

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Besfamille, Martin<sup>1</sup>; Lockwood, Ben

Article Title: Bailouts in federations: is a hard budget constraint always best?

Year of publication: 2008

Link to published version: <http://dx.doi.org/10.1111/j.1468-2354.2008.00490.x>

Publisher statement: The definitive version is available at [www.blackwell-synergy.com](http://www.blackwell-synergy.com)

# Bailouts in Federations: is a Hard Budget Constraint Always Best?\*

Martin Besfamille<sup>†</sup>

Ben Lockwood<sup>‡</sup>

Manuscript received July 2005, revised April 2007, Accepted June 2007.

## Abstract

This paper analyses hard and soft budget constraints in a federation, where there is a moral hazard problem between the central and the regional governments. Regional governments can avoid a bailout from the centre by exerting costly effort. In this setting, a hard budget constraint is not always optimal because it can provide excessive incentives for high effort, and thus discourage investment that is socially efficient. Thus, a hard budget constraint can imply the opposite kind of inefficiency that emerges under a soft budget constraint, where it is well-known that the common pool problem can give rise to low effort and overinvestment.

**JEL Classification:** D82 - H77.

**Keywords:** Fiscal federalism - Moral Hazard - Bailouts - Hard budget constraints.

---

\*The first version of this paper was published as Warwick Economics Research Paper 717, under the title “Are Hard Budget Constraints for Sub-National Governments Always Efficient?”. We thank F. Alvarez, L. Arozamena, M. Ellman, E. Espino, E. Levy Yeyati, A. Neumeyer, P. Sanguinetti, M. Solá, M. Tommasi, A. Trannoy, M. Vigneault, F. Weinschelbaum, I. Werning, three referees and two associate editors for the very constructive comments. We also thank seminar participants at APET (Beijing, 2004), Workshop on Fiscal Federalism (Barcelona, 2005), CSGR (University of Warwick), IAE, Universidad de San Andrés and Universidad Torcuato Di Tella. This paper started when M. Besfamille visited CSGR, whose hospitality and financial support is gratefully acknowledged.

<sup>†</sup>Universidad Torcuato Di Tella, Argentina.

<sup>‡</sup>University of Warwick, United Kingdom.

# 1 Introduction

Fiscal decentralization, the allocation of tax and spending powers to lower levels of government, is now an established policy objective, in many developed and developing countries. Moreover, it is actively promoted as a development strategy by organizations such as the World Bank (Azfar et al. (2001), World Bank (2000)). However, in many countries, fiscal decentralization is not balanced in terms of tax and expenditure assignments, generating vertical fiscal imbalances. For example, in 1998, across OECD countries, the average share of subnational government expenditure not covered by subnational government revenues was 35.4%, with some countries having much higher shares than this (e.g. Belgium, with 54%). Vertical imbalance is often much higher in non-OECD countries, with Peru and South Africa for example, having shares of 73% and 82% respectively in 2001 (World Bank (2005)).

In practice, vertical fiscal imbalances are resolved either by centrally provided transfers to local governments or by sub-national borrowing. But this may generate new problems. Indeed, it is increasingly claimed that one of the costs of fiscal decentralization is that subnational governments (henceforth regional governments, RGs) may face soft budget constraints (SBCs)<sup>1</sup>. Hence they have an incentive to over-borrow, and/or pay insufficient attention to the quality of the investments that their borrowing finances. For example, Ter-Minassian et al. (2004) say, in a survey of national experiences of subnational borrowing management, “Over the last two decades, the deepening of the decentralization process has resulted in many countries in a significant increase in subnational public expenditure and debt.” Similarly, Jin and Zou (2002) find, in a panel study of 32 industrial and developing countries, that vertical imbalance tends to increase the size of government in the aggregate.

The common pool fiscal externality by which over-borrowing can occur has been explored quite thoroughly in the theoretical literature (Wildasin (1997), Goodspeed (2002)). The willingness of the central government (CG) to bail out RGs who cannot pay creates a negative externality if, as is usual, the cost of bailout is met through increases in taxes or reductions in spending nationally; other regions will then partially finance one region’s bailout. In turn, this induces excessive borrowing initially.

In the policy literature, it seems to be widely assumed that, because of this common pool fiscal externality generated by a SBC, a hard budget constraint (HBC) for RGs is always desirable (Ter-Minassian (1997)). Indeed, much of the literature is concerned with the design of institutional mechanisms that will harden the budget constraint. For example, Rodden et al. (2003) and Ter-Minassian et al. (2004) distinguish rules-based and market discipline HBCs. Generally, a rules-based HBC is one where the central government places legal restrictions (or in extreme cases, a ban) on sub-national government borrowing. A market discipline HBC (Lane (1993)) is one where no restrictions *ex ante* are placed on sub-national governments, but where the central government can precommit not to bail out a sub-national government if it cannot repay its loans.

This paper provides a simple model where a SBC, *not* a HBC, may be the only efficient choice. In our model, regions decide whether to provide discrete public goods (projects). The initial cost of the project is covered by the financial resources of the RG. Regions can also, at a cost, exert effort: high effort raises the probability that the project is good. If the project is good, it immediately

---

<sup>1</sup>Probably the best general definition of a soft budget constraint appears in Kornai et al. (2003): “A budget-constrained organization faces a hard budget constraint (HBC) as long as it does not receive support from other organizations to cover its deficit and is obliged to reduce or cease its activity if the deficit persists. The SBC phenomenon occurs if one or more supporting organizations are ready to cover all or part of the deficit.”

generates a non-monetary benefit for the region greater than the initial cost. If it is bad, it generates no benefit unless additional resources are invested (refinancing).

With refinancing, the project pays a non-monetary benefit to the region greater than the additional resource cost, but possibly less than the total resource cost. Thus, once initiated, a bad project should be refinanced, even if it may be best not to initiate it. The RG does not have enough tax revenue to pay for refinancing, but the CG can refinance a bad project in any region via a uniform non-distortionary national income tax. So, unless the CG has precommitted to not refinance the project (a HBC), it will, in equilibrium, refinance it. Without precommitment, there will be de facto a SBC.

In this setting, the SBC need not provide efficient incentives for effort due to the usual common pool fiscal externality. Because the bailout in any region is paid for by all regions, that region may not have sufficient incentive to put in high effort. More importantly, when this is the case, projects may be initiated in the first place, even when it is inefficient to do so. So, a SBC may lead to overprovision of the public good, a well-known result.

