I N S T I T U T O   D E   E C O N O M Í A

DOCUMENTO de TRABAJO

**460**

2015

# Regional State Capacity and the Optimal Degree of Fiscal Decentralization

**Antonio A. Bellofatto, Martín Besfamille**

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

# Regional State Capacity and the Optimal Degree of Fiscal Decentralization*

Antonio A. Bellofatto† Martín Besfamille‡

September 9, 2015

**Abstract**

This paper studies the optimal degree of fiscal decentralization in a federation. In our environment, regional governments are characterized by two dimensions of state capacity; namely, administrative and fiscal. These gauge the ability to deliver public goods and to raise tax revenues, respectively. Two regimes are compared: partial and full decentralization. Under partial decentralization, regional governments have no tax powers and rely on central bailouts to refinance incomplete projects. Under full decentralization, regional governments refinance incomplete projects through capital taxes, in a context of tax competition. We provide a normative comparison between partial and full decentralization and show how the optimal degree of fiscal decentralization hinges on the relative magnitudes of each type of capacity. Specifically, for sufficiently low levels of fiscal capacity bailing out regional governments is optimal, regardless of the level of administrative ability. However, a combination of low levels of administrative capacity and high levels of fiscal capacity calls for fully decentralizing tax powers.

**Keywords:** Fiscal federalism - State capacity - Partial and full fiscal decentralization - Bailouts - Hard budget constraints.

**JEL Codes**: H77.

*We thank C. Agostini, L. Ales, B. Allison, L. Arozamena, A. Auerbach, D. Bogart, P. Boyer, J. Brueckner, L. Cohen, J. Cravino, P. Dal Bo, E. Dardati, R. de Elejalde, A. de Paula, J. Dubra, N. Dwenger, E. Espino, I. Fainmesser, N. Figueroa, F. Forges, E. Giolito, M. González Eiras, B. Guimaraes, C. Hevia, J. Hines, B. Knight, J. C. Hallak, E. Janeba, K. Lee, B. Lockwood, A. Manelli, E. Mattos, P. Mongin, L. Navarro, A. Neumeyer, E. Ornelas, H. Piffano, F. Piguillem, C. Ponce, A. Porto, E. Rezk, C. Ruzzier, M. Sacks, E. Saez, R. Serrano, S. Skaperdas, W. Sosa Escudero, M. Tommasi, U. Troiano, F. Weinschelbaum, seminar participants at Berkeley, Brown, EESP-FGV, Irvine, Michigan, San Andrés, Pontificia Universidad Católica de Chile, Universidad Alberto Hurtado and Universidad Torcuato Di Tella, and attendants to the Annual Congress of the Association for Public Economic Theory (2014), Latin American Meetings of the Econometric Society (2011), Jornadas Internacionales de Finanzas Públicas (2010), Reunión Anual de la Asociación Argentina de Economía Política (2010) and ZEW Public Finance Conference (2015) for useful comments and suggestions. We are also grateful to S. Briole and H. Barahona for excellent research assistance. M. Besfamille gratefully acknowledges financial support from FONDECYT Grant # 1150170. The usual disclaimer applies.

†Tepper School of Business, Carnegie Mellon University. Email: abellofa@alumni.cmu.edu.

‡Corresponding author. Instituto de Economía, Facultad de Ciencias Económicas y Administrativas, Pontificia Universidad Católica de Chile. E-mail: mbesfamille@uc.cl.

# 1 Introduction

In many developed and developing countries, it is common that tax and expenditure assignments to subnational or regional governments be unbalanced. In particular, regional governments are often in charge of delivering local public services, but cannot raise the required revenues to finance these expenditures.[1] These so called "vertical fiscal imbalances" should either be covered by centrally provided transfers,[2] or just bypassed by decentralizing tax powers to regional governments.

The first alternative is widely used in practice and gives rise to an institutional setting defined as *partial decentralization*.[3] As shown by Wildasin (1997) and Goodspeed (2002), under this scenario regional governments may face "soft budget constraints,"[4] which can create negative externalities across regions and induce excessive spending or borrowing.[5] To cope with this problem, a vast literature in public finance has put forward different institutional mechanisms so that regional governments face "hard budget constraints."[6] One mechanism that has attracted significant attention is the complete decentralization of tax powers to subnational governments. In a seminal contribution, Qian and Roland (1998) argue that such a system of *full decentralization* gives rise to tax competition, which in turn raises the perceived marginal costs of public funds at the regional level. But while this regime boosts fiscal discipline, hard budget constraints may also lead to underprovison of local projects, as argued by Besfamille and Lockwood (2008).

The goal of this paper is to provide a novel framework for comparing partial and full decentralization from a normative perspective. Our central departure relative to previous analyzes is to model local governments as being indexed by their level of *state capacity*, defined by Besley and Persson (2010) as the "state's ability to implement a range of policies." This institutional ingredient is of central importance for the study of the optimal degree of fiscal decentralization. In fact, a number of descriptive reviews of decentralization reforms (e.g., Bird (1995), Litvack et al. (1998)) and recent quantitative evaluations of such processes (e.g., Loayza et al. (2014)) argue that the benefits accruing from a decentralized form of government crucially hinge on these abilities. Or, more generally, Prud'homme (1995) and Bardhan and Mookherjee (2006), state that models formalizing pro-decentralization arguments usually ignore key institutional features of local governments.

[1]Eyraud and Lusinyan (2013) report that across OECD countries, the average share of subnational government expenditure not financed through own revenues was 40 percent between 1995 and 2005. In Belgium and Mexico these shares climb to 60 and 83 percent, respectively. Corbacho et al. (2013) document that vertical fiscal imbalances in Latin America are the highest among developing nations.

[2]See Boadway and Shah (2007).

[3]This term has been coined by Brueckner (2009).

[4]According to Kornai et al. (2004), "*A budget-constrained organization faces a hard budget constraint 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 soft budget constraint phenomenon occurs if one or more supporting organizations are ready to cover all or part of the deficit*" (page 1097).

[5]Among others, Pettersson-Lidbom (2010) confirms this theoretical result by estimating that between 1979 and 1992, Swedish local governments increased their debt by more than 20 percent when they expected to receive future bailouts.

[6]See, among others, Rodden et al. (2003), Oates (2005) and Weingast (2009).

We consider an environment in which regional governments decide whether or not to provide a discrete, local public good or project. The initial cost of the project is covered by the regional government's financial resources. If the project is initiated, it is carried out by regional bureaucracies, and there is a probability that the project is finished in due time and generates a social benefit greater than the initial cost. With the complementary probability, the project is delayed and needs a second round of financing to be completed. In this case, it generates a social benefit to the region, but lower than when the project is carried out on time. In the partially decentralized regime, no regional government has tax powers to refinance its incomplete project but the central government can bail out regions. Bailouts are financed via a uniform national tax on local capital, which is imperfectly mobile. Under full decentralization, on the other hand, regional governments have to refinance incomplete projects through a tax on capital invested in their jurisdiction, in a context of tax competition. Equilibrium outcomes under partial and full decentralization can be inefficient, as these regimes can generate overprovision or underprovision of projects, respectively.

Crucially, we make a distinction between two dimensions of local state capacity; namely, *administrative* and *fiscal*.[7] The former measures the ability of subnational governments to produce and deliver public goods and services, and it is proxied by the probability that a project is finished on time. The latter gauges the capacity to raise revenues through local taxes, and it is modeled as the fraction of the potential tax base that ends up as fiscal revenues of the subnational governments. The model is symmetric *ex ante*, in the sense that all regional governments have the same level of state capacity, and all costs and benefits are identical across regions. However, outcomes can be different *ex post* because some projects are completed before others.

Our main results indicate how the optimal degree of decentralization hinges on the relative magnitudes of each dimension of state ability. First, we show that the level of regional fiscal capacity that prevails in the federation is sufficiently low, partial decentralization dominates. Intuitively, refinancing incomplete projects under full decentralization is too costly, regardless of the level of regional administrative capacity. Second, and more interestingly, we demonstrate that full decentralization dominates when combining high levels of fiscal capacity with low levels of administrative capacity. Essentially, under this scenario many projects remain incomplete, but the likelihood of recurring to distortionary refinancing under full decentralization is low. Thus, expected distortions under full decentralization are lower than under partial decentralization. This finding contradicts the views of certain policy proposals suggesting that high levels of administrative capacity are necessary for successful decentralization reforms (see, e.g., Bird (1995)).

We then undertake a comparative statics analysis. We show that when the regional capital stock increases, full decentralization dominates more frequently. Conversely, partial decentralization dominates more often when the highest possible benefit of the project rises. In addition, we show that when the number of regions increases, partial decentralization

[7]Such a decomposition of state capacity has already been considered in other works (see Hanson and Sigman (2013) for a survey). We also provide an empirical motivation for this distinction in the body of the paper.

dominates in a parametric area where full decentralization was initially the optimal regime, and vice versa. Finally, we discuss a special case with unidimensional state capacity, and evaluate the robustness of our results by developing an extension that incorporates distortionary national taxation.

## 1.1 Related Literature

This paper is related to various strands of literature. First, other contributions have compared partially and fully decentralized regimes. Brueckner (2009) presents a Tiebout-type model, where local governments provide a local public good, private developers build houses, and then heterogeneous consumers decide on their location. He shows that when local governments are benevolent, full decentralization always dominates because uniform transfers under partial decentralization generate less variety of local public goods, and thus a worse preference matching. Peralta (2011) presents a political economy model, with both benevolent and rent-seeking local politicians. In her environment, voters cannot observe neither the politician's type nor the cost of the public good being provided. She finds that partial decentralization improves politician's selection (i.e., voting out rent-seekers) whereas full decentralization fosters discipline (i.e., giving incentives to rent seekers to behave as benevolent), and that this last regime dominates when the proportion of rent seekers is low.[8] The main differences between our paper and these contributions is that we study a different trade-off between partial and full decentralization, namely, the right balance between inefficient bailouts and project overprovision vs. capital tax competition and project underprovision. Moreover, neither of the previous articles incorporate regional state capacity into the analysis.

The paper is also related to an important set of contributions that analyze the advantages and disadvantages of different types of regional budget constraints in federations. The optimality of hard budget constraints has been studied by Qian and Roland (1998) and Inman (2003); whereas the possibility that they may be inefficient has been raised by Besfamille and Lockwood (2008). The main differences between Besfamille and Lockwood (2008) and our paper are the following. These authors compare, from a normative point of view, soft and hard budget constraints at the *interim* stage (i.e., project by project), whereas we take an *ex ante* perspective, more suitable for an institutional comparison. Moreover, they do not consider how different levels of regional administrative capacity affect the trade-off between partial and full decentralization, which is one of our main objectives. On the other hand, Wildasin (1997), Goodspeed (2002), Akai and Sato (2008) and Crivelli and Staal (2013) describe how bailouts in federations distort, via a common-pool fiscal externality, decisions at the regional level. Silva and Caplan (1997), Caplan et al. (2000) and Köthenbürger

[8] Other authors have considered environments where partial decentralization is the optimal regime, but they adopt different definitions for partial decentralization. Janeba and Wilson (2011), and Hatfield and Padró i Miquel (2012), for example, define partial decentralization as a regime in which a subset of public goods are exclusively funded and provided by local governments. Joanis (2014), on the other hand, defines partial decentralization as "shared responsibility," that is, an institutional regime where both the central and the regional government participate in the funding of a given public good.

(2004) claim that a regime with decentralized leadership, where the central government sets intergovernmental transfers after regional governments have adopted their own policy, may give a more efficient outcome than a regime with hard budget constraints. This result relies on a second best argument, and thus needs some pre-existing distortion in the form of public goods or tax spillovers to hold. Sanguinetti and Tommasi (2004) analyze the trade-off between hard and soft budget constraints, but in the "rules vs. discretion" tradition.

Finally, the paper is related to a recent literature that empirically studies how, in contexts of decentralized regimes, local state capacity has a positive impact on public outcomes. Steiner (2010) measures local governments' capacity using an index of resources available to local governments, and another one that captures the level of technical and administrative capacity of district governments. Loayza et al. (2014) quantitatively evaluate how local capacity affects municipal budget execution rates in Peru, among other things. Bandyopadhyay and Green (2012) show that a higher percentage of residents from centralized ethnic groups imply higher development indicators at the local level in Uganda. Finally, Acemoglu et al. (2014) study how Colombian municipalities choose the size of their local bureaucracy, as well as its impact on public goods provision. Contrary to what we do, these papers take the intergovernmental institutional setting as given.

The layout of the remainder of the paper is as follows. Section 2 presents the model, and Section 3 describes the efficient benchmark. In Section 4 we analyze partial decentralization outcomes. Section 5 studies the equilibrium under full decentralization. Section 6 contains the normative comparison between partial and full decentralization, comparative statics and robustness results. Section 7 concludes. The main proofs are contained in the Appendix. Additional proofs and derivations are relegated to the Online Appendix.

# 2 Model

## 2.1 Preliminaries

The economy lasts for three periods $t = \{1, 2, 3\}$ and is composed of $L \geq 2$ regions. Each region $\ell \in \{1, ..., L\}$ has a continuum (of measure one) of risk-neutral, immobile residents, each of whom has an endowment $\kappa$ of capital. In the last period, each resident derives utility from consumption of a numéraire good, produced in every region by competitive firms that operate a constant returns technology. Capital is the only input and units are chosen so that one unit of capital produces one unit of output. Following Persson and Tabellini (1992), we assume that capital is mobile between regions, but at a cost. Specifically, a resident of a region that invests $f$ units of capital in other region(s) incurs a mobility cost $f^2/2$. As we explain below, residents may also benefit from a discrete local public good, or project.

There are two levels of government: central and regional. Throughout the paper, we assume that both levels of government are benevolent and choose policies so as to maximize the sum of utilities of their residents. For simplicity, there is no discounting of future payoffs.

## 2.2 Timing

The order of events is as follows. At $t = 1$, a political body (e.g., a Congress) chooses between partial ($PD$) and full decentralization ($FD$). These institutional regimes rule all fiscal interactions between the central and regional governments, in a way specified below.

