paper, however, differs in a basic way from the literature on the structure of credit markets. In particular, it emphasizes 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.

Appendix References

- 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.
- 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.
- 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. (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.
- 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).
- Owens, J. and G. Panella (eds.) (1991), *Local Government Finance: An International Perspective* (Amsterdam: North-Holland).
- 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.
- von Thadden, E.-L. (1995), "Long-Term Contracts, Short-Term Investment, and Monitoring," *Review of Economic Studies* 62, 557-575.
- Wildasin, D.E. (1996), "Comments on '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, 1995* (Washington, DC: The World Bank, 1996), 323-328.
- 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).

Appendix B

Proof of Proposition 1:

Suppose that locality i chooses its contribution at the first-best optimal level $c'_i = z^*$. From (6), (10) must hold as a strict inequality at a solution to (B). By continuity, this inequality must hold for any c'_i sufficiently close to z^* .

Q.E.D.

Proof of Proposition 2:

- (a) As noted in the text, (12) is sufficient to insure that $\gamma(0) > 0$. By continuity, $\gamma(c'_i) > 0$ for any c'_i sufficiently close to zero.
- (b) This follows from (a) and as a corollary to Proposition 1.
- (c) For all c'_i such that $g'_i > 0$, (10) must hold as an equality. Since $D < 0$, one can solve (10) implicitly for g'_i as a function $g'_i = \gamma(c'_i)$ of c'_i . Totally differentiating (10) and using (9),

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

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.

Q.E.D.

Proof of Proposition 3: The result follows directly from (14').

Q.E.D.

Proof of Proposition 4:

Assuming that preferences satisfy (A), the unique bailout level $\bar{c}'_i$ that the center provides is determined by the first-order condition (10), which now takes the form

$$C'(Z)\beta'(\bar{c}'_i) \leq 1 \tag{10'}$$

where

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

This condition is satisfied as an equality if the bailout $\bar{c}'_i$ is strictly positive, which will be true if and only if (12) holds, i.e.,

$$C'(Z^* - n(\beta(z^*) - \beta(0)))\beta'(0) > 1. \tag{12'}$$

For a given utility function satisfying (A) and for a given external benefit function $\beta(\cdot)$, whether or not condition (12') holds depends only on the value of n . In particular, by (6), (12') cannot hold for n sufficiently small. However, since

$$\frac{dZ}{dn} = -(\beta(z^*) - \beta(0)) < 0$$

and since $C(\cdot)$ is strictly concave in Z , it follows that there is a critical value of jurisdiction size $\bar{n} \in (0, N]$ such that (12') does not hold for all $n \leq \bar{n}$ while it does hold for all $n > \bar{n}$. This proves (a) and (b).

To prove (c), use (10') to solve implicitly for $\bar{c}'_i$ as a function of n . Implicitly differentiating,

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

Q.E.D.

Proof of Proposition 5:

The proof follows from the analysis of the symmetric case and is only sketched briefly. Note first that the formal analysis underlying the characterization of the central government bailout policy focuses on deviations by a single locality, holding the policies of all other localities fixed. Also, since there are no technological scale economies in the model, the first-best consumption vector of each household is independent of the size of the jurisdiction in which it resides. It is simple to verify, therefore, that the bailout policy of the central government in relation to a locality of a particular size is independent of the sizes of the jurisdictions that partition the *remaining* population. (Thus, for example, the first-order condition (10) for the optimal bailout policy in dealing with a locality i is completely unaffected by the size of jurisdiction i , except that the matching rate in (10) becomes the size-adjusted m_i^* rather m^* .) Thus, Proposition 4 shows that jurisdictions below a certain size will not be offered bailouts, *irrespective* of the sizes of other jurisdictions.

Q.E.D.