Our novel finding is that with a HBC there is a symmetric inefficiency. Now, RGs may be overincentivised to provide effort, as the payoff to the RG in the event of a bad project is excessively low (i.e. zero), as the project is terminated. Thus, effort exerted may be inefficiently high. More importantly, when this is the case, projects will not be initiated in the first place, even when it is efficient to do so. So, a HBC may lead to underprovision of the public good. Therefore, even in the case where the CG can credibly and costlessly precommit to a HBC, under some circumstances, it may strictly prefer not to do so. This finding is robust to a number of extensions of the basic model, e.g. distortionary taxation, a continuum of effort levels, as discussed in Section 5.

Related literature is discussed in detail in Section 6. Specifically, we address the issue of how our model and results are related to the important papers of Dewatripont and Maskin (1995) and Qian and Roland (1998). In these two papers, a HBC set by the principal always induces efficient behavior by the agent, whereas in our model, this is not the case.

The layout of the remainder of the paper is as follows. Section 2 presents the model. Section 3 derives the main results in the simplest case, where bad projects are ex ante efficient. In this case, it is always efficient to initiate the project, even if it is bad, but this may not occur in equilibrium with a HBC. Section 4 shows that these results, notably the inefficiency of the HBC, extend straightforwardly to the case where, as in Kornai et al. (2003), bad projects are ex ante inefficient, i.e. it would be best not to initiate them if they could be identified. In this case, not initiating a bad project can also be efficient. But, in spite of this additional complexity, the qualitative results on the inefficiency of the HBC continue to hold. Section 5 presents some extensions and Section 6 discusses related literature. Finally, Section 7 concludes.

## 2 The Model

The model has three periods  $t = 0, 1, 2$  and several regions  $i \in N = \{1, \dots, n\}$ , with  $n \geq 2$ . Each region has a continuum of identical residents of measure 1, each of whom has an endowment  $\omega$  of a private consumption good in the last period,  $t = 2$ . In the last period, each resident derives utility from consumption of this good and in periods  $t = 1, 2$  may also benefit from a discrete public good (a *project*), provided by the regional government (RG).

Projects have the following characteristics. At the beginning of  $t = 1$ , initial investment in a project costs  $c$ . Projects may be good or bad: the determination of the project type is described

below. If the project in region  $i$  is good, it generates a non-financial benefit for all the residents of region  $i$  of  $b_g > c$  at the end of  $t = 1$ . If the project is bad, it generates no benefit at this date, but requires an additional input of  $c$  (refinancing) in order to be completed at  $t = 2$ . Upon completion, a bad project generates a non-financial benefit  $c < b_b < b_g$ .

We assume that RGs do not have enough resources to refinance a bad project and thus it is the central government (CG) that decides on refinancing in every region. In that case, the CG refinances using a uniform<sup>2</sup> national proportional tax  $\tau$  on the endowment  $\omega$ . Throughout, we assume that governments are benevolent, i.e. both regional and central governments maximize the sum or average of the welfares of the residents of their jurisdictions over the three periods. There is no discounting of future payoffs.

The order of events, and relevant features of the model in more detail, are as follows. At  $t = 0$ , the central government decides whether to commit to no refinancing (a hard budget constraint, HBC) or not. If it does not, because  $c < b_b$ , in equilibrium, it will refinance bad projects in all regions at  $t = 2$ , i.e. set a soft budget constraint (SBC).

At  $t = 1$ , first the RGs  $i \in N$  simultaneously choose  $d_i \in \{0, 1\}$ , where  $d_i = 1$  denotes a decision to invest in a project in region  $i$ , at a cost of  $c$  of the private good. We assume that RG  $i$  has just initial resources of  $c$ , i.e. no resources left for bailout, should the project be bad.

Then, having observed  $d_1, \dots, d_n$ , the RGs  $i \in N$  who have set  $d_i = 1$  simultaneously take some action  $e_i \in \{0, \Delta\}$ , interpreted as effort, which can improve the quality of the project. If region  $i$  exerts high effort  $e_i = \Delta$ , its project will be good. If effort is low ( $e_i = 0$ ), its project is good with probability  $1 > p \geq 0$  and bad with probability  $1 - p$ . To make it clear that our argument does not rely on asymmetric information, we assume that effort levels  $e_1, \dots, e_n$  can be observed by the CG and the RGs. Finally, the benefit of a good project in region  $i$  is realized and observed by all players. Thus the project type (good or bad) is revealed at this stage. We take the cost of action  $e$  to be just  $e$ . For convenience only<sup>3</sup>, we assume that this cost is non-monetary (e.g. disutility of effort for the RG).

At  $t = 2$ , if the project is bad, the central government, depending on the decision adopted at  $t = 0$ , refinances it or not. If it refinances, the additional cost is met by the tax  $\tau$  as described above. Finally, the bad project's benefit in region  $i$  is realized and consumption of the private good takes place.

Under the assumptions made, the above is a game of complete information played between the  $n$  RGs and the CG. We focus on the (generically unique<sup>4</sup>) sub-game perfect equilibrium of this game: we refer to this just as the *equilibrium* in what follows. Moreover, as regions are ex ante identical, this equilibrium will be symmetric: that is, whatever action the CG takes at  $t = 0$ ,  $d_i = d_j$ ,  $e_i = e_j$  in equilibrium.

Finally, two points are worth noting. First, unlike much of the literature, we assume that the

---

<sup>2</sup>To be more specific, we assume that the CG can neither differentiate regional taxes nor pay side-transfers to the regions.

<sup>3</sup>If the cost is monetary, then we suppose that the initial resources of the RG are just enough to finance the initiation and possible quality upgrading of the project, i.e.  $c + \Delta$ . In turn,  $\Delta$  must therefore be too small to allow the RG to refinance a bad project by itself, i.e.  $\Delta < c$ . Then, subject to this constraint, the analysis goes through as before.

<sup>4</sup>This equilibrium is generically unique because the decision problem of RG  $i$  is always independent of the decision problem of RG  $j$ : this point is further discussed below. Moreover, we resolve non-uniqueness in the special cases where parameters are such that players are indifferent between actions by adopting the following tie-breaking rules: RGs always make initial investments and supply high effort if they are indifferent between doing so and not, and the CG refinances when it is indifferent between doing so and not.