At the beginning of $t = 2$, Nature draws the cost of the projects according to a strictly positive probability density function $h(c)$ on $[0, b]$. Based on the realized cost $c$, regional governments choose whether or not to initiate a project in their region. This decision is denoted by $i_\ell \in \{I, NI\}$, where $I$ ($NI$) stands for initiation (not initiation). Regional governments have just enough resources to fund the initial investment $c$.

If initiated, a project is carried out by the regional bureaucracy. With an exogenous probability $\pi \in [0, 1]$, a project generates a social benefit $B > 0$ for all residents of the region at the end of the current period.[9] With probability $(1 - \pi)$, the project remains incomplete and yields no benefit during this period.[10] Projects' outcomes are the result of independent draws from the probability distribution $(\pi, 1 - \pi)$, and are observable.

At $t = 3$, central or regional governments, depending on the institutional regime in place, decide whether to shut down or continue incomplete projects. In the last case, a project requires an additional input of $c$ of the consumption good to be completed. It is worth emphasizing that, by construction, $c \leq b$; otherwise, continuation would never be optimal.

Under partial decentralization, the central government decides on refinancing incomplete projects through a uniform tax $\tau$ on capital, collected by the national tax authority. Under full decentralization, each regional government decides whether or not to refinance its incomplete project, using a per unit tax levied on capital invested in its region at the rate $\tau_\ell$.[11]

Once taxes are set, capital owners invest in the region(s) with the highest net return(s), central or regional governments raise their taxes, production takes place, and private consumption (net of mobility costs) occurs.[12] When the project is completed, it generates a social benefit $b$ for all residents of the region. We assume that $b < B$.[13]

In our environment, the federation is characterized by two dimensions of state capacity: *administrative* and *fiscal*.[14] The former measures the ability of regional bureaucracies to carry out projects in due time, and it is encapsulated by the probability $\pi$.[15] The latter

[9]To focus on the trade off between soft and hard budget constraints, we rule out spillovers across regions.

[10]Delays in local public works are prevalent both in developed and developing countries. See Guccio et al. (2014).

[11]Regions do not have access to credit markets to refinance their incomplete projects. But, in any case, if regions could issue debt they would ultimately have to raise taxes to pay back their obligations.

[12]Such a timing is also adopted in the bulk of the literature on tax competition. See Wilson (1999) for a survey.

[13]The difference between $B$ and $b$ reflects that some extra costs arise during the project's delay. For example, an incomplete park may affect pedestrian movement.

[14]Mann (1984) provides a general definition of state capacity as the infrastructural power of the state to enforce policy within its territory. Snyder (2001) and Ziblatt (2008) apply this concept to regional governments.

[15]Patil et al. (2013) document that public projects' delays in Indian states are mainly caused by admin-

gauges the capacity to collect local taxes. This dimension is modeled by assuming that local governments can only collect a fraction $\theta \in [0, 1]$ of their potential tax base, where $\theta$ measures the level of fiscal capacity.[16]

We summarize the timing in Figure 1. Values in terminal nodes represent the benefit of the project.

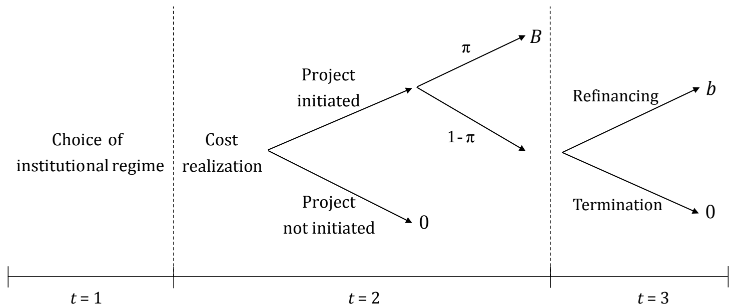


**Figure 1: Timing**

## 2.3 Discussion

Some features of the model deserve further comments. First, our distinction between two dimensions of state capacity (fiscal and administrative) can be motivated both on theoretical and on empirical grounds. Theoretically, our results indicate that the optimal degree of fiscal decentralization crucially hinges on the relative magnitudes of fiscal and administrative capacities prevailing in the federation (see Section 6). Empirically, our data analysis suggests that such dimensions of state capacities are *not* highly correlated, unlike what an aprioristic view could indicate. This point is illustrated in Figure 2. Here we proxy fiscal capacity using the "Political Extraction Index" elaborated by Arbetman-Rabinowitz et al. (2012), and we summarize administrative capacity using the "Government Effectiveness Index" estimated by the World Bank.[17] Each point in the scatter plot represents a country, where each dimension of state capacity is measured as the average of its corresponding index between 2002-2011.[18] As it is evident from the plot, the correlation between each dimension of state ability is positive but particularly low: the correlation coefficient is 0.175.[19] Taken together,

istrative problems that arise during the land acquisition process.

[16]This measure is related to the concept of "tax effort," which is widely used in the empirical public finance literature. See Le et al. (2008). Besley and Persson (2010), on the other hand, define fiscal capacity as an upper bound on the tax rate that a government can levy.

[17]The former index essentially estimates the actual to the potential share of taxes to GDP (see http://thedata.harvard.edu/dvn/dv/rpc). The latter, on the other hand, captures perceptions of the quality of public services and the quality of policy implementation, among other things (see http://info.worldbank.org/governance/wgi/index.aspx#home). Both statistics have already been used to measure fiscal and administrative capacities in previous studies. See Hanson and Sigman (2013).

[18]Our sample contains 170 countries, both from the developed and the developing world, at an annual frequency.

[19]This estimate is significant at a 5% level, with a $p$-value of 0.023.

the previous arguments imply that a thorough analysis of optimal fiscal decentralization should allow for a separation between fiscal and administrative capacities.

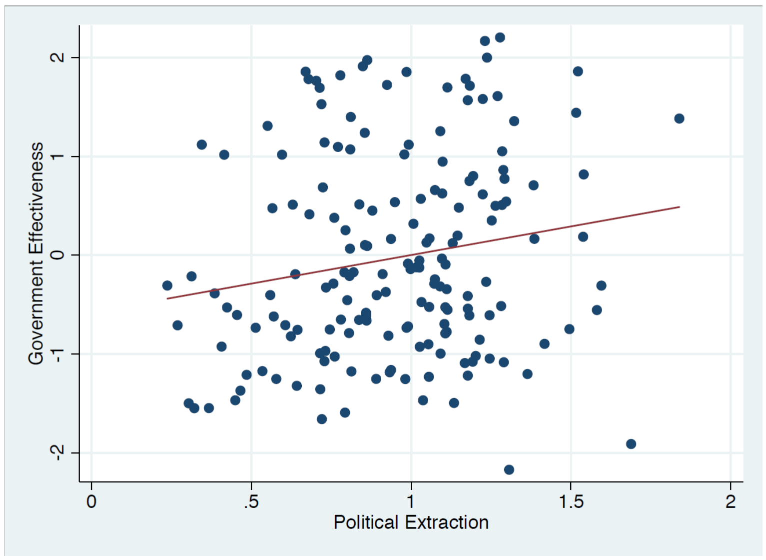


**Figure 2: Fiscal and administrative capacities in the data**

Second, we assume that national taxation under partial decentralization is uniform across regions. This assumption has a counterpart in many federal countries, where national taxes (like the income tax or the VAT) are required to be set uniformly across subnational governments due to constitutional reasons.[20] In our model, it is straightforward to show that if one allowed national taxes to be non-uniform (for example, if they were contingent on which region requires additional funds), the central government would be able to replicate the efficient outcome. But if that was the case, it would certainly be pointless to compare partial and full decentralization, which is one of the main goals of this paper.

Third, we assume that the central government can commit not to bailout regions under full decentralization. If that wasn't the case, we would need to consider a "hybrid" regime, as in Köthenbürger (2004), where the central government keeps the option to bailout regions that have decided to refinance a fraction of their incomplete project. However, it can be shown that the equilibrium of such a "hybrid" regime replicates the partial decentralization outcome. Indeed, anticipating a bailout from the central government (which is, by assumption, ex post optimal) and in order to avoid the use of distortionary taxation, each region would decide not to refinance its incomplete project. In the end, the central government would bailout all regions.

Fourth, we do not incorporate considerations of administrative capacity at the central level. The reason is that, in order to focus on bailouts under partial decentralization, we

[20] Among developed countries, US, Australia and Switzerland incorporate uniformity of national taxation explicitly in their constitutions. Analogous examples among developing countries include Argentina, Brazil and South Africa.

suppose that the central government cannot intervene a region to avoid its project's delay. On the other hand, while in the base model the central government is assumed to be fully efficient along its fiscal ability, this assumption is relaxed in Section 6.4.

Fifth, we consider a discrete, regional public project, instead of a continuous public good as it is common in most of the literature on tax competition.[21] Indivisibility fixes the type of competition between regions. As in Wildasin (1988), regions compete first in refinancing decisions, and then taxes are set accordingly in a context of imperfect capital mobility.[22] Moreover, this assumption combined with our specification of administrative competence is a simple way to analyze, via refinancing decisions, the interaction between levels of regional state capacity and different intergovernmental fiscal arrangements.

Finally, regional governments and projects in the model are *ex ante* and *interim* identical. More precisely, all regional governments share the same exogenous levels of state capacity $\pi$ and $\theta$. Concerning projects, ex ante (i.e., in period 1) they are all characterized by the same configuration of exogenous social benefits $b$ and $B$, and by the same probability density function $h(c)$. Interim (i.e., at the beginning of period 2), the cost $c$ is realized and applies for all projects in all regions. However, we still introduce *ex post* heterogeneity given that projects' outcomes can be different across regions.[23]

# 3 First Best

In this section we analyze a benchmark for efficiency. Consider a social planner who makes all decisions, but cannot anticipate whether a project will be completed at the end of $t = 2$ (i.e., he has to carry out projects through the regional bureaucracies). We solve his decision problem backwards.

First, the refinancing decision in any region is independent of the planner's choice in any other region. This is because individual utilities are linear in income and the planner maximizes the sum of utilities. Thus, in any region, continuing an incomplete project is always optimal because $c \leq b$.

Moving back to the initial investment decision, the planner faces another separable problem between regions. Knowing the cost $c$, he initiates projects provided their expected benefit is higher than their expected cost (which includes a possible second round of financing). Let

$$c^*(\pi) \equiv \frac{\pi B + (1-\pi)b}{2-\pi}$$

denote the cost that makes the net expected regional welfare from initiating a project equal to zero. If $c \leq c^*(\pi)$, initial investment is efficient in any region $\ell$; otherwise, not investing is the optimal decision.

[21] To the best of our knowledge, Cremer et al. (1997) and Lockwood (2002) are the other contributions to the local public finance literature which deal with discrete projects.

[22] Akai and Sato (2008) and Köthenbürger (2011) also analyze models with this timing, but they consider income or wage taxation instead.

[23] An important feature of the model is risk neutrality. If individuals were risk averse, this could imply an insurance motive for bailouts under partial decentralization.

For a given configuration of parameters $(b, B)$, efficient investment decisions are depicted in Figure 3. A point in the $(\pi, c)$ plane represents a project that costs $c$ in a region with administrative capacity $\pi$.

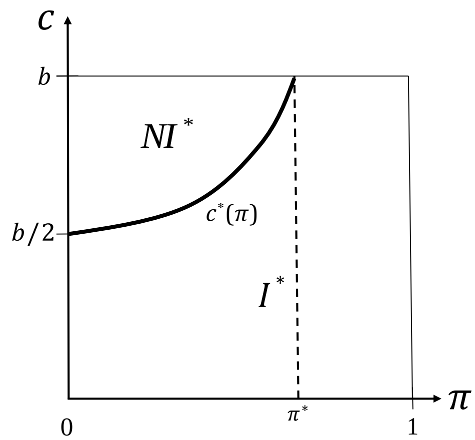


**Figure 3: First best outcomes**

When $0 \leq \pi \leq \pi^* \equiv b/B$, we have that $0 < c^*(\pi) \leq b$. Therefore, there exists a non-empty area $NI^*$, delimited by the thick curve $c^*(\pi)$, where it is optimal not to initiate projects. Below this curve, in the area $I^*$, it is efficient to undertake all projects. Ceteris paribus, as $\pi$ increases the projects' net expected benefit increases. Thus $c^*(\pi)$ increases as well. When $\pi^* < \pi \leq 1$, $c^*(\pi) > b$: all projects are efficiently initiated.

# 4 Partial Decentralization

This section studies the partially decentralized regime. Given that $c \leq b$, under this regime incomplete projects are always refinanced by the central government using a uniform tax on capital. This implies that project initiation decisions of any given region may ultimately impact welfare of others, thus giving rise to a simultaneous game between regions in the second period (i.e., when the initial investment decision is made).

Region $\ell$'s expected welfare at the beginning of the second period is given by

$$\mathbb{E}W_\ell^{PD}(i_\ell, \mathbf{i}_{\boldsymbol{m}}) = \kappa(1 - \tau^e) + \mathbb{1}_{\{i_\ell = I\}}[\pi B + (1 - \pi)b - c], \tag{1}$$

where $\mathbf{i}_{\boldsymbol{m}}$ is the profile of investment decisions chosen by regions $m \neq \ell$, $\mathbb{1}_{\{i_\ell = I\}}$ is equal to 1 if region $\ell$ has initiated the project and to 0 otherwise, and $\tau^e$ is the expected tax.

The expected tax $\tau^e$ is obtained as follows. At the end of the second period, all projects' outcomes are realized. Let $\omega$ be a profile of outcomes, and denote by $\mathcal{B}(\omega)$ the number of completed projects in this particular realization of outcomes. For any profile $\omega$, the central government mechanically sets a tax $\tau_\omega$ to cover, at the beginning of the third period, the cost of refinancing $\sum_\ell \mathbb{1}_{\{i_\ell = I\}} - \mathcal{B}(\omega)$ incomplete projects. As this tax is uniform and exporting capital is costly, every household will invest in its own region. This implies that the tax base is $L\kappa$, and taxation is non distortionary. Hence, under the profile $\omega$, the central government's budget constraint is

$$\tau_\omega . L\kappa = \left[ \sum_\ell \mathbb{I}_{\{i_\ell = I\}} - \mathcal{B}(\omega) \right] c.$$

Therefore, when deciding on initial investment before the outcome of projects are realized, each region faces the expected tax $\tau^e \equiv \mathbb{E}_\omega \tau_\omega$, which satisfies

$$\tau^e . L\kappa = \left[ \sum_\ell \mathbb{I}_{\{i_\ell = I\}} (1-\pi) \right] c, \tag{2}$$

where the term in square brackets gives the expected number of bailouts. Substituting (2) into (1) and rearranging, we obtain

$$\mathbb{E}W_\ell^{PD}(i_\ell, \mathbf{i}_{\boldsymbol{m}}) = \kappa + \mathbb{I}_{\{i_\ell = I\}} \left[ \pi B + (1-\pi) \left( b - \frac{c}{L} \right) - c \right] - \sum_{m \neq \ell} \mathbb{I}_{\{i_m = I\}} (1-\pi) \frac{c}{L}. \tag{3}$$

By inspection of (3), the effect of $i_\ell$ on $\mathbb{E}W_\ell^{PD}$ (captured by the term in square brackets) is independent of $\mathbf{i}_{\boldsymbol{m}}$. So, we can analyze the choice of $i_\ell$ just for a representative region $\ell$.

Notice that each region only pays $1/L$ of the cost of refinancing its incomplete project, as this cost is shared through national taxation. Therefore, the central government's budget constraint generates a *common-pool fiscal externality*: any resident of $\ell$ is negatively affected by the possibility of an incomplete project in a region $m \neq \ell$. Let

$$c^{PD}(\pi) \equiv \frac{L[\pi B + (1-\pi) b]}{L + 1 - \pi}$$

be the cost that makes the net expected regional welfare from initiating the project under partial decentralization equal to zero. The next proposition completely characterizes regional project initiation decisions under this institutional regime.

**Proposition 1** *Consider the project initiation game under partial decentralization. Symmetric equilibria are as follows. If $c \leq c^{PD}(\pi)$, initial investment takes place in all regions. Otherwise, no region invests in equilibrium.*

In the Appendix we show that $c^{PD}(\pi) > c^*(\pi)$. Hence, we can establish the form of the inefficiencies that emerge under this institutional regime as follows.

**Corollary 1** *Under partial decentralization, initial investments may occur in equilibrium when it is inefficient to do so.*

Inefficiencies involve *over investment*. This kind of inefficiency, driven by the common-pool fiscal externality, is well known (See Wildasin (1997) and Goodspeed (2002)). The following figure depicts equilibrium outcomes that emerge under partial decentralization.

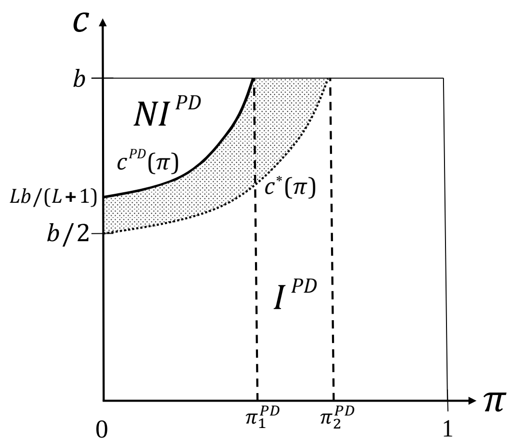


**Figure 4: Partial decentralization outcomes**

Analogous to the first best, when $0 \leq \pi < \pi_1^{PD} \equiv \frac{b}{L(B-b)+b}$, there exists a non-empty area ($NI^{PD}$) where projects are efficiently not initiated in any region, and another one ($I^{PD}$) where all projects are initiated. In the white area, regional decisions are optimal and so, in equilibrium, each region is expected to contribute an amount equal to the cost of refinancing by itself its incomplete project. Thus, interim expected welfare in each region coincides with the first best level. In the shaded area, when $c \in [c^*(\pi), c^{PD}(\pi)]$, the inefficient investment decision is adopted in equilibrium by all regions. When $\pi_1^{PD} \leq \pi \leq \pi_2^{PD} \equiv b/B$, $c^*(\pi) \leq b \leq c^{PD}(\pi)$, which implies that all projects are initiated in equilibrium. Projects in the shaded area are inefficiently initiated. Finally, when $\pi_2^{PD} < \pi \leq 1$, inefficient investments cannot emerge because the model is biased towards project initiation. Under these parameter conditions, partial decentralization replicates the first best outcome.

# 5 Full Decentralization

In this section we analyze the fully decentralized regime. In this case, a three-stage simultaneous game between regions emerges. First, regional governments take the initial investment decision. Second, the continuation decision is made. Finally, refinancing is achieved by levying taxes on capital employed in the region, in a context of tax competition.

But before solving the game backwards, it is convenient to describe how capital reacts to different tax rates. Given a profile of tax rates $\tau = \{\tau_1, ..., \tau_L\}$ set by all regions, a household resident in region $\ell$ decides where to invest its capital endowment. Let $f_{\ell m}$ denote the amount of capital that this household invests in a region $m \neq \ell$, and let $\widetilde{M}$ be the set of regions $\widetilde{m} \neq \ell$ that have chosen the minimum tax rate $\widetilde{\tau}_\ell = \min\{\tau_m\}_{m \neq \ell}$. The following proposition characterizes the household's investment decision.

**Proposition 2** *If* $\tau_\ell \geq \widetilde{\tau}_\ell$, $\sum_{\widetilde{m} \in \widetilde{M}} f_{\ell \widetilde{m}} = \tau_\ell - \widetilde{\tau}_\ell \geq 0$. *Otherwise,* $f_{\ell \widetilde{m}} = 0$.

As a household in region $\ell$ seeks to maximize net returns from its investments, its portfolio decision only depends upon the comparison between $\tau_\ell$ and $\widetilde{\tau}_\ell$, and not between $\tau_\ell$ and the whole profile of tax rates chosen by regions $m \neq \ell$. As expected, a household in region $\ell$ invests "abroad" in a region $\widetilde{m} \in \widetilde{M}$ provided $\tau_\ell > \widetilde{\tau}_\ell$. When only one region sets the minimum tax rate $\widetilde{\tau}_\ell$, it attracts all capital that leaves region $\ell$. But if two or more regions choose the same tax rate $\widetilde{\tau}_\ell$, the value of the capital outflow $f_{\ell\widetilde{m}}$ that goes to each of these regions is undetermined. The intuition behind the expression in the proposition is straightforward, as capital leaves region $\ell$ until the marginal tax savings equal the marginal mobility cost. As expected, this flow increases with the difference $\tau_\ell - \widetilde{\tau}_\ell$.

When the regional tax rate $\tau_\ell$ is lower or equal than $\widetilde{\tau}_\ell$, there is no capital outflow from region $\ell$ and its residents invest all their endowment "at home." Moreover, in this case, region $\ell$ receives capital inflows from other regions. But this does not benefit directly its residents because returns from these investments are consumed abroad, by residents in regions $m \neq \ell, \widetilde{m}$. Despite this fact, in the next section we show that these capital inflows have an important role in the determination of the equilibrium tax rates.

## 5.1 Equilibrium in Tax Rates

In the remainder of the section, we solve the game between regions backwards. We start with the last stage, where regional governments set tax rates to refinance incomplete projects.

To obtain the equilibrium tax rates, in the Appendix we derive region $\ell$'s reaction function $\tau_\ell(\boldsymbol{\tau_m})$, where $\boldsymbol{\tau_m}$ denotes the profile of tax rates chosen by regions $m \neq \ell$. We have previously discussed that, depending upon the whole profile of tax rates, capital may leave or enter region $\ell$ from other regions. Despite these different possibilities, region $\ell$'s after-tax consumption monotonically decreases with $\tau_\ell$, and so does regional welfare. Hence, for any profile $\boldsymbol{\tau_m}$, the tax rate chosen in region $\ell$ should be the lowest tax rate that enables the regional government to raise $c$.

The reaction function $\tau_\ell(\boldsymbol{\tau_m})$ is built around the value $c/\theta\kappa$, which is the tax rate that a regional government would choose in autarky. When the profile $\boldsymbol{\tau_m}$ is such that $\tau_m < c/\theta\kappa$ for all $m$, region $\ell$'s optimal response is to tax strictly above the minimum tax $\widetilde{\tau}_\ell$. Despite the fact that this decision will trigger a capital outflow, this is the only way to ensure the project's refinancing. When all tax rates $\tau_m$ are set equal to $c/\theta\kappa$, region $\ell$'s optimal response is to replicate this level. Due to the way we model imperfect capital mobility, the tax collection's elasticity with respect to $\tau_\ell$ is lower than one. This combined with the fact that region $\ell$ needs to collect enough revenues to refinance its incomplete project, makes tax undercutting not a profitable deviation. Finally, when the profile of tax rates is such that $\widetilde{\tau}_\ell > c/\theta\kappa$ or $\widetilde{\tau}_\ell = c/\theta\kappa$ but at least one region $n \neq \ell$ has chosen a tax rate $\tau_n > c/\theta\kappa$, region $\ell$'s optimal response is to tax strictly below $\widetilde{\tau}_\ell$. This decision generates an inflow of capital that allows the government to raise sufficient revenues to pursue its incomplete project, thus moderating the tax burden on its residents.

Another important feature of region $\ell$'s reaction function is that it is non-continuous. Despite this fact, we can still characterize the Nash equilibria of this subgame as follows.

**Proposition 3** *When all regions have decided to refinance their incomplete project, the unique symmetric Nash equilibrium in pure strategies is such that $\widehat{\tau}_\ell = c/\theta\kappa$. When there is at least one region that does not refinance, then all regions that refinance set*

$$\widehat{\tau}_\ell(0) \equiv \frac{1}{2}\left[\kappa - \sqrt{\kappa^2 - (4c/\theta)}\right].$$

Consider the tax competition subgame that emerges when all regions have decided to refinance their incomplete project. The equilibrium tax rate $\widehat{\tau}_\ell$ increases with the cost of the project $c$. Moreover, as in equilibrium nobody invests abroad, regions tax their own capital endowment without bearing any deadweight loss due to its mobility. Thus, the higher this endowment, the lower the equilibrium tax rate. Similarly, the higher the fiscal capacity $\theta$, the lower the equilibrium tax $\widehat{\tau}_\ell$. In this case, region $\ell$'s welfare (net of the initial cost $c$) is $W_\ell^{FD} = \kappa + b - \frac{c}{\theta}$ in equilibrium. When $\theta < 1$, imperfect fiscal capacity implies that regions do not pay only the technical cost of completing the project $c$, but a higher, effective refinancing cost $c/\theta$. The difference between these values corresponds to the cost of tax collection.

The proposition also shows that, when at least one region does not refinance (in which case it does not need to tax its population), region $\ell$ has to set the tax rate $\widehat{\tau}_\ell(0)$ to pursue its ongoing project. Thus, asymmetric taxation emerges as a possible equilibrium, as in Bucovetsky (1991) and Wilson (1991).[24] The tax rate $\widehat{\tau}_\ell(0)$ also decreases with the capital endowment $\kappa$ and the fiscal capacity $\theta$. With this tax rate, the resulting capital outflow is $\sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}} = \widehat{\tau}_\ell(0)$ and the equilibrium regional welfare (net of the initial cost $c$) is

$$\begin{aligned} W_\ell^{FD} &= \left(\kappa - \textstyle\sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}}\right)(1 - \widehat{\tau}_\ell(0)) + \textstyle\sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}} - \frac{1}{2}\left(\textstyle\sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}}\right)^2 + b \\ &= \kappa + b - T(c,\theta), \end{aligned}$$

where

$$T(c,\theta) = \frac{c}{\theta} + \frac{[\widehat{\tau}_\ell(0)]^2}{2}$$

measures the total refinancing cost, which comprises the effective refinancing cost $c/\theta$, plus the deadweight loss $[\widehat{\tau}_\ell(0)]^2/2$ of financing the continuation of the project through a distortionary tax. This distortion is due to mobility costs incurred by owners of capital seeking to avoid taxation in region $\ell$. Therefore, distortionary taxation only emerges in some final nodes of the tax competition subgame, when at least one region does not refinance. More importantly, the likelihood of these final nodes depends upon the level of regional administrative capacity $\pi$.

[24] The main difference with their result is that variations in tax rates across regions are not originated from an ex ante regional asymmetry, but rather from the possibility that some regions may end up with incomplete projects ex post.

## 5.2 Refinancing

At the beginning of the second stage, the strategy for region $\ell$ is $r_\ell \in \{R, NR\}$, where $R$ ($NR$) denotes "refinancing" ("not refinancing"). Conditional on $\mathbf{i} = (i_1, ..., i_L)$, and given $\mathbf{r}_{\boldsymbol{m}}$, the profile of refinancing decisions chosen by regions $m \neq \ell$, region $\ell$'s welfare is

$$W_\ell^{FD}(r_\ell = R, \mathbf{r}_{\boldsymbol{m}}, \mathbf{i}) = \kappa + \mathbb{1}^\ell_{\{I,inc\}}\{[b - \mathbb{1}^m_{\{I,inc,R\}}\frac{c}{\theta} - (1 - \mathbb{1}^m_{\{I,inc,R\}})T(c,\theta)] - c\},$$

and

$$W_\ell^{FD}(r_\ell = NR, \mathbf{r}_{\boldsymbol{m}}, \mathbf{i}) = \kappa - \mathbb{1}^\ell_{\{I,inc\}}c$$

where $\mathbb{1}^\ell_{\{I,inc\}}$ is equal to 1 if region $\ell$ initiated a project that has remained incomplete at $t = 2$, and to 0 otherwise, and $\mathbb{1}^m_{\{I,inc,R\}}$ is equal to 1 if all regions $m \neq \ell$ initiated their project, have not completed it in due time but decided to refinance them in $t = 3$, and to 0 otherwise.