CG can commit to a HBC if it wishes before investment decisions are taken. This is because we want to address the implicit assumption made in the policy literature, i.e. that if it is *possible* to design institutions to achieve a HBC, then it is always *desirable* to do so. Second, note that it makes no difference to our analysis whether the CG can directly observe effort or only infer it from the project outcome: it only has two instruments (the uniform tax and the commitment or not to a HBC) to control the RG. So, for concreteness, we assume that effort is observable and verifiable.

### 3 The Base Case

To make the basic argument as clearly as possible, for the moment, we make the following assumption:

**A1.**  $b_b > 2c$

This assumption says that bad projects are in fact ex ante worth initiating. To further simplify, and make comparisons with Qian and Roland (1998), we also assume in this section that  $p = 0$ , i.e. low effort always generates a bad project.

Consider first a social planner who maximises the sum of regional utilities, and makes all the decisions that are above divided between the two levels of government. Solving the social planner's decision problem backward, first, note that as  $b_b > c$ , refinancing of a bad project<sup>5</sup> is always better than termination. Thus, the net social benefit from low effort, given that the project will be continued, is  $b_b - 2c$ . The net social benefit from a high level of effort in any region is  $b_g - \Delta - c$ . High effort is optimal if

$$b_g - \Delta - c \geq b_b - 2c \iff \Delta \leq b_g - b_b + c \equiv \Delta^*(c) \quad (1)$$

This condition is shown in Figure 1. First, the possible parameter values satisfying A1 comprise the region below the dotted line. In regions A and B of that figure, high effort is optimal, whereas in the other regions C, D and E of that figure, low effort is optimal. Thus, the boundary between regions B and C is given by the relationship  $\Delta^*(c)$  in (1).

Next, we consider the investment decision. When  $\Delta \leq \Delta^*(c)$ , the social planner invests in the project if  $\Delta \leq b_g - c$ . But as  $\Delta^*(c) < b_g - c$ , due to A1, investment always occurs when  $\Delta \leq \Delta^*(c)$ . When  $\Delta > \Delta^*(c)$ , the social planner undertakes the project if  $b_b \geq 2c$ , which also holds by A1. So, *investment is always efficient*. Observe that, when  $\Delta > \Delta^*(c)$ , it is socially optimal to invest and then to exert low effort.

Now we solve for the equilibrium outcome, first in the case where the CG does commit to a HBC at period 0. With a HBC imposed, there is no interaction between the regional governments when they choose their project funding and effort decisions. Each region  $i$  faces a simple decision problem: choose  $d_i \in \{0, 1\}$ , followed by  $e_i \in \{0, \Delta\}$ .

So, the (unique) equilibrium of the model is simply the solution to this decision problem for any region. Assuming that investment has occurred, the benefit to the region from a high level of effort is  $b_g - \Delta - c$ . The benefit from low effort, given that the project will be terminated if bad, is  $-c$ . High effort is thus optimal for the region if

$$b_g - \Delta - c \geq -c \iff \Delta \leq b_g \equiv \Delta^{HBC} \quad (2)$$

---

<sup>5</sup>Note that this observation states that the refinancing decision in any region are independent of what happens in any other region. This is due to the fact that utility is linear in income and the social planner maximizes the sum of regional utilities.

From (1) and (2), obviously,  $\Delta^{HBC} > \Delta^*(c)$  : as termination of the bad project is worse than refinancing, with a HBC, the incentive to put in high effort increases. This simple fact drives all the really new results in this paper.

Next, we consider the investment decision. When  $\Delta \leq \Delta^{HBC}$ , the benefit from the project is  $b_g - \Delta - c$ , so the region undertakes it if  $\Delta \leq b_g - c$ . When  $\Delta > \Delta^{HBC}$ , the benefit from the project is  $-c$ , so the region never undertakes the project.

To summarize, the unique HBC equilibrium outcome is as follows. If  $\Delta \leq b_g - c$ , investment occurs, and high effort is supplied in all regions. Otherwise, there is no investment in equilibrium in any region. This is also shown in Figure 1. In regions A, B and C of that figure, the project will be undertaken and high effort is chosen; whereas in regions D and E, the project will simply never be initiated.

Finally, consider the case of a SBC. Now, a bad project will be refinanced by the CG, so the region pays a bailout cost of only  $c/n$ , because the cost of the bailout is shared equally across all regions via the uniform tax<sup>6</sup>. Note the *common pool fiscal externality* generated by the aggregate budget constraint: any resident of  $i$  is affected negatively by the possibility of a bad project in regions  $j \neq i$ , as all regions bear  $1/n$  of the cost of bailout of any particular region.

So, assuming that initial project investment has occurred, the benefit to the region from a high level of effort is  $b_g - \Delta$ . The benefit from low effort, given that the project will be bailed out, is  $b_b - c/n$ . High effort is thus optimal for the region if

$$b_g - \Delta \geq b_b - c/n \iff \Delta \leq b_g - b_b + c/n \equiv \Delta^{SBC}(c) \quad (3)$$

Clearly, from (1) and (3),  $\Delta^{SBC}(c) < \Delta^*(c)$ . Next, we consider the investment decision. When  $\Delta \leq \Delta^{SBC}(c)$ , the RG of  $i$  anticipates that all regions - including region  $i$  itself - will exert high effort if investment is made and so no bailout will occur. So it invests if  $\Delta \leq b_g - c$ . But as  $\Delta^{SBC}(c) < b_g - c$ , due to assumption A1, investment always occurs when  $\Delta \leq \Delta^{SBC}(c)$ . When  $\Delta > \Delta^{SBC}(c)$ , the RG of  $i$  anticipates that all regions - including region  $i$  itself - will exert low effort if investment is made. In this case, the project will be bad and it will impose expected bailout cost  $c/n$  on region  $i$  itself. So, the benefit to region  $i$  of investment is  $b_b - c - c/n$ , independently of how many other regions invest. Again, due to assumption A1, this benefit is always positive.

Hence, to summarize, the unique SBC equilibrium is as follows. First, investment always occurs. Then, if  $\Delta \leq \Delta^{SBC}(c)$ , high effort is made, whereas if  $\Delta > \Delta^{SBC}(c)$ , low effort is made. This is shown in Figure 1, where high effort is made in region A, and low effort is made in all other regions. Moreover, the boundary between regions A and B is given by the relationship  $\Delta^{SBC}(c)$  in (3).