The following proposition characterizes the Nash equilibria of this subgame.

**Proposition 4** *Let $c_1$ denote the value of $c$ that makes the total refinancing cost $T(c,\theta)$ equal to the benefit $b$*, and $c_2 > c_1$ *the value of $c$ that makes the effective refinancing cost $c/\theta$ equal to the benefit $b$. When all regions face an incomplete project, they refinance them in equilibrium provided $c \leq c_2$. Otherwise, no region refinances. When there is at least one region that does not need refinancing, region $\ell$ refinances its incomplete project provided $c \leq c_1$.*

In the Appendix we show that $c \leq c_1$ ($c \leq c_2$) implies $T(c,\theta) \leq b$ ($c/\theta \leq b$). When all regions failed to complete their project in due time, the Nash equilibria of this refinancing subgame depend upon the cost $c$. When $c < c_1$, refinancing is a dominant strategy. But when $c_1 \leq c \leq c_2$, a standard coordination game emerges, with two Nash equilibria. The first one is trivial: despite the fact that $c \geq c_1$, when all regions decide to refinance, these strategies form a Nash equilibrium because, as there will be no capital flows (and thus no deadweight loss due to distortionary taxation), only the fact that $c/\theta \leq b$ matters. But if at least one region has decided not to refinance, the other regions also shut down their project because $c \geq c_1$. In this case, these regions face an endogenous hard budget constraint due to other regions' decisions, as in Qian and Roland (1998). We focus on the Nash equilibrium in which all regions refinance because it is strong (Aumann (1959)) and coalition-proof (Bernheim et al. (1987)). Finally, when $c_2 \leq c$, not refinancing is a dominant strategy. Regions face again an endogenous hard budget constraint, but this time as a consequence of their imperfect regional fiscal capacity. As $c \leq b$, this outcome is inefficient.

In all other subgames, when there is at least one region that does not need refinancing because it executed its project in due time, region $\ell$ refinances its incomplete project provided $c \leq c_1$. If $c > c_1$, the total cost from refinancing is higher than the benefit $b$, pushing region $\ell$ to shutdown its incomplete project. Once more, this is an inefficient outcome.

## 5.3 Project Initiation

Anticipating refinancing equilibria, regional governments simultaneously adopt the project initiation decision. The next proposition characterizes the Nash equilibria of the investment stage under this regime.

**Proposition 5** *Consider the project initiation game under full decentralization. Let* $c_R^{FD}(\pi)$, $c_{NAR}^{FD}(\pi)$ *and* $c_{NR}^{FD}(\pi)$ *be the costs that make the net expected regional welfare equal to zero (i) when incomplete projects are refinanced regardless of the continuation decisions of other regions, (ii) when incomplete projects are refinanced only if all other regions do so, and (iii) when incomplete projects are never refinanced, respectively. There exists thresholds* $\pi_1^{FD}, \pi_2^{FD}$ *and* $\pi_3^{FD}$ *such that the symmetric Nash equilibria are as follows:*

1. *When* $0 \leq \pi \leq \pi_1^{FD}$, *initial investment takes place in all regions provided* $c \leq c_R^{FD}(\pi)$. *Otherwise, no region invests in equilibrium.*

2. *When* $\pi_1^{FD} < \pi \leq \pi_2^{FD}$, *initial investment takes place in all regions provided* $c \leq c_{NAR}^{FD}(\pi)$. *Otherwise, no region invests in equilibrium.*

3. *When* $\pi_2^{FD} < \pi \leq \pi_3^{FD}$, *initial investment takes place in all regions provided* $c \leq c_{NR}^{FD}(\pi)$. *Otherwise, no region invests in equilibrium.*

4. *When* $\pi_3^{FD} < \pi \leq 1$, *initial investment takes place in all regions.*

In the Appendix we characterize the probability thresholds $\pi_1^{FD}, \pi_2^{FD}$ and $\pi_3^{FD}$, and we show that cost thresholds $c_R^{FD}(\pi), c_{NAR}^{FD}(\pi)$ and $c_{NR}^{FD}(\pi)$ are lower than $c^*(\pi)$. Hence, we can establish the types of inefficiencies that emerge under this institutional regime, as follows.

**Corollary 2** *Under full decentralization, equilibrium outcomes can be inefficient for three reasons: (i) initial investments do not take place in equilibrium when it is efficient to do so, (ii) initial investments take place in equilibrium and are efficient, but incomplete projects are not refinanced when it is efficient to do so, or (iii) initial investments take place in equilibrium and are efficient, but incomplete projects are refinanced using distortionary capital taxes.*

The following figure depicts equilibrium outcomes that emerge under full decentralization.

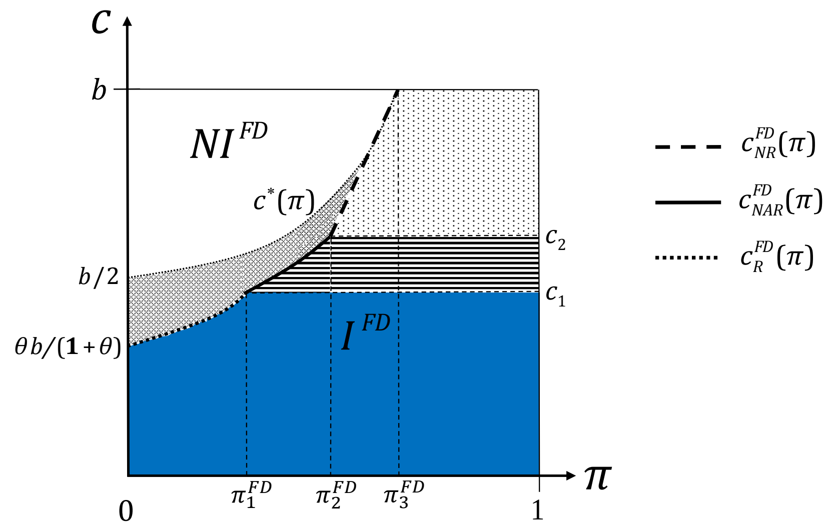


**Figure 5: Full decentralization outcomes**

In the non-empty area $NI^{FD}$, delimited from below by the thick curves representing the cost thresholds $c_R^{FD}(\pi), c_{NAR}^{FD}(\pi)$ and $c_{NR}^{FD}(\pi)$, projects are not initiated. In the complementary area, denoted by $I^{FD}$, all projects are initiated.

When $c \in [c^*(\pi), b]$, efficient decisions are adopted. In all other areas, either the projects' initiation and continuation decisions are distorted or the region refinances bearing deadweight losses. When $0 \leq \pi \leq \pi_1^{FD}$, two different types of inefficiency emerge. First, in the darker area when $c \in [0, c_R^{FD}(\pi)]$, initiation and continuation decisions are optimal, but refinancing is done bearing deadweight losses generated by imperfect regional fiscal capacity or distortionary capital taxation. Second, as $c_R^{FD}(\pi) < c^*(\pi)$, the condition for project initiation is stricter with full decentralization than for the social planner. Therefore, in the shaded area when $c \in [c_R^{FD}(\pi), c^*(\pi)]$, investments are not initiated in equilibrium, despite the fact that they are efficient. Underinvestment is due to i) distortionary refinancing, when $c \in [c_R^{FD}(\pi), c_1]$, and ii) endogenous hard budget constraints, when $c \in [c_1, c^*(\pi)]$.[25]

When $\pi_1^{FD} < \pi \leq \pi_2^{FD}$, a new type of inefficiency emerges. In the dashed area, projects are initiated but only refinanced when all regions do so. Again, there is underinvestment when $c \in [c_{NAR}^{FD}(\pi), c^*(\pi)]$, but only due to endogenous hard budget constraints.

As $\pi_2^{FD} < \pi \leq \pi_3^{FD}$, the last type of inefficiency emerges. In the dotted area projects are initiated but never refinanced. There is still a shaded area where, because of endogenous hard budget constraints, projects are not initiated.[26]

[25] The former is the analogue of the Zodrow and Mieszkowski's (1986) result, while the latter has been analyzed by Besfamille and Lockwood (2008), in a setting where an exogenous hard budget constraint is imposed to regional governments.

[26] Figure 5 depicts full decentralization outcomes when $\theta < 1$. If $\theta = 1$, $c_2 = b$ and $\pi_2^{FD} = \pi_3^{FD}$: the area where projects are never refinanced vanishes.

Finally, when $\pi \geq \pi_3^{FD}$, the model is biased towards project initiation. Incomplete projects are either refinanced in a distortionary way, shut down in some terminal nodes of the tax competition subgame, or never finished.

# 6 Optimal Institutional Regime

In the initial period, there is an institutional choice between partial and full decentralization. At this stage, the Congress observes projects' payoffs $b, B$, the regional capital endowment $\kappa$ and state capacities $(\pi, \theta)$, and knows that the cost $c$ is distributed according to a strictly positive probability density function $h(c)$ on $[0, b]$. The Congress chooses the optimal regime by maximizing

$$\mathbb{E}W^{IR} = \int_0^b \mathbb{E}W_\ell^{IR} h(c) dc,$$

where $\mathbb{E}W^{IR}$ is the expected welfare of a representative region, under the institutional regime $IR \in \{PD, FD\}$.[27]

The main goal of this paper is to evaluate how the regime choice is affected by the level of regional state capacity. We first characterize the relationship between $\mathbb{E}W^{IR}$ and the pair $(\pi, \theta)$. Taking into account equilibrium decisions and outcomes, we show that $\mathbb{E}W^{FD}$ is a continuous, everywhere differentiable, increasing, convex function of the administrative capacity $\pi$. Moreover, it also increases with the fiscal capacity $\theta$. Therefore, under full decentralization, an increase in either $\pi$ or $\theta$ increases $\mathbb{E}W^{FD}$, as suggested by the decentralization literature.[28] But this assertion does not necessarily imply that full decentralization dominates for high levels of state capacity. The reason is that $\mathbb{E}W^{PD}$ is a continuous, increasing, convex function of the administrative capacity $\pi$.[29] Hence, the comparison between both regimes is not a priori evident.

To make progress, we introduce the following assumption which ensures that the intersection between $\mathbb{E}W^{FD}$ and $\mathbb{E}W^{PD}$ is unique.

**Assumption A**

1. The cost $c$ is distributed uniformly on $[0, b]$.
2. The benefit $b$ satisfies $B/2 \leq b < B$.
3. The number of regions $L$ satisfies $\underline{L} \leq L \leq \overline{L}$, where $\underline{L}$ and $\overline{L}$ are defined in the Appendix.

[27] In a one-shot version of the model, the Congress could wait until the realization of the cost to choose the optimal interim regime. But in a repeated version of the model, it seems realistic to assume that changing the institutional regime after each realization of $c$ would be too costly, which justifies our focus on the optimal ex-ante regime.

[28] See, among others, Bird (1995), Litvack et al. (1998) and Loayza et al. (2014).

[29] See the Online Appendix.

The following proposition characterizes the optimal choice between partial and full decentralization.

**Proposition 6** *Suppose Assumption A holds and let* $\theta_0 \equiv 2L/(1+L^2)$ *be the regional fiscal capacity that makes* $\mathbb{E}W^{PD}$ *equal to* $\mathbb{E}W^{FD}$ *when* $\pi = 0$. *Then,*

1. *When* $\theta < \theta_0$, *partial decentralization dominates for all values of* $\pi$.

2. *When* $\theta \geq \theta_0$, *there exists a unique threshold* $\widehat{\pi}(\theta)$ *such that, when* $\pi \leq \widehat{\pi}(\theta)$, *full decentralization dominates. Otherwise, partial decentralization dominates.*

3. *When* $\pi = 1$, *both regimes are efficient.*

Figure 6 illustrates this result. Each point in the $(\theta, \pi)$ plane represents the regional state capacity that prevails in the federation. In the Appendix, we show that $\hat{\pi}(\theta)$ increases with $\theta$.[30]

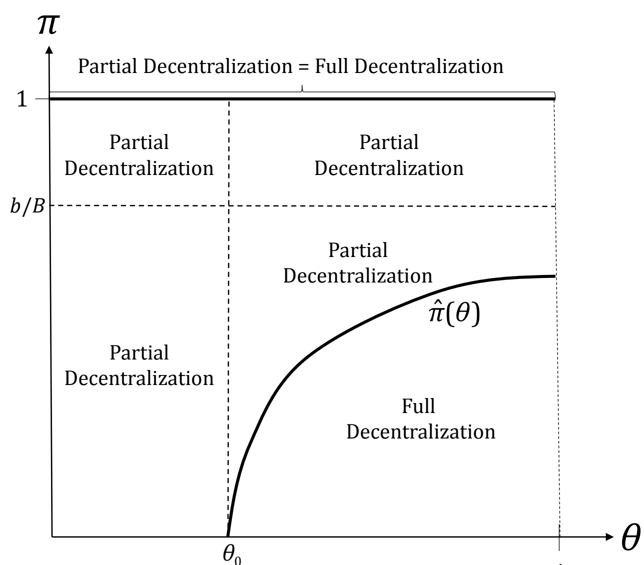


**Figure 6: Partial vs. full decentralization**

When regional fiscal capacity satisfies $\theta < \theta_0$, refinancing incomplete projects under full decentralization is too costly. Therefore, regardless of the level of regional administrative capacity, partial decentralization dominates.

As expected, full decentralization may dominate when regional fiscal capacity is relatively high. But interestingly, this occurs for relatively low levels of the regional administrative

[30]When Assumption A does not hold, the model becomes analytically untractable and thus unicity of $\hat{\pi}(\theta)$ cannot be ensured. In order to verify whether our results would hold under parameter configurations of the model that do not satisfy those conditions, we have simulated the model and replicated Figure 6. All simulations confirmed the results presented in Proposition 6, and are available upon request.

capacity, i.e., when $\pi \leq \hat{\pi}(\theta)$. The intuition for this result is the following. Under full decentralization, tax distortions emerge if at least one region does not need refinancing. But when $\pi$ is low, the likelihood of this event is also low. Thus, distortions under full decentralization cost less than inefficient investments made under partial decentralization, where investments need to be refinanced with a relatively high probability. On the other hand, when $\pi$ is relatively high, the expected welfare cost of inefficiently initiated projects under partial decentralization is low (or even zero) because it is very likely that they will be completed at the end of $t = 2$.

When $\pi$ is above $b/B$ the model is biased towards project initiation, and under partial decentralization outcomes are efficient. On the other hand, under full decentralization, the likelihood of facing capital mobility costs or shutting down the project is still positive. Thus, partial decentralization dominates. But if $\pi$ increases further and converges to one, the likelihood and welfare cost of full decentralization's distortions decrease, attenuating partial decentralization's dominance. In the limit, when $\pi$ is equal to one, both regimes yield optimal outcomes.

These results clarify how the level of regional state capacity prevalent in a federation affects the trade-off between partial and full decentralization. On the one hand, our results confirm that a high level of fiscal capacity is a necessary condition for full decentralization to be the optimal institutional regime. But, we also caution against the position held by some authors suggesting that high levels of regional state capacity are necessary for successful decentralization reforms. First, our model shows that high levels of regional administrative or fiscal capacity do not always imply that full decentralization should dominate. Second, it is still plausible that full decentralization dominates even for low levels of administrative capacity.

## 6.1 Comparative Statics

In this subsection we analyze how changes in the key parameters of the model affect the comparison between partial and full decentralization.[31]

[31]Proofs corresponding to this section are contained in the Online Appendix.

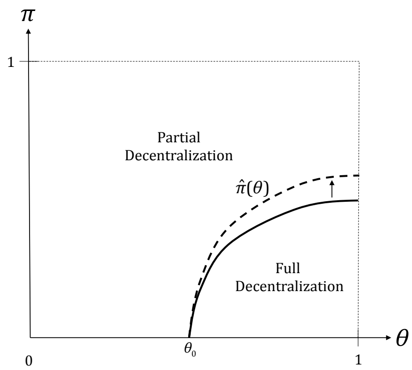

Figure 7a: Increase in capital endowment $\kappa$

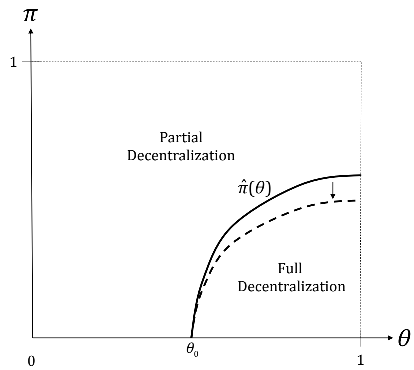

Figure 7b: Increase in project's benefit $B$

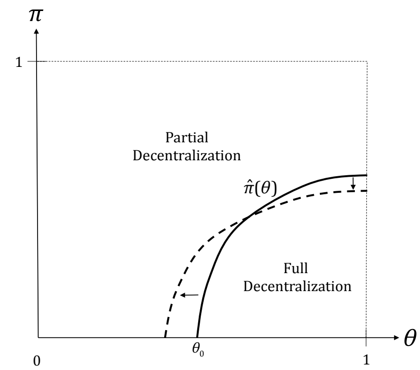

Figure 7c: Increase in the number of regions $L$

**Change in the capital endowment** $\kappa$. Figure 7a shows that an increase in $\kappa$ favors full decentralization, in the sense that this regime dominates in a larger area of the plane $(\theta, \pi)$. The intuition of this result hinges on two facts. First, the cost of distortions generated under partial decentralization do not depend upon the value of $\kappa$. Second, an increase in the regional stock of capital increases expected welfare under full decentralization, because as the tax $\widehat{\tau}_\ell(0)$ decreases with $\kappa$, and so does the mobility cost of capital.

**Change in the project's benefit** $B$. Figure 7b shows that an increase in $B$ always favors partial decentralization. The reason is that when $B$ increases, the welfare loss due to overinvestment under partial decentralization increases less than the corresponding welfare loss due to underinvestment under full decentralization.

**Change in the number of regions** $L$. Figure 7c shows that an increase in $L$ has two effects. On the one hand, for relatively high levels of the regional fiscal capacity, partial decentralization dominates in an area where full decentralization used to be the optimal regime, and vice versa for low levels of regional administrative capacity. The intuition for this result hinges on the relative costs of an increase in $L$. On the one hand, the common-pool fiscal externality under partial decentralization gets larger, and thus more projects are inefficiently initiated. On the other hand, under full decentralization, the likelihood that a given region has to bear deadweight losses due to capital mobility or to its project's shutdown increases as well. When the administrative capacity $\pi$ is relatively low, the cost of the latter is lower; the opposite holds for relatively high values of $\pi$.

## 6.2 A Special Case: $\theta = \pi$

In this section we analyze the special case in which there is only one dimension of state capacity, say $\sigma$, with $\sigma \equiv \theta = \pi$. Using a simple geometrical argument, we can show the following result:

**Corollary 3** *Suppose $\theta = \pi$ and let $\sigma \equiv \theta = \pi$. Then two cases can emerge:*

1. *There exists two thresholds, $\sigma_1$ and $\sigma_2$, such that full decentralization dominates when $\sigma \in [\sigma_1, \sigma_2]$. Otherwise, partial decentralization dominates.*

2. *Partial decentralization always dominates.*

The proof follows by simply inspecting Figure 6. More precisely, case 1 in the corollary emerges when the curve $\hat{\pi}(\theta)$ crosses the 45° line. Case 2, on the other hand, is obtained whenever the intersection between those two curves is empty.[32] The intuition behind the first case is the following. When $\sigma$ is low, fiscal capacity is low and thus tax distortions under full decentralization generate more costs than those arising from bailing out regions under partial decentralization. Conversely, if $\sigma$ is high the likelihood of distortionary tax competition under full decentralization is also high, whereas bailouts under partial decentralization are unusual. It then follows that full decentralization can only dominate when $\sigma$ has an intermediate value: in that case, the likelihood of recurring to distortionary tax competition under full decentralization is sufficiently low, while the inefficiency cost of project overprovision under partial decentralization is sufficiently high. Case 2 in Corollary 3 corresponds to particular case in which the area $[\sigma_1, \sigma_2]$ vanishes.

These results confirm that assuming a unique dimension of state capacity is not only at odds with the empirical evidence presented previously (see Figure 2), but it is also with loss of generality. Indeed, in the unidimensional formulation, full decentralization dominates for intermediate levels of state capacity, while in the general case full decentralization dominates in the corner of the domain.

## 6.3 An Extension: Distortionary National Taxation

So far, the tax set by the central government to bail out regions under partial decentralization was non distortionary. This is arguably a strong assumption: typically national taxes also generate deadweight losses, so assuming non-distortionary national taxation underestimates the cost of the common-pool fiscal externality. To address this issue, in this section we model distortions in national taxation by assuming that the cost of bailing out an incomplete project under partial decentralization is $c + \lambda$, where $0 < \lambda < b$.[33] Although this change does not affect the fully decentralized regime, outcomes in the other regime are modified as shown by the following figure.[34]

[32] If the intersection is non-empty, $\hat{\pi}(\theta)$ crosses the 45° line twice since $\theta_0 > 0$ and $\hat{\pi}(1) < 1$.

[33] This assumption is a shortcut that captures, among other things, cost differences in complying with central and regional tax systems. As stressed by Slemrod and Venkatesh (2002), these differences are substantial: nearly 70% of their compliance spending was devoted by firms to federal government's compliance, whereas almost 25% was spent on regional and local compliance. If, for the sake of simplicity, we normalize regional compliance costs to zero, $\lambda$ measures the cost difference in complying with the central government.

[34] Proofs corresponding to this section are contained in the Online Appendix.

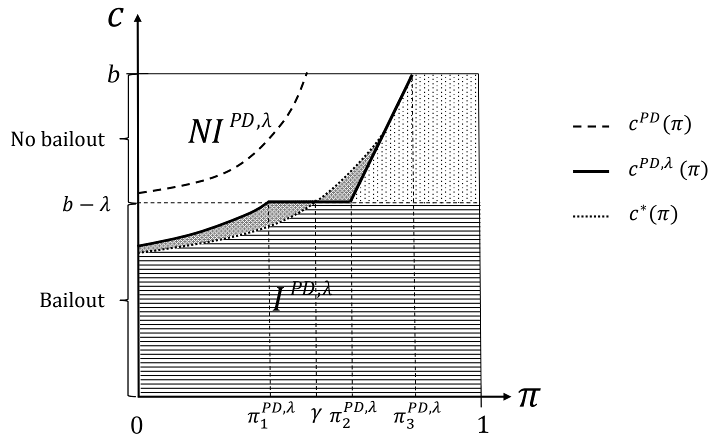


**Figure 8: Outcomes under partial decentralization with distortionary national taxation**

As expected, the central government refinances fewer projects than before. Indeed, only when $c \leq b - \lambda < b$, incomplete projects are bailed out; otherwise, the central government faces an endogenous hard budget constraint. When $0 \leq \pi \leq \pi_1^{PD,\lambda} \equiv \pi_1^{PD} - \frac{\lambda L}{L(B-b)+b}$, despite the fact that the central government bails out all projects with costs lower than $b - \lambda$, only those for which

$$c \leq c^{PD,\lambda}(\pi) \equiv c^{PD}(\pi) - \frac{(1-\pi)\lambda}{L+1-\pi}$$

are initiated and refinanced. Otherwise, projects are not undertaken because refinancing them is too costly. Projects falling within the shaded area are inefficiently initiated. But note that these are necessarily fewer than in the case in which national taxation is non distortionary because $c^{PD,\lambda}(\pi) < c^{PD}(\pi)$. As $\pi$ increases, the cost interval where these inefficient projects are initiated when the central government refinances vanishes. Thus, when the central government does not bail out regions, no project is undertaken in equilibrium, as shown when $\pi_1^{PD,\lambda} < \pi \leq \pi_2^{PD,\lambda} \equiv \pi_2^{PD} - \frac{\lambda}{B}$. Note that, for $\pi \in [\gamma, \pi_2^{PD,\lambda}]$ we have $c^{PD,\lambda}(\pi) < c^*(\pi)$. This implies the opposite type of distortion: in the shaded area, regions inefficiently underinvest. Then, as $\pi_2^{PD,\lambda} \leq \pi \leq \pi_3^{PD,\lambda} \equiv b/B$, all projects that would be refinanced are initiated. But this does not hold for projects where $c > b - \lambda$. These projects are undertaken provided $c \leq \pi B$; otherwise they are not initiated, and again there is inefficient underinvestment in the shaded area. Finally, when $\pi_3^{PD,\lambda} \leq \pi \leq 1$, all projects are initiated, despite the fact that some with high cost would not be refinanced if they remain incomplete in the second period.

The following proposition compares this regime with full decentralization.

**Proposition 7** *When the value of the deadweight loss $\lambda$ increases, full decentralization dominates in a larger area of the plane $(\theta, \pi)$.*

Clearly, increasing $\lambda$ favors full decentralization. Although expected, the intuition of this result should not be based on the mere assertion that by increasing the cost of bailouts, expected welfare under partial decentralization should automatically decrease. In fact, when $\lambda$ increases, this institutional regime does generate more costly bailouts than before, inefficient project's shutdown and underinvestment. But when the administrative capacity is relatively low, a higher $\lambda$ also reduces the size of the cost interval where projects are inefficiently initiated. Despite these countervailing impacts, we can show that the negative effects mentioned previously dominate, and thus expected welfare under this regime decreases with the deadweight loss $\lambda$.

# 7 Conclusion

This paper presents a model featuring a central government and regional authorities. The latter are characterized their levels of administrative and fiscal capacities. We analyze two fiscal regimes. Under partial decentralization, regional governments rely on central bailouts to refinance previously started projects. Hence, regions face soft budget constraints and can overinvest in local public projects. Under full decentralization, regional governments cannot rely on central bailouts and face hard budget constraints. In this scenario, capital tax competition increases the marginal cost of public funds and regional governments may underinvest.

The main goal of the paper is to conduct a normative comparison between these regimes and determine how different levels of regional state capacity affect this comparison. As expected, when the regional fiscal capacity is low, partial decentralization dominates. But contrary to the common wisdom, we find that full decentralization may be optimal even when the regional administrative capacity is low.

An interesting route for further research is to incorporate ex ante asymmetries between regions (either in capital endowments or in state capacities). This would yield a more suitable framework to deliver quantitative assessments, and would be an essential feature to endogenize regional state capacity formation under different fiscal regimes.

# References

[1] Acemoglu, D., C. García-Jimeno and J. Robinson (2015), "State Capacity and Economic Development: A Network Approach," *American Economic Review*, **105**, 2364-2409.

[2] Akai, N. and M. Sato (2008), "Too Big or Too Small? A Synthetic View of the Commitment Problem of Interregional Transfers," *Journal of Urban Economics*, **64**, 551-559.

[3] Arbetman-Rabinowitz, M., J. Kugler, M. Abdollahian, K. Kang, H. Nelson and R. Tammen (2012), "Political Performance" in Kugler, J. and R. Tammen (Eds.) *The Performance of Nations*, Rowman and Littlefield Publishers, Maryland.

[4] Aumann, R. (1959), "Acceptable Points in General Cooperative $n$-Person Games" in Luce, R. and A. Tucker (Eds.) *Contributions to the Theory of Games, Volume IV*, Princeton University Press.

[5] Bandyopadhyay, S. and E. Green (2012), "Pre-Colonial Political Centralization and Contemporary Development in Uganda," mimeo, LSE.

[6] Bardhan, P. and D. Mookherjee (2006), "Decentralization and Accountability in Infrastructure Delivery in Developing Countries," *The Economic Journal*, **116**, 101-127.

[7] Bernheim, D., D. Peleg and M. Whinston (1987), "Coalition-Proof Nash Equilibria. I: Concepts," *Journal of Economic Theory*, **42**, 1-12.