So, summarizing our findings so far, we have the main result of the paper:

**Proposition 1.** *When  $b_b > 2c$  and  $p = 0$ , it is always efficient to initiate the project. This will always occur in equilibrium with a SBC, but not always with a HBC (regions D and E in Figure 1). If the project is initiated, equilibrium effort will be inefficiently low with a SBC and inefficiently high with a HBC. So, there are parameter ranges where only a SBC is efficient (regions C, D and E in Figure 1), and also where only a HBC is efficient (region B in Figure 1). One of the two is always efficient, and so the efficient regime will always be chosen at  $t = 0$  by the CG.*

This is the key result of the paper: central government may prefer a SBC to a HBC, even if it can commit to a HBC at zero cost.

---

<sup>6</sup>This is proved more rigorously and generally in the Appendix: see the Proof of Proposition 4.

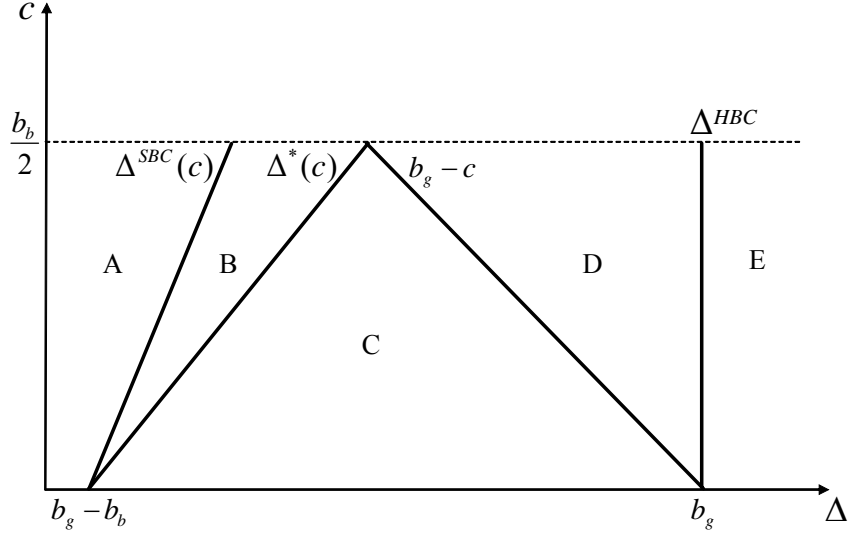


Figure 1: The base case

How does the relative size of the regions in Figure 1 vary with  $n$ , the number of regions? One might think that as  $n \rightarrow \infty$ , so the common pool fiscal externality generated by the SBC gets larger and larger, the HBC might always be efficient, i.e. region B might dominate. In fact, this is not the case. As  $n$  increases,  $\Delta^{SBC}(c)$  decreases at any given  $c$ , and eventually becomes vertical in the diagram. This implies that as  $n$  increases, the region B where the HBC is the only efficient regime increases in size. But, on the other hand, the size of the regions where the SBC equilibrium is the only *efficient* regime (i.e. C, D, E) do not depend upon the number of regions. Therefore, when  $n$  increases, there will always exist a parametric region of fixed size where only the SBC is optimal.

## 4 The Case of Dynamic Inconsistency

The basic argument has been made under the assumption A1 that initiating a bad project was ex ante efficient. This assumption, while convenient, is restrictive for two reasons. First, in this case, it is *always* efficient to initiate the project, even if it turns out to be bad, but initial investment may not occur in equilibrium with a HBC. So, it may be that under assumption A1, the model is biased against finding efficiency under the HBC. Second, an important branch of the literature on the SBC, surveyed in Kornai et al. (2003) [KMR], assumes that bad projects are ex ante inefficient, i.e. it would be best not to initiate them if they could be identified. This assumption is in fact crucial in the KMR model, for reasons explained in Section 6 below. In this section, following KMR, we assume that:

**A2.**  $b_b < 2c$

Together with the initial assumption  $c < b_b$ , A2 generates a dynamic inconsistency problem: if a bad project could be identified ex ante, it would be inefficient to invest in it, but once initiated, it is ex post efficient to refinance it. In this case, if  $p = 0$ , it is of course always efficient to choose



high effort if the project is initiated<sup>7</sup>, in which case a HBC cannot (by assumption) be inefficient. So, to make the analysis interesting, we assume  $p > 0$  in this section.

It turns out that, with some minor exceptions, the results of this section are qualitatively the same as in the base case above<sup>8</sup>. So, here, we simply present the results, highlighting any differences with the base case. All proofs are in the Appendix. We begin with the case of the social planner. The optimal effort and investment decisions are described as follows:

**Proposition 2.** *When  $0 < \frac{b_b}{2} < c < b_b$  and  $0 < p < 1$ , efficiency is characterized as follows. Let  $\Delta^*(c) = (1-p)(b_g + c - b_b)$  and  $c^* = \frac{pb_g + (1-p)b_b}{2-p} \in [\frac{b_b}{2}, b_b)$ . If  $\Delta \leq \min\{\Delta^*(c), b_g - c\}$ , investment and high effort are efficient in all regions. If  $\Delta > \Delta^*(c)$  and  $c \leq c^*$ , investment and low effort are efficient in all regions. Otherwise, no investment in any region is efficient.*

The main difference with respect to the previous section is that now initial investment is not always efficient. This is simply because now the option of investing followed by low effort is no longer guaranteed to generate a positive net benefit.

Now we solve for the equilibrium outcome, first with a HBC. With a HBC imposed, there is again no interaction between the regional governments. So, the (unique) equilibrium of the model, and its (in)efficiency, is easily characterized.

**Proposition 3.** (i) *When  $0 < \frac{b_b}{2} < c < b_b$  and  $0 < p < 1$ , with a HBC, the unique equilibrium outcome is as follows. Let  $\Delta^{HBC} = (1-p)b_g$  and  $c^{HBC} = pb_g$ . If  $\Delta \leq \min\{\Delta^{HBC}, b_g - c\}$ , investment occurs and high effort is supplied in all regions. If  $\Delta > \Delta^{HBC}$  and  $c \leq c^{HBC}$ , investment occurs and low effort is supplied in all regions. Otherwise, there is no investment in equilibrium in any region.*

(ii) *With a HBC, if the equilibrium is inefficient, the inefficiency takes one of the following two forms: either (a) if  $\Delta > b_g - c$  and  $c \in [c^{HBC}, c^*]$ , investments are not made in equilibrium when it is efficient to do so; or (b) if  $\Delta^*(c) \leq \Delta \leq \min\{b_g - c, \Delta^{HBC}\}$ , investments are made in equilibrium when it is efficient to do so, but high effort is supplied when it is efficient to supply low effort.*

As in the base case, for some parameter values, inefficiency *can* occur under a HBC, and involves either *undersupply of projects* or *oversupply of effort*. So, qualitatively, the results for the HBC are the same as in the base case.