[8] Besfamille, M. and B. Lockwood (2008), "Bailouts in Federations: is a Hard Budget Constraint Always Best?," *International Economic Review*, **49**, 577-593.

[9] Besley, T. and T. Persson (2010), "State Capacity, Conflict and Development," *Econometrica*, **78**, 1-34.

[10] Bird, R. (1995), "Decentralizing Infrastructure: For Good or for Ill?" in Estache, A. (Ed.) *Decentralizing Infrastructure Advances and Limitations*, WB Discussion Papers 290, The World Bank, Washington DC.

[11] Boadway, R. and A. Shah (Eds.) (2007), *Intergovernmental Fiscal Transfers Principles and Practice*, The World Bank, Washington DC.

[12] Brueckner, J. (2009), "Partial Fiscal Decentralization," *Regional Science and Urban Economics*, **39**, 23-32.

[13] Bucovetsky, S. (1991), "Asymmetric Tax Competition," *Journal of Urban Economics*, **30**, 167-181.

[14] 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.

[15] Corbacho, A. , V. Fretes Cibils and E. Lora (Eds.) (2013), *More than Revenue: Taxation as a Development Tool*, Palgrave Macmillan, Washington DC.

[16] Cremer, H., M. Marchand and P. Pestieau (1997), "Investment in Local Public Services: Nash Equilibrium and Social Optimum," *Journal of Public Economics*, **65**, 23-35.

[17] Crivelli, E. and K. Staal (2013), "Size, Spillovers and Soft Budget Constraints," *International Tax and Public Finance*, **20**, 338-356.

[18] Eyraud, L. and L. Lusinyan (2013), "Vertical Fiscal Imbalances and Fiscal Performance in Advanced Economies," *Journal of Monetary Economics*, **60,** 571-587.

[19] Goodspeed, T. (2002), "Bailouts in a Federation," *International Tax and Public Finance*, **9**, 409-421.

[20] Guccio, C., G. Pignataro and I. Rizzo (2014), "Do Local Governments Do it Better? Analysis of Time Performance in the Execution of Public Works," *European Journal of Political Economy*, **34**, 237-252.

[21] Hanson, J. and R. Sigman (2013), "Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research", mimeo, Syracuse University.

[22] Hatfield, J. and G. Padró i Miquel (2012), "A Political Economy Theory of Partial Decentralization," *Journal of the European Economic Association*, **10**, 605-633.

[23] Inman, R. (2003), "Transfers and Bailouts: Enforcing Local Fiscal Discipline with Lessons from U.S. Federalism" in Rodden, J., G. Eskeland and J. Litvack (Eds.) *Fiscal Decentralization and the Challenge of Hard Budget Constraints*, The MIT Press, Cambridge Massachussets.

[24] Janeba, E. and J. Wilson (2011), "Optimal Fiscal Federalism in the Presence of Tax Competition," *Journal of Public Economics*, **95**, 1032-1311.

[25] Joanis, M. (2014), "Shared Accountability and Partial Decentralization in Local Public Good Provision," *Journal of Development Economics*, **107**, 28-37.

[26] Kornai, J., E. Maskin and G. Roland (2004), "Understanding the Soft Budget Constraint," *Journal of Economic Literature*, **41**, 1095-1136.

[27] Köthenbürger, M. (2004), "Tax Competition in a Fiscal Union with Decentralized Leadership," *Journal of Urban Economics*, **55**, 498-513.

[28] Köthenbürger, M. (2011), "How Do Local governments Decide on Public Policy in Fiscal Federalism? Tax vs. expenditure optimization," *Journal of Public Economics*, **95**, 1516-1522.

[29] Le, T., B. Moreno-Dodson and J. Rojchaichaninthorn (2008), "Expanding Taxable Capacity and Reaching Revenue Potential: Cross-Country Analysis," Policy Research Working Paper 4559, The World Bank, Washington DC.

[30] Litvack, J., J. Ahmad and R. Bird (1998), "Rethinking Decentralization in Developing Countries," Sector Studies Series, The World Bank, Washington DC.

[31] Loayza, N., J. Rigolini and O. Calvo-González (2014), "More Than You Can Handle: Decentralization and Spending Ability of Peruvian Municipalities," *Economics and Politics*, **26**, 56-78.

[32] Lockwood, B. (2002), "Distributive Politics and the Costs of Centralization," *The Review of Economic Studies*, **69**, 313-337.

[33] Mann, M. (1984), "The Autonomous Power of the State: Its Origins, Mechanisms and Results," *Archives Européennes de Sociologie*, **25**, 185-213.

[34] Oates, W. (2005), "Toward a Second-Generation Theory of Fiscal Federalism," *International Tax and Public Finance*, **12**, 349-373.

[35] Patil, S., A. Gupta, D. Desai and A. Sajane (2013), "Causes of Delay in Indian Transportation Infrastructure Projects," *International Journal of Research in Engineering and Technology*, **2**, 71-80.

[36] Peralta, S. (2011), "Partial Fiscal Decentralization, Local Elections and Accountability," mimeo, Nova School of Business and Economics.

[37] Persson, T. and G. Tabellini (1992), "The Politics of 1992: Fiscal Policy and European Integration," *Review of Economic Studies*, **59**, 689-701.

[38] Pettersson-Lidbom, P. (2010), "Dynamic Commitment and the Soft Budget Constraint: An Empirical Test," *American Economic Journal: Economic Policy*, **2**, 154–179.

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

[40] Qian, Y. and G. Roland (1998), "Federalism and the Soft-Budget Constraint," *American Economic Review*, **88**, 1146-1162.

[41] Rodden, J., G. Eskeland and J. Litvack (2003), *Fiscal Decentralization and the Challenge of Hard Budget Constraints*, The MIT Press, Cambridge Massachusetts.

[42] Sanguinetti, P. and M. Tommasi (2004) "Intergovernmental Transfers and Fiscal Behavior Insurance Versus Aggregate Discipline," *Journal of International Economics*, **62**, 149-170.

[43] Silva, E. and A. Caplan (1997), "Transboundary Pollution Control in Federal Systems," *Journal of Environmental Economics and Management*, **34**, 173-186.

[44] Slemrod, J. and V. Venkatesh (2002), "The Income Tax Compliance Cost of Large and Mid-Size Businesses," mimeo, Office of the Tax Policy Research, University of Michigan.

[45] Snyder, R. (2001), "Scaling Down: the Subnational Comparative Method," *Studies in Comparative International Development*, **36**, 93-110.

[46] Steiner, S. (2010), "How Important is the Capacity of Local Governments for Improvements in Welfare? Evidence from Decentralised Uganda," *Journal of Development Studies*, **46**, 644–661.

[47] Weingast, B. (2009), "Second Generation Fiscal Federalism: The Implications of Fiscal Incentives," *Journal of Urban Economics*, **65**, 279-293.

[48] Wildasin, D. (1988), "Nash Equilibria in Models of Fiscal Competition," *Journal of Public Economics*, **35**, 229-240.

[49] Wildasin, D. (1997), "Externalities and Bailouts Hard and Soft Budget Constraints in Intergovernmental Fiscal Relations," mimeo, Vanderbilt University.

[50] Wilson, J. (1991), "Tax Competition with Interregional Differences in Factor Endowments," *Regional Science and Urban Economics*, **21**, 423-451.

[51] Wilson, J. (1999), "Theories of Tax Competition," *National Tax Journal*, **52**, 269-304.

[52] Ziblatt, D. (2008), "Why Some Cities Provide More Public Goods than Others: A Subnational Comparison of the Provision of Public Goods in German Cities in 1912," *Studies in Comparative International Development*, **43**, 273-289.

[53] Zodrow, G. and P. Mieszkowski (1986), "Pigou, Tiebout, Property Taxation, and the Underprovision of Local Public Goods," *Journal of Urban Economics*, **19**, 356-370.

# Appendix

## Proof of Proposition 1

The government of region $\ell$ anticipates that its net expected welfare from investing in the project is

$$\kappa + \mathbb{I}_{\{i_\ell = I\}} \left[\pi B + (1-\pi)\left(b - \frac{c}{L}\right) - c\right] - \sum_{m \neq \ell} \mathbb{I}_{\{i_m = I\}} (1-\pi)\frac{c}{L}, \tag{4}$$

whereas its net expected welfare from not investing is

$$\kappa - \sum_{m \neq \ell} \mathbb{I}_{\{i_m = I\}} (1-\pi)\frac{c}{L}.$$

So, for any region $\ell$, initiating the project is a dominant strategy if

$$c \leq c^{PD}(\pi) \equiv \frac{L[\pi B + (1-\pi)b]}{L + 1 - \pi}.$$

Since

$$c^{PD}(\pi) - c^*(\pi) = \frac{(L-1)(1-\pi)[\pi B + (1-\pi)b]}{(L+1-\pi)(2-\pi)},$$

it follows that $c^{PD}(\pi) > c^*(\pi)$. Moreover, we can show that $\frac{\partial}{\partial \pi} c^{PD}(\pi) = L\frac{L(B-b)+B}{(L+1-p)^2} > 0$ and that $c^{PD} = b$ when $\pi = \frac{b}{L(B-b)+b}$. ■

## Proof of Proposition 2

Given a profile of tax rates $\tau = \{\tau_1, ..., \tau_L\}$, a household resident in region $\ell$ decides where to invest its capital endowment by solving the following problem:

$$\max_{h_\ell, \{f_{\ell m}\}_{m \neq \ell}} \; h_\ell (1 - \tau_\ell) + \sum_{m \neq \ell} f_{\ell m} (1 - \tau_m) - \frac{1}{2} (\sum_{m \neq \ell} f_{\ell m})^2$$

subject to its portfolio constraint

$$h_\ell + \sum_{m \neq \ell} f_{\ell m} = \kappa$$

and $(L-1)$ non-negativity constraints

$$f_{\ell m} \geq 0 \qquad \forall m \neq \ell,$$

where $h_\ell$ is capital invested in region $\ell$, and $f_{\ell m}$ is capital invested in region $m \neq \ell$. Denote by $\lambda_{\ell m}$ the multipliers associated with the non-negativity constraints. Using the portfolio constraint to replace $h_\ell$ in the maximand of the household's problem, we obtain the first-order conditions for $f_{\ell m}$ and the complementary slackness conditions

$$\begin{cases} \tau_\ell - \tau_m + \lambda_{\ell m} = \sum\limits_{m \neq \ell} f_{\ell m} \\ \\ \lambda_{\ell m} f_{\ell m} = 0 \quad \lambda_{\ell m} \geq 0 \quad \forall m \neq \ell. \end{cases}$$

The proof of the proposition uses the following two lemmas.

**Lemma 1** *Assume that there are two regions $m, n \neq \ell$, with $\tau_m > \tau_n$. Then $f_{\ell m} = 0$.*

**Proof.** Subtracting $m$'s first-order condition from $n$'s first-order condition, we obtain

$$\tau_m - \tau_n + \lambda_{\ell n} = \lambda_{\ell m}.$$

As $\tau_m > \tau_n$ and $\lambda_{\ell n} \geq 0$, $\lambda_{\ell m} > 0$. Hence, by the corresponding complementary slackness condition, $f_{\ell m} = 0$. ■

Let $\widetilde{M}$ be the set of regions $\widetilde{m} \neq \ell$ that have chosen the minimum tax rate $\widetilde{\tau}_\ell = \min\{\tau_m\}_{m \neq \ell}$. Then, as an immediate consequence of Lemma 1, for all regions $m \neq \ell, \widetilde{m}$, $f_{\ell m} = 0$. Hence, tax rates $\tau_m$ become redundant; from now on, pertinent comparisons should be done only between $\tau_\ell$ and $\widetilde{\tau}_\ell$.

**Lemma 2** *Assume that $\tau_\ell \geq \widetilde{\tau}_\ell$. If $\widetilde{m} \in \widetilde{M}$ then $\lambda_{\ell \widetilde{m}} = 0$.*

**Proof.** We consider two cases: (i) $\text{Card}\{\widetilde{M}\} = 1$, (ii) $\text{Card}\{\widetilde{M}\} \geq 2$.
(i) First, assume that $\text{Card}\{\widetilde{M}\} = 1$ and consider the first-order condition

$$\tau_\ell - \widetilde{\tau}_\ell + \lambda_{\ell \widetilde{m}} = f_{\ell \widetilde{m}}. \tag{5}$$

If $\tau_\ell > \widetilde{\tau}_\ell$, as $\lambda_{\ell\widetilde{m}} \geq 0$, then $f_{\ell\widetilde{m}} > 0$. Thus, by the complementary slackness condition, $\lambda_{\ell\widetilde{m}} = 0$. If $\tau_\ell = \widetilde{\tau}_\ell$, (5) becomes

$$\lambda_{\ell\widetilde{m}} = f_{\ell\widetilde{m}}.$$

If $\lambda_{\ell\widetilde{m}} > 0$, then $f_{\ell\widetilde{m}} > 0$, which implies that $\lambda_{\ell\widetilde{m}} = 0$, which is a contradiction. Hence $\lambda_{\ell\widetilde{m}} = 0$.

(ii) Now assume that Card$\{\widetilde{M}\} \geq 2$. First-order conditions that characterize flows $f_{\ell\widetilde{m}}$ are

$$\tau_\ell - \widetilde{\tau}_\ell + \lambda_{\ell\widetilde{m}} = \sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}}.$$

In order to satisfy these first-order conditions, all multipliers $\lambda_{\ell\widetilde{m}}$ should have the same value. If they were all strictly positive, then all outflows $f_{\ell\widetilde{m}}$ should be equal to 0, implying that $\sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}} = 0$. But, as $\tau_\ell \geq \widetilde{\tau}_\ell$, all first-order conditions would yield a contradiction. Hence, all multipliers are zero. ■

Therefore, from any first-order condition that characterizes a flow to a region $\widetilde{m} \in \widetilde{M}$, we obtain

$$\sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}} = \tau_\ell - \widetilde{\tau}_\ell. \tag{6}$$

## Proof of Proposition 3

To obtain the equilibrium tax rates, first we derive region $\ell$'s reaction function.[35]

### Scenario 1: All regions have decided to refinance their incomplete project

Denote by $\boldsymbol{\tau_m}$ the profile of tax rates chosen by regions $m \neq \ell$. For any profile $\boldsymbol{\tau_m}$, we need to consider three cases.