Finally consider the case of a SBC. Again, characterization of the equilibrium, and its (in)efficiency properties, is straightforward:

**Proposition 4.** (i) *When  $0 < \frac{b_b}{2} < c < b_b$  and  $0 < p < 1$ , with a SBC, the unique equilibrium outcome is as follows. Let  $\Delta^{SBC}(c) = (1-p)(b_g - b_b + \frac{c}{n})$  and  $c^{SBC} = \frac{n[pb_g + (1-p)b_b]}{n+1-p}$ . If  $\Delta \leq \min\{\Delta^{SBC}(c), b_g - c\}$ , investment occurs and high effort is supplied in all regions. If  $\Delta > \Delta^{SBC}(c)$  and  $c \leq c^{SBC}$ , investment occurs and low effort is supplied in all regions. Otherwise, there is no investment in equilibrium in any region.*

(ii) *With a SBC, if the equilibrium is inefficient, the inefficiency takes one of the following two forms: either (a) if  $\Delta \geq b_g - c$  and  $c \in [c^*, c^{SBC}]$ , investments are made in equilibrium when*

<sup>7</sup>Because of  $p = 0$ , it is always preferable not to initiate the project (thus saving set-up cost  $c$ ) rather than initiate it and subsequently choose low effort.

<sup>8</sup>Throughout this section we assume that the parameters  $n, p, b_b$  and  $b_g$  are such that

$$\frac{b_b}{2p} < b_g < \frac{1+p(n-1)}{np} b_b$$

holds. These conditions enable us to obtain clear-cut results. But if they were not to hold, the qualitative results of this sections would still remain.

it is inefficient to do so; or (b) if  $\Delta^{SBC}(c) \leq \Delta \leq \min\{b_g - c, \Delta^*(c)\}$ , investments are made in equilibrium when it is efficient to do so, but low effort is supplied when it is efficient to supply high effort.

Comparing Propositions 2, 3 and 4 to Proposition 1, we see that both the equilibrium possibilities and the nature of inefficiency are qualitatively different to the base case, but these differences are relatively minor. In particular, first, it is possible that investment does not occur in equilibrium with a SBC. Second, there can be investment in equilibrium with a SBC when it is not efficient, a possibility that could not occur in the base case. This result is well-known and is driven by the common pool fiscal externality that a SBC generates. But, the central point, i.e. that under some conditions, the government may not wish to precommit costlessly to a HBC, still holds.

To investigate this point further, we can now construct the analog of Figure 1, to see in which regions of the parameter space the CG may prefer a SBC instead of a HBC.

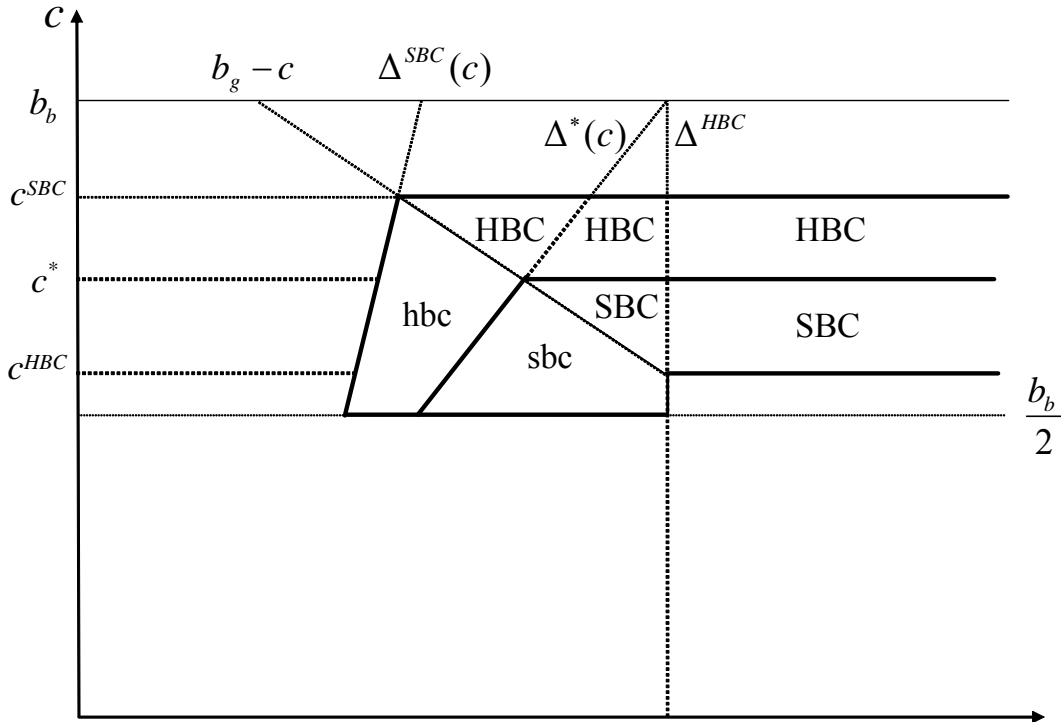


Figure 2: Comparison of regimes in the dynamic inconsistency case

In the parametric regions denoted by “HBC”, only the HBC is efficient because with a SBC, investments are made when it is inefficient to do so. On the other hand, in the parametric regions denoted by “SBC”, only the SBC is efficient because with a HBC, investments are *not* made when it *is* inefficient to do so. Next, in the parametric region denoted by “hbc”, a HBC is optimal because all regions exert high effort (the socially efficient solution) and thus generate good projects. Similarly, in the parametric region denoted by “sbc”, a SBC is optimal because regions exert low effort, which is socially efficient. As in the base case, one of the two possibilities is always efficient, and so the efficient regime will always be chosen at  $t = 0$  by the CG.

## 5 Extensions

### 5.1 Distortionary Taxes

So far, we have assumed that the tax set by the CG is lump-sum. This is of course a strong assumption: in practice, taxes used to fund bailouts will have distortionary effects. Let us model this in the simplest possible way by supposing that the cost of financing a project initially, or bailing out a bad project, is  $c + \delta$ , rather than  $c$ . If this is the case, then the effect on equilibrium is simply to shift the ‘ $c$ ’ on the vertical axis of Figures 1 and 2 upwards, while holding  $\Delta$  fixed. The positive effects of this are obvious. In the case of a HBC, this change will prevent initial investment if  $\delta$  is large enough. With a SBC, this can also occur, and additionally (because  $\Delta^{SBC}(c)$  is upward-sloping, not vertical) there can be a switch from low effort to high effort.

The normative effects are also clear from Figure 1 and 2. First, an increase in  $c$ , holding  $\Delta$  fixed, generally moves the economy from a regime where the SBC is optimal to one where the HBC is optimal. Intuitively, this is because distortionary taxes increase the severity of the common pool problem. Second, given that the CG can precommit, an increase in  $c$  makes the CG worse off, because from Proposition 1, social welfare, given efficient investment and effort decisions, is decreasing<sup>9</sup> in  $c$ .

### 5.2 Continuous Effort

So far, we have assumed only two effort levels. While this is a convenient simplification, it does have one implication that is not robust, namely, that one of the two regimes (HBC or SBC) is always fully efficient. But, when many effort levels are possible (e.g. a continuous effort variable), generally, the level of effort will be inefficient in either regime. To see this, consider the simple case where effort  $e \in [0, 1]$  is a continuous variable, the probability of a good project is  $e$ , and the cost of effort is  $\frac{e^2}{2}$ . Also, to rule out corner solutions, assume  $b_g < 1$ . Finally, again for simplicity, assume A1, i.e. that it is always efficient not to terminate bad projects.

Then, given that the social planner always continues bad projects, he chooses  $e$  to maximize

$$eb_g + (1 - e)(b_b - c) - \frac{e^2}{2} \quad (4)$$

which is maximized by  $e^* = b_g - b_b + c$ ; as  $1 > b_g > b_b - c > 0$ ,  $0 < e^* < 1$ .

Next, with a HBC, the region chooses  $e$  to maximize (4), but where  $b_b - c$  is replaced by 0, implying  $e^{HBC} = b_g > e^*$ . Finally, with a SBC, the region chooses  $e$  to maximize (4), but where  $b_b - c$  is replaced by  $b_b - c/n$ , implying  $e^{SBC} = b_g - b_b + c/n < e^*$ . So, as claimed, effort is inefficient in either regime.

### 5.3 Towards Empirical Testing

Although the model is quite stylized, it does have some distinctive empirical implications. Our model, along with several others (e.g. the Dewatripont and Maskin (1995) model<sup>10</sup>, applied to local public finance) predicts that a HBC will lead to less investment than a SBC. However, *only* our model predicts that conditional on investment, a HBC leads to higher ‘effort’ by the RG than does a SBC. To test the distinctive prediction of our model, then, we need two things.

<sup>9</sup>Specifically, welfare per region is  $W(c) = \max\{b_g - c - \Delta, b_b - 2c, 0\}$ .

<sup>10</sup>In their model, with a HBC, RGs will seek finance only for projects that they know to be good.

First, it is necessary to identify a historical episode where regional governments face switches from a soft budget constraint to a hard budget constraint (or vice versa), preferably in different periods of time. A possible case for such an exercise could be Brazil. It has been acknowledged in many policy papers (Burki et al. (1999) and Rodden (2003) among others), that Brazilian States were, before 2000, under a SBC regime. For example, it was very easy for them to obtain an authorization from the Senate to borrow more than their legal limit. Starting in the late 90's, the Federal Government launched a series of measures aiming to harden the States' budget constraints<sup>11</sup>. Up to now, the Brazilian Federal Government seems to be successful in enforcing these measures. Data on state public investment as a percentage of state income do show a downward trend after 2000, consistently with the theory.

Then, we need to identify empirical proxies for the level of 'effort' in our model. The yearly budget of each State's administration that it is in charge with the monitoring of undertaken public projects could be used as a proxy for the 'effort' variable. Unfortunately, this data is not currently available. Collecting it seems indeed to be a worthy step for future empirical research.

## 6 Related Literature

A key question is how our model and results relate to the standard model of the SBC of Dewatripont and Maskin (1995) [DM] and KMR (2003). In these papers, a HBC is always optimal. The most general statement of the model is in KMR; and we follow their exposition, but using our notation. There, there are two agents or organizations, the centre and the manager. The manager has a randomly determined good or bad project (with probabilities  $\alpha, 1 - \alpha$  respectively). Project types are initially observable only by the manager. Each type of project initially costs  $c$ . Good projects generate payoffs  $(r_g, b_g)$  for the centre and the manager respectively. Bad projects require refinancing of amount  $c$  before they yield benefits  $(r_b, b_b)$  again for the centre and the manager.<sup>12</sup> The order of events is: 1. the manager decides whether to present a project to the centre for funding or not; 2. the centre either funds the projects presented or not; 3. if the project turns out to be bad, the centre decides whether to refinance it or not.

As in our model under assumption A2, refinancing is ex post optimal but ex ante sub-optimal (i.e.  $c < r_b + b_b < 2c$ ), and good projects are ex ante optimal (i.e.  $b_g > c$ ). Also, there is a conflict of interest between the centre and the manager, in that if refinancing occurs, the manager prefers a bad project to a good project  $b_b > b_g$ . So, for high enough  $\alpha$ , the unique equilibrium of this simple model is an inefficient SBC one where (i) the centre finances all projects and refinances bad projects, and (ii) anticipating this, the manager presents a project for funding even if it is bad. Imposing a HBC changes the equilibrium to an efficient one where managers do not present bad projects. So, in their model, a HBC is always welfare-improving.

In our model, the centre can be interpreted as the CG, and the manager as a RG. So, we can see a number of important differences between our model and KMR. First, in our model, the 'manager' makes the initial decision to fund the project or not, and also pays the initial cost. Second, there is no direct project payoff to the centre (the projects are discrete regional public goods), i.e.

---

<sup>11</sup>Law 9496/1997, Senate Resolutions 40 and 43 and National Monetary Council Resolution 2653 controlled both the demand and supply of State credit. In May 2000, the Fiscal Responsibility Law (LRF) and its companion legislation Law 10.028/2000 (that amends the penal code and the laws dealing with impeachable offenses) capped all these control measures. Many observers agree that all these laws represent a clear signal for future governors and mayors that easy bailouts have come to an end in Brazil.