1. If the regional government of $\ell$ plans to set its tax rate strictly above $\widetilde{\tau}_\ell$, there will be capital outflows to regions $\widetilde{m} \in \widetilde{M}$. Hence, the regional welfare would be

$$W_\ell^{FD} = (\kappa - \sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}})\,(1-\tau_\ell) + \sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}}\,(1-\widetilde{\tau}_\ell) - \frac{1}{2}(\sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}})^2 + b.$$

   By the Envelope Theorem, $\partial W_\ell^{FD}/\partial\tau_\ell = -(\kappa - \sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}}) < 0$. So the regional government of $\ell$ should set the lowest tax rate that satisfies its budget constraint

$$\theta\tau_\ell(\kappa - \sum_{\widetilde{m}\in\widetilde{M}} f_{\ell\widetilde{m}}) = c. \tag{7}$$

[35]Due to specific features of this model, we cannot apply Wildasin's (1988) methodology to derive the equilibrium tax rates.

Using (6), the smallest root of (7) is given by

$$\overline{\tau}_\ell = \frac{1}{2}\left[\kappa + \widetilde{\tau}_\ell - \sqrt{(\kappa + \widetilde{\tau}_\ell)^2 - \frac{4c}{\theta}}\right]. \tag{8}$$

Throughout the paper, we assume that $\kappa$ is so large so that this square root always exists (see footnote 36).

2. If the regional government of $\ell$ plans to set its tax rate strictly below $\widetilde{\tau}_\ell$, there will be no capital outflows to regions $m \neq \ell$. Thus, regional welfare would be

$$W_\ell^{FD} = \kappa\,(1 - \tau_\ell) + b.$$

Again, by the Envelope Theorem, $\partial W_\ell^{FD}/\partial \tau_\ell = -\kappa < 0$. So, regional government of $\ell$ should choose the lowest tax rate that satisfies its budget constraint

$$\theta\tau_\ell(\kappa + \sum_{m\neq\ell} f_{m\ell}) = c, \tag{9}$$

where

$$\sum_{m\neq\ell} f_{m\ell} = \sum_{m\neq\ell} \tau_m - (L-1)\tau_\ell \tag{10}$$

represents all capital inflows that leave regions $m \neq \ell$, and enter region $\ell$. Rearranging terms, the smallest root of (9) is given by

$$\underline{\tau}_\ell = \frac{1}{2}\left[\frac{1}{L-1}\left(\kappa + \sum_{m\neq\ell}\tau_m\right) - \sqrt{\left(\frac{\kappa + \sum_{m\neq\ell}\tau_m}{L-1}\right)^2 - \frac{4c}{\theta(L-1)}}\right]. \tag{11}$$

3. If the regional government of $\ell$ plans to replicate $\widetilde{\tau}_\ell$, there will be capital outflows from regions $m \notin \widetilde{M}$ to regions $\widetilde{m} \in \widetilde{M}$. Since $\ell \in \widetilde{M}$, the regional welfare of $\ell$ would be

$$W_\ell^{FD} = \kappa\,(1 - \widetilde{\tau}_\ell) + \mathbb{1}_{\{ETC \geq c\}} b,$$

where

$$ETC = \theta\widetilde{\tau}_\ell(\kappa + \sigma_\ell \sum_{m\notin\widetilde{M}} \sum_{\widetilde{m}\in\widetilde{M}} f_{m\widetilde{m}}),$$

is the effective tax collection, and $\sigma_\ell$, $0 < \sigma_\ell < 1$, represents the fraction of all capital outflows that leave regions $m \notin \widetilde{M}$ and move to region $\ell$.

Now, in order to completely characterize the reaction function $\tau_\ell(\boldsymbol{\tau_m})$, we need to identify profiles of tax rates $\boldsymbol{\tau_m}$ for which $\underline{\tau}_\ell$, $\overline{\tau}_\ell$ and $\widetilde{\tau}_\ell$ are region $\ell$'s best responses.

First suppose that, facing a profile of tax rates $\boldsymbol{\tau_m}$, region $\ell$ wishes to set $\underline{\tau}_\ell$. It is straightforward to show that, for any change in one particular tax rate $\tau_m$, $m \neq \ell$ that does not modify $\widetilde{\tau}_\ell$,

$$\frac{\partial \underline{\tau}_\ell}{\partial \tau_m} < 0.$$

Therefore, given $\widetilde{\tau}_\ell$, the highest tax rate $\underline{\tau}_\ell < \widetilde{\tau}_\ell$ is set when all regions $m \neq \ell$ have chosen $\tau_m = \widetilde{\tau}_\ell$. Denote by $\underline{\tau}_\ell(\widetilde{\tau}_\ell)$ this tax rate, which is given by

$$\underline{\tau}_\ell(\widetilde{\tau}_\ell) = \frac{1}{2}\left[\frac{\kappa}{L-1} + \widetilde{\tau}_\ell - \sqrt{\left(\frac{\kappa}{L-1} + \widetilde{\tau}_\ell\right)^2 - \frac{4c}{\theta(L-1)}}\right].$$

We can show that

$$\frac{\partial \underline{\tau}_\ell(\widetilde{\tau}_\ell)}{\partial \widetilde{\tau}_\ell} < 0$$

and

$$\lim_{\widetilde{\tau}_\ell \downarrow c/\theta\kappa} \underline{\tau}_\ell(\widetilde{\tau}_\ell) = c/\theta\kappa.$$

If $\widetilde{\tau}_\ell \leq c/\theta\kappa$, $\underline{\tau}_\ell(\widetilde{\tau}_\ell) \geq \widetilde{\tau}_\ell$, contradicting the definition of $\underline{\tau}_\ell(\widetilde{\tau}_\ell)$ as the highest tax rate that region $\ell$ can set strictly below $\widetilde{\tau}_\ell$. Therefore, if $\widetilde{\tau}_\ell < c/\theta\kappa$, the government in region $\ell$ cannot set a tax rate strictly less than $\widetilde{\tau}_\ell$ while, at the same time, satisfying its budget constraint. A similar argument applies if the regional government wishes to set $\tau_\ell = \widetilde{\tau}_\ell < c/\theta\kappa$. If this were the case, the effective tax collection would be

$$ETC = \theta\widetilde{\tau}_\ell\left(\kappa + \sigma_\ell \sum_{m \notin \widetilde{M}} \sum_{\widetilde{m} \in \widetilde{M}} f_{m\widetilde{m}}\right).$$

Consider the most favorable case for region $\ell$: if $\sigma_\ell = 1$ and $(L-2)$ regions have set a tax rate equal to one, region $\ell$ receives the largest capital flow. Under this circumstance,

$$ETC < \frac{c}{\kappa}\left(\kappa + (L-2)(1-\widetilde{\tau}_\ell)\right) \simeq c$$

because $\tau_\ell < c/\theta\kappa$ and $\kappa$ is assumed to be very large. Hence, when $\widetilde{\tau}_\ell < c/\theta\kappa$, the regional government of $\ell$ can only set a tax rate higher than $\widetilde{\tau}_\ell$ to cover the refinancing cost.

Now consider that, facing a profile of tax rates $\boldsymbol{\tau_m}$, region $\ell$ wishes to set $\overline{\tau}_\ell > \widetilde{\tau}_\ell$. It is straightforward to show that

$$\lim_{\widetilde{\tau}_\ell \to 0} \overline{\tau}_\ell = \frac{1}{2}\left[\kappa - \sqrt{\kappa^2 - \frac{4c}{\theta}}\right] \equiv \tau_\ell(0) > 0,$$

$$\frac{\partial \overline{\tau}_\ell}{\partial \widetilde{\tau}_\ell} < 0\,, \quad \frac{\partial^2 \overline{\tau}_\ell}{\partial \widetilde{\tau}_\ell^{\,2}} > 0$$

and

$$\lim_{\widetilde{\tau}_\ell \uparrow c/\theta\kappa} \overline{\tau}_\ell = c/\theta\kappa.$$

If $\widetilde{\tau}_\ell \geq c/\theta\kappa$, $\overline{\tau}_\ell(\widetilde{\tau}_\ell) \leq \widetilde{\tau}_\ell$, which again contradicts the definition of $\overline{\tau}_\ell$ as the lowest tax rate that region $\ell$ can set strictly above $\widetilde{\tau}_\ell$. Hence, if $\widetilde{\tau}_\ell > c/\theta\kappa$, the government in region $\ell$ cannot set a tax higher than $\widetilde{\tau}_\ell$ while, at the same time, satisfying its budget constraint. If the regional government sets $\tau_\ell = \widetilde{\tau}_\ell > c/\theta\kappa$, there will be no outflows $f_{\ell m}$. Thus, region $\ell$'s welfare will amount to $W_\ell^{FD} = \kappa - c$. But this region will receive a capital inflow $\sigma_\ell \sum_{m \notin \widetilde{M}} \sum_{\widetilde{m} \in \widetilde{M}} f_{m\widetilde{m}}$, and so its tax collection will be

$$\theta \frac{c}{\theta\kappa}(\kappa + \sigma_\ell \sum_{m \notin \widetilde{M}} \sum_{\widetilde{m} \in \widetilde{M}} f_{m\widetilde{m}}) > c.$$

Therefore, there will be room for a decrease in $\tau_\ell$. Indeed, region $\ell$ can set its tax rate $\underline{\tau}_\ell < c/\theta\kappa$, satisfying its budget constraint and increasing its welfare (with respect to the choice of $\tau_\ell = \widetilde{\tau}_\ell > c/\theta\kappa$).

Two more cases remain to be analyzed. First, consider a profile of tax rates $\boldsymbol{\tau_m}$ such that $\widetilde{\tau}_\ell = c/\theta\kappa$ but when at least one region $n \neq \ell$ has set $\tau_n > c/\theta\kappa$. By a similar argument than before, region $\ell$ can set its tax rate $\underline{\tau}_\ell < c/\theta\kappa$, satisfying its budget constraint and increasing its welfare (with respect to the choice of $\tau_\ell = \widetilde{\tau}_\ell = c/\theta\kappa$.

Finally, when all regions $m \neq \ell$ have set the same tax rate $\tau_m = c/\theta\kappa$, we have already shown that region $\ell$ cannot tax strictly above or below $c/\theta\kappa$. Hence, the optimal response is to set $\tau_\ell = c/\theta\kappa$.

Region $\ell$'s reaction function is thus characterized as follows:

$$\tau_\ell\left(\boldsymbol{\tau_m}\right) = \begin{cases} \frac{1}{2}\left[\kappa + \widetilde{\tau}_\ell - \sqrt{\left(\kappa + \widetilde{\tau}_\ell\right)^2 - \frac{4c}{\theta}}\right] & \text{if } \tau_m < c/\theta\kappa \ \forall m \neq \ell \\ c/\theta\kappa & \text{if } \tau_m = c/\theta\kappa \ \forall m \neq \ell \\ \frac{1}{2}\left[\frac{\kappa + \sum\limits_{m \neq \ell} \tau_m}{L-1} - \sqrt{\left(\frac{\kappa + \sum\limits_{m \neq \ell} \tau_m}{L-1}\right)^2 - \frac{4c}{\theta(L-1)}}\right] & \begin{array}{l} \text{if } \tau_m > c/\theta\kappa \ \forall m \neq \ell \text{ or} \\ \tau_m = c/\theta\kappa \text{and } \exists n \neq \ell : \tau_n > c/\theta\kappa. \end{array} \end{cases} \tag{12}$$

Region $\ell$'s reaction function is non-continuous. When $\widetilde{\tau}_\ell$ converges to $c/\theta\kappa$ from below, $\tau_\ell\left(\boldsymbol{\tau_m}\right)$ converges to this limit from above. But when the distribution of taxes is such that $\widetilde{\tau}_\ell = c/\theta\kappa$ but at least one region $n \neq \ell$ has set $\tau_n > c/\theta\kappa$, the optimal response is to set $\tau_\ell\left(\boldsymbol{\tau_m}\right) < c/\theta\kappa$.

Clearly, $\widehat{\tau}_1 = .... = \widehat{\tau}_L = c/\theta\kappa$ is a Nash equilibrium of this subgame because it is a fixed point of the best response correspondence. To prove uniqueness, we proceed in two steps. First, by simple inspection of (12), it is immediate that an asymmetric choice of taxes cannot be a Nash equilibrium. Second, there cannot be another symmetric equilibrium. Assume the contrary: let $\widehat{\tau}'_\ell = \widehat{\tau}'_m$ and, without any loss of generality,

$$\widehat{\tau}'_\ell = \widehat{\tau}'_m = \frac{c}{\theta\kappa} + \varepsilon, \tag{13}$$

with $\varepsilon \neq 0$ be another symmetric equilibrium. By substituting (13) into (12), we obtain a contradiction. ■

**Scenario 2: At least one region has decided not to refinance**

In this case, to refinance its incomplete project the regional government of $\ell$ has to set the tax rate $\widehat{\tau}_\ell(0)$. This tax rate is obtained replacing $\widetilde{\tau}_\ell$ by 0 in the definition of $\overline{\tau}_\ell$.[36] ■

## Proof of Proposition 4

First, we prove the existence of $c_1$ and $c_2$. When there is at least one region that does not need refinancing, the total cost from completing a project in any region $T(c,\theta)$ is a strictly increasing and convex function of $c$, that satisfies

$$\lim_{c\to 0} T(c,\theta) = 0 \quad \text{and} \quad \lim_{c\to b} T(c,\theta) = \frac{b}{\theta} + \frac{\left[\lim_{c\to b} \widehat{\tau}_\ell(0)\right]^2}{2} > b.$$

Hence, by Bolzano's Theorem, there exists a threshold $0 < c_1 < b$ such that, when $c \leq c_1$, $b - T(c,\theta) \geq 0$. Also, as $c/\theta > c$, there exists a threshold $c_2$ such that $c_2/\theta = b$. Moreover, as $c/\theta < T(c,\theta)$, $c_1 < c_2$.