<sup>12</sup>For simplicity, assume project termination gives both agents zero.

$r_g = r_b = 0$ . Third, in our model, the ‘manager’ does pay a share of the refinancing cost, via the national tax (generally,  $1/n$ ). Finally, the project quality is not just randomly determined, but also depends on an investment (effort) by the ‘manager’.

Upon reflection, it is clear that the final assumption is the one driving our key result that a HBC can be inefficient. To see this, make the other three assumptions of our model, but retain the key KMR assumption that ‘managers’ i.e. RGs have randomly determined good and bad projects. Then, an inefficient SBC equilibrium as in the KMR model arises if  $0 < b_b - c - c/n$ . Under this condition, the RG gets a positive benefit from presenting the bad project, anticipating that it will be refinanced.<sup>13</sup> Under these conditions, as in KMR, the imposition of a HBC always restores efficiency.

Next we turn to Qian and Roland (1998) [QR]. How does their model compare to ours? Again, there are a number of differences, but only some are crucial. First, in the terminology of KMR, in QR the ‘centre’ is the RG and the ‘manager’ is the state enterprise, unlike in our model. Second, unlike in our paper, there is no initial project funding decision: projects are assumed already initiated, modelling a legacy from the period of socialist planning. Thirdly, as in our base case of Section 3, project quality can be improved from bad to good by an action taken by the manager (effort). Fourthly, there are two related assumptions on parameters: (i) given a SBC, the manager wishes to put in low effort i.e.  $b_g - \Delta < b_b$  in our notation<sup>14</sup>; (ii) the gains to the manager from low effort are exceeded by the cost to RG in terms of lower payoffs<sup>15</sup>. So, from these two assumptions, in the QR model, the possibility of the SBC being optimal is ruled out by assumption. This means, of course, that a HBC can never be inefficient in their model. This seems to us very restrictive.

Finally, there is also a small public finance literature<sup>16</sup> explaining why, in some circumstances, a SBC may be optimal. This class of models assumes (unlike us) that public projects have positive inter-regional spillovers, and (somewhat<sup>17</sup> like us), in equilibrium with a SBC, RGs anticipate that additional expenditure will be partially subsidized by transfers from other regions. With positive spillovers, this common pool externality offsets the failure to internalize inter-regional spillovers<sup>18</sup>, and may lead to the SBC outcome being more efficient. This is rather different from our story: we do not need inter-regional spillovers from projects to yield the conclusion that only the SBC can be efficient.

## 7 Conclusions

This paper has shown that, in the context of a model of fiscal federalism, hard budget constraints do not necessarily ‘solve’ the problems generated by soft budget constraints. Our argument was the following. First, by definition, a hard budget constraint necessarily implies that ‘bad’ invest-

<sup>13</sup>The only other equilibrium is where  $0 > b_b - c(1 + \frac{1}{n})$ , in which case only good projects are presented. Then there is no need for a HBC to restore efficiency.

<sup>14</sup>In their notation,  $B_s > B_q$ , where implicitly,  $B_q$  is the benefit net of any cost of effort.

<sup>15</sup>This is implicit in this statement: “Notice that because  $E^H > E^S$ , hard budget constraints would yield more tax revenue and thus allow a higher welfare level. Thus, the government would prefer a commitment to no bailout” (p1149). In their paper,  $E^H, E^S$  denote net transfers from the enterprise to RG.

<sup>16</sup>Silva and Caplan (1997), Caplan et al. (2000), Köthenbürger (2004).

<sup>17</sup>In this class of models, the mechanism is slightly different : grants are chosen ex post by the central government to equalize the marginal utility of private consumption between regions.

<sup>18</sup>Remarkably, if the spillovers are such that the regional public good is effectively a contribution to a pure national public good, the two effects exactly offset each other and public good provision is efficient, even in the absence of a matching grant instrument at the federal level (Caplan et al. (2000)).

ment projects will be terminated, when ex post, it would be desirable to continue them through additional financing. Now suppose that by exerting some effort, regional governments can lower the probability that projects are 'bad'. Then, with a hard budget constraint, regional governments are overincentivised to provide effort, as payoff to them in the event of a bad project is excessively low. Thus, effort exerted may be inefficiently high. More importantly, when this is the case, some projects will not be initiated in the first place, even when it is efficient to do so. So, a hard budget constraint may lead to underinvestment. The argument developed in this paper is quite general, and may apply to areas other than fiscal federalism, like, for example, in a lender-borrower relationship or in a government-public enterprise regulatory context. That is a topic for future research.

## A Appendix

**Proof of Proposition 2.** The net social benefit from high effort in any region is  $b_g - \Delta - c$ . For a social planner, refinancing is always better. So the net social benefit from low effort, given that the project will be continued, is  $pb_g + (1-p)(b_b - c) - c$ . Thus high effort is optimal if  $\Delta \leq (1-p)(b_g + c - b_b) = \Delta^*(c)$ . Next, consider the decision whether to invest in the project. When  $\Delta \leq \Delta^*(c)$ , the social planner undertakes the project if  $\Delta \leq b_g - c$ . When  $\Delta > \Delta^*(c)$ , the social planner undertakes the project if  $pb_g + (1-p)(b_b - c) - c \geq 0$ , or  $c \leq \frac{pb_g + (1-p)b_b}{2-p} = c^*$ . By assumption,  $c^*$  lies between  $\frac{b_b}{2}$  and  $b_b$   $\square$

**Proof of Proposition 3.** To prove (i), note first that high effort is optimal for the region if  $\Delta \leq (1-p)b_g = \Delta^{HBC}$ . Next, consider the decision whether to initiate the project. When  $\Delta \leq \Delta^{HBC}$ , the region undertakes the project if  $\Delta \leq b_g - c < \Delta^{HBC}$ . When  $\Delta > \Delta^{HBC}$ , the region undertakes the project if  $c \leq pb_g = c^{HBC}$ . Comparing critical values  $\Delta^*(c)$ ,  $\Delta^{HBC}$  and  $c^*$ ,  $c^{HBC}$ , and noting that  $c^{HBC} < c^*$ ,  $\Delta^*(c) < \Delta^{HBC}$ , we can establish part (ii)  $\square$

**Proof of Proposition 4.** To prove (i) note first that conditional on  $d = (d_i)_{i \in N}$ ,  $e = (e_i)_{i \in N}$ , the expected utility of a resident of region  $i$  is

$$u_i(d, e) = \omega(1 - E\tau) + d_i[p(e_i)b_g + (1 - p(e_i))b_b] - e_i - d_i c \quad (5)$$

where

$$p(e) = \begin{cases} 1 & \text{if } e = \Delta \\ p & \text{if } e = 0 \end{cases} \quad (6)$$

is the probability of a good project as a function of  $e$  and  $E\tau$  is the expected tax. The latter satisfies, from the CG budget constraint:

$$n\omega E\tau = c \sum_{j=1}^n d_j(1 - p(e_j)) \quad (7)$$

That is, the tax must fund refinancing of a project in every region  $j$  in the event that it is bad, which occurs with probability  $d_j(1 - p(e_j))$ . As this tax is uniform and wealth  $\omega$  is immobile, the base of the tax  $\tau$  is simply  $n\omega$ . Substituting (6) and (7) into (5), we obtain, suppressing the constant  $\omega$ :

$$\begin{aligned} u_i(d, e) &= d_i[p(e_i)b_g + (1 - p(e_i))b_b] - \frac{c}{n} \sum_{j=1}^n d_j(1 - p(e_j)) - e_i - d_i c \\ &= \left\{ d_i[p(e_i)b_g + (1 - p(e_i))(b_b - \frac{c}{n})] - e_i - d_i c \right\} - \frac{c}{n} \sum_{j \neq i} d_j(1 - p(e_j)) \end{aligned} \quad (8)$$

By inspection of (8), the effect of  $d_i, e_i$  on  $u_i$  - measured by the bracketed term - is independent of  $d_j, e_j$ . So we can analyse the choice of  $d_i, e_i$  just for a representative region  $i$ .

Assuming  $d_i = 1$ , the gain to region  $i$  from high relative to low effort is easily calculated from (8) and (6) to be  $(1-p)(b_g - b_b + \frac{c}{n}) - \Delta$ . So, for region  $i$ , it is a dominant strategy to put in high effort in project implementation if

$$\Delta \leq (1-p)(b_g - b_b + \frac{c}{n}) = \Delta^{SBC}(c)$$

and low effort otherwise, conditional on the project having been initiated.

Now we turn to the project initiation decision. First, assume that  $\Delta \leq \Delta^{SBC}(c)$ . Then, the government of region  $i$  anticipates that  $p(e_i) = 1$ . From (8), the gain to region  $i$  from initiating a project is  $b_g - \Delta - c$ . So, for region  $i$  is a dominant strategy to invest if  $\Delta \leq b_g - c$ .

Now suppose that  $\Delta > \Delta^{SBC}(c)$ . Then the government of region  $i$  anticipates that  $p(e_i) = p$ . From (8), the benefit to region  $i$  of  $d_i = 1$  is

$$p(b_g - c) + (1 - p)(b_b - c - \frac{c}{n}) \quad (9)$$

So, for any region it is a dominant strategy to invest if (9) is positive, or

$$c \leq \frac{n[pb_g + (1 - p)b_b]}{n + 1 - p} = c^{SBC}$$

as claimed.

Comparing critical values  $\Delta^*(c)$ ,  $\Delta^{SBC}(c)$  and  $c^*$ ,  $c^{SBC}$ , and noting that  $c^{SBC} > c^*$ ,  $\Delta^*(c) > \Delta^{SBC}(c)$ , we can establish part (ii)  $\square$



## References

- [1] Azfar, O., S. Kähkönen and P. Meagher (2001), “Conditions for Effective Decentralized Governance: A Synthesis of Research Findings,” Working Paper 256, IRIS Center, University of Maryland.
- [2] Burki, S., G. Perry and W. Dillinger (1999), “Beyond the Center: Decentralizing the State,” World Bank Latin American and Caribbean Studies 19639.
- [3] Caplan, A., R. Cornes and E. Silva (2000), “Pure Public Goods and Income Redistribution in a Federation with Decentralized Leadership and Imperfect Labour Mobility”, *Journal of Public Economics*, 77, 265-284.
- [4] Dewatripont, M. and E. Maskin (1995), “Credit and Efficiency in Centralized and Decentralized Economies,” *Review of Economic Studies*, 62, 541-555.
- [5] Goodspeed, T. (2002), “Bailouts in a Federation,” *International Tax and Public Finance*, 9, 409-421.
- [6] Jin, J. and H. Zou (2002), “How Does Fiscal Decentralization Affect Aggregate, National, and Subnational Government Size?,” *Journal of Urban Economics*, 52, 270-293.
- [7] Kornai, J., E. Maskin and G. Roland (2003), “Understanding the Soft Budget Constraint,” *Journal of Economic Literature*, 41, 1095-1136.
- [8] Köthenbürger, M. (2004), “Tax Competition in a Fiscal Union with Decentralised Leadership,” *Journal of Urban Economics*, 55, 498-513.
- [9] Lane, T. (1993), “Market Discipline,” *IMF Staff Papers*, 40, 53-88.
- [10] Qian, Y. and G. Roland (1998), “Federalism and the soft-budget constraint”, *American Economic Review*, 88, 1146-1162.
- [11] Rodden, J. (2003), “Federalism and Bailouts in Brazil” in Rodden et al. (Eds.) *Fiscal Decentralization and the Challenge of Hard Budget Constraints*, The MIT Press, Cambridge Massachusetts.
- [12] Rodden, J., G. Eskeland and J. Litvack (2003), *Fiscal Decentralization and the Challenge of Hard Budget Constraints*, The MIT Press, Cambridge Massachusetts.
- [13] Silva, E. and A. Caplan (1997), “Transboundary Pollution Control in Federal Systems”, *Journal of Environmental Economics and Management*, 34, 173-186.
- [14] Ter-Minassian, T. (1997), “Decentralization and Macroeconomic Management”, IMF Working Paper 97/155.
- [15] Ter-Minassian, T., M. Albino-War and R. Singh (2004), “Managing Sub-National Borrowing”, unpublished paper, IMF.
- [16] Wildasin, D. (1997), “Externalities and Bailouts Hard and Soft Budget Constraints in Intergovernmental Fiscal Relations”, mimeo, Vanderbilt University.

[17] World Bank (2000), *Entering the 21st Century*, World Development Report.

[18] World Bank (2005), *Country Database on Decentralization*.