Now consider the first case, when all regions face an incomplete project. When $0 \leq c \leq c_1$,

$$W_\ell^{FD}(r_\ell = R, \mathbf{r_m}, \mathbf{i}) = \kappa + b - c/\theta - c \geq W_\ell^{FD}(r_\ell = NR, \mathbf{r_m}, \mathbf{i}) = \kappa - c.$$

Here refinancing is a dominant strategy.

When $c_1 < c \leq c_2$,

$$W_\ell^{FD}(r_\ell = R, r_m = R, \mathbf{i}) = \kappa + b - c/\theta - c \geq W_\ell^{FD}(r_\ell = NR, r_m = R, \mathbf{i}) = \kappa - c \qquad \forall m \neq \ell$$

but

$$W_\ell^{FD}(r_\ell = R, r_m = NR, \mathbf{i}) = \kappa + b - T(c,\theta) - c \leq W_\ell^{FD}(r_\ell = NR, r_m = NR, \mathbf{i}) = \kappa - c \qquad \forall m \neq \ell.$$

These payoffs define a coordination (sub)game between regions, with two Nash equilibria. In the first equilibrium all regions refinance, while in the second one no region refinances.

In fact, we prove that the equilibrium where all regions refinance is both the strong (Aumann, 1959) and the coalition-proof Nash equilibrium (Bernheim et al., 1987), as follows.

[36] A sufficient condition for the existence of the square roots in (8) and (11), and to ensure that $\widehat{\tau}_\ell < 1$ is $\kappa \geq \max\{2\sqrt{b/\theta}, b/\theta, (L-1)4c/\theta\}$.

1. As both equilibria are Nash equilibria, no region can do better by unilaterally changing its equilibrium strategy. Then consider $\mathcal{L}-$regions coalitions, with $1 < \mathcal{L} < L$. If the other regions refinance (not refinance), the $\mathcal{L}$ regions do not want to deviate because $b - c/\theta \geq 0$ ($b - T(c, \theta) < 0$). Finally, consider the $L$-regions coalition. If all regions refinance, they do not want to deviate since $b - c/\theta \geq 0$. But, when no region refinances, they all wish to deviate because the first Nash equilibrium is Pareto optimal. Hence, the unique equilibrium that is strong is the first one.

2. Again, as both equilibria are Nash equilibria, no region can do better by unilaterally changing its equilibrium strategy. Then consider $\mathcal{L}-$regions coalitions, with $1 < \mathcal{L} < L$. If the other regions refinance, the coalition of $\mathcal{L}$ regions can jointly decide not to refinance. Although this deviation is self-enforcing, it is not worthy because $b - c/\theta \geq 0$. But if the other regions do not refinance, the unique available deviation to the coalition is to refinance. However this deviation is not self-enforcing because $b - T(c, \theta) < 0$. Finally, consider the $L$-regions coalition. If all regions refinance, they can jointly decide to not refinance. Although this deviation is self-enforcing, it is not worthy because $b - c/\theta \geq 0$. But when no region refinances, they all wish to deviate because the first Nash equilibrium is Pareto optimal. Hence, the unique coalition-proof Nash equilibrium is the first one.

Therefore, when $c_1 < c \leq c_2$, we choose the equilibrium where all regions refinance as the Nash equilibrium of this subgame.

Finally, when $c_2 < c \leq b$,

$$W_\ell^{FD}(r_\ell = R, \mathbf{r}_{\boldsymbol{m}}, \mathbf{i}) = \kappa + b - c/\theta - c \leq W_\ell^{FD}(r_\ell = NR, \mathbf{r}_{\boldsymbol{m}}, \mathbf{i}) = \kappa - c.$$

So not refinancing is a dominant strategy.

The proof of the second part of the proposition is immediate, and thus omitted. ■

## Proof of Proposition 5

First, we proceed to evaluate net expected regional welfares under different parameter conditions. If $c \leq c_1$, region $\ell$'s net expected welfare is:

$$\begin{aligned}
\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = I) &= \kappa - c + \pi B + (1-\pi)^L[b - \tfrac{c}{\theta}] \\
&\quad -(1-\pi)(1-(1-\pi)^{L-1})[b - T(c,\theta)] && \text{if } i_m = I\ \forall m \neq \ell \\
\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = NI) &= \kappa - c + \pi B + (1-\pi)[b - T(c,\theta)] && \text{if } \exists m \neq \ell : i_m = NI_\ell \\
\mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m) &= \kappa && \forall i_m \in \{I, NI\}
\end{aligned}$$

If $c_1 < c \leq c_2$, region $\ell$'s net expected welfare is:

$$\begin{aligned}
\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = NI) &= \kappa - c + \pi B + (1-\pi)^L[b - \tfrac{c}{\theta}] && \text{if } i_m = I\ \forall m \neq \ell \\
\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = NI) &= \kappa - c + \pi B && \text{if } \exists m \neq \ell : i_m = NI_\ell \\
\mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m) &= \kappa && \forall i_m \in \{I, NI\}
\end{aligned}$$

When $c_2 < c \leq b$, region $\ell$'s net expected welfare is:

$$\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m) = \kappa - c + \pi B \quad \forall i_m \in \{I, NI\}$$
$$\mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m) = \kappa \quad \forall i_m \in \{I, NI\}$$

In order to characterize the Nash equilibria, we need to evaluate these different levels of regional welfare. We proceed as follows. First, define $\Delta^*(c) \equiv (1-\pi)[B - (b-c)]$. $\Delta^*(c)$ is an increasing, linear function of $c$ that satisfies

$$\lim_{c \to 0} \Delta^*(c) = \Delta^*(0) \equiv (1-\pi)[B - b] \quad \text{and} \quad \lim_{c \to b} \Delta^*(c) = \Delta^*(b) \equiv (1-\pi)B.$$

Then, when $c \leq c_1$, we can show that:

1. $\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = I) \geq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m) \Leftrightarrow \Delta_R^{FD}(c) \leq B - c$, where
$$\Delta_R^{FD}(c) \equiv (1-\pi)\left[B - (1-\pi)^{L-1}\left(b - \frac{c}{\theta}\right) - (1-(1-\pi)^{L-1})(b - T(c,\theta))\right]$$
is a continuous, increasing, convex function of $c$ that satisfies
$$\lim_{c \to 0} \Delta_R^{FD}(c) \equiv \Delta_R^{FD}(0) = (1-\pi)[B - b] = \Delta^*(0) \quad , \quad \Delta_R^{FD}(c) > \Delta^*(c)$$
and
$$\lim_{c \to c_1} \Delta_R^{FD}(c) = \Delta_R^{FD}(c_1) \equiv (1-\pi)\left[B - (1-\pi)^{L-1}\left(b - \frac{c_1}{\theta}\right)\right] < \Delta^*(b).$$

2. $\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = NI) \geq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m) \Leftrightarrow \Delta'(c) \leq B - c$, where
$$\Delta'(c) \equiv (1-\pi)[B - (b - T(c,\theta))]$$
is another continuous, increasing, convex function of $c$ that satisfies
$$\lim_{c \to 0} \Delta'(c) = (1-\pi)[B - b] = \Delta^*(0) \quad , \quad \Delta'(c) > \Delta_R^{FD}(c) > \Delta^*(c)$$
and
$$\lim_{c \to c_1} \Delta'(c) = \Delta'(c_1) \equiv (1-\pi)B = \Delta^*(b).$$

When $c_1 < c \leq c_2$, we can also show that:

1. $\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = I) \geq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m) \Leftrightarrow \Delta_{NAR}^{FD}(c) \leq B - c$, where
$$\Delta_{NAR}^{FD}(c) \equiv (1-\pi)\left[B - (1-\pi)^{L-1}\left(b - \frac{c}{\theta}\right)\right]$$
is a continuous, increasing, linear function of $c$ that satisfies
$$\lim_{c \to c_1} \Delta_{NAR}^{FD}(c) = \Delta_R^{FD}(c_1) \quad , \quad \Delta_{NAR}^{FD}(c) > \Delta^*(c)$$
and
$$\lim_{c \to c_2} \Delta_{NAR}^{FD}(c) = \Delta_{NAR}^{FD}(c_2) \equiv (1-\pi)B = \Delta^*(b).$$

2. $\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = NI) \geq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m) \Leftrightarrow \Delta''(c) \leq B - c$, where

$$\Delta''(c) \equiv (1-\pi)B$$

satisfies

$$\Delta''(c) = \Delta_{NAR}^{FD}(c_2) = \Delta'(c_1) = \Delta^*(b).$$

When $c_2 < c \leq b$, we show that $\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m) \geq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m) \Leftrightarrow \Delta_{NR}^{FD}(c) \leq B - c$, where

$$\Delta_{NR}^{FD}(c) \equiv (1-\pi)B = \Delta''(c).$$

The functions $\Delta^*(c)$, $\Delta_R^{FD}(c)$, $\Delta'(c)$, $\Delta_{NAR}^{FD}(c)$ and $\Delta_{NR}^{FD}(c)$ are (weakly) increasing in $c$, while $B-c$ decreases with $c$. Therefore, at some point, these five functions intersect $B-c$. Next, we first characterize these intersections as cost thresholds and divide the cost range $[0,b]$ in sub-intervals, according to these thresholds. Then we find the Nash equilibria in each corresponding sub-interval.

1. $0 \leq \pi \leq \pi_1^{FD}$

   Let $\pi_1^{FD}$ be implicitly defined by $\Delta_R^{FD}(c_1) = B - c_1$. As $\Delta'(c) \geq \Delta_R^{FD}(c) \geq \Delta^*(c)$, the intersection between $\Delta'(c)$ and $B-c$ defines a threshold $c'(\pi) \leq c^*(\pi)$, whereas the intersection between $\Delta_R^{FD}(c)$ and $B-c$ defines another threshold $c_R^{FD}(\pi)$ that satisfies $c'(\pi) \leq c_R^{FD}(\pi) \leq c^*(\pi)$. The Nash equilibria are the following:[37]

   (a) When $0 \leq c \leq c'(\pi)$, $\Delta_R^{FD}(c) \leq \Delta'(c) \leq B - c$, which implies that

   $$\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = I) \geq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m)$$

   and

   $$\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = NI) \geq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m).$$

   Hence all regions initiate their project and, if it remains incomplete at the end of $t=2$, they refinance it in $t=3$.

   (b) When $c'(\pi) < c \leq c_R^{FD}(\pi)$, $\Delta_R^{FD}(c) \leq B - c < \Delta'(c)$, which implies that

   $$\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = I) \geq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m)$$

   but

   $$\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = NI) \leq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m).$$

   Two Nash equilibria emerge: i) all regions initiate their project and, if it remains incomplete at the end of $t=2$, they refinance it in $t=3$, or ii) no region invests. We choose the first equilibrium as it is strong and coalition-proof.

[37] We only present the complete proof for this case. The remaining cases are analogous.

(c) When $c_R^{FD} < c$, $B - c < \Delta_R^{FD}(c) < \Delta'(c)$, which implies that

$$\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = I) \leq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m)$$

and

$$\mathbb{E}W_\ell^{FD}(i_\ell = I, i_m = NI) \leq \mathbb{E}W_\ell^{FD}(i_\ell = NI, i_m).$$

Hence, no region initiates a project.

2. $\pi_1^{FD} < \pi \leq \pi_2^{FD}$

Let $\pi_2^{FD} \equiv c_1/B$ and $c_{NAR}^{FD}(\pi) \leq c^*(\pi)$ be defined by the intersection between $\Delta_{NAR}^{FD}(c)$ and $B - c$. The Nash equilibria are the following:

(a) When $0 \leq c \leq c_1$, all regions initiate their project and, if it remains incomplete at the end of $t = 2$, they refinance it in $t = 3$.

(b) When $c_1 < c \leq c_{NAR}^{FD}(\pi)$, all regions initiate their project and, if it remains incomplete at the end of $t = 2$, they refinance it in $t = 3$, provided all other regions do the same.

(c) When $c_{NAR}^{FD}(\pi) < c \leq b$, no region initiates a project.

3. $\pi_2^{FD} < \pi \leq \pi_3^{FD} \equiv b/B$

Let $c_{NR}^{FD}(\pi) \equiv \pi B \leq c^*(\pi)$ be defined by the intersection between $\Delta_{NR}^{FD}(c)$ and $B - c$. The Nash equilibria are the following:

(a) When $0 \leq c \leq c_1$, all regions initiate their project and, if it remains incomplete at the end of $t = 2$, they refinance it in $t = 3$.

(b) When $c_1 < c \leq c_2$, all regions initiate their project and, if it remains incomplete at the end of $t = 2$, they refinance it in $t = 3$, provided all other regions do the same.

(c) When $c_2 < c \leq c_{NR}^{FD}(\pi)$, all regions initiate their project but, if it remains incomplete at the end of $t = 2$, they do not refinance it.

(d) When $c_{NR}^{FD}(\pi) < c \leq b$, no region initiates a project.

4. $\pi_3^{FD} < \pi \leq 1$

The Nash equilibria are the following:

(a) When $0 \leq c \leq c_1$, all regions initiate their project and, if it remains incomplete at the end of $t = 2$, they refinance it in $t = 3$.

(b) When $c_1 < c \leq c_2$, all regions initiate their project and, if it remains incomplete at the end of $t = 2$, they refinance it in $t = 3$, provided all other regions do the same.

(c) When $c_2 < c \leq b$, all regions initiate their project but, if it remains incomplete at the end of $t = 2$, they do not refinance it ■

## Proof of Proposition 6

The proof of this proposition uses the expected welfare expressions derived in the Online Appendix.

### Welfare analysis when $\pi$ adopts extreme values

We consider two cases: (i) $\pi = 0$, (ii) $b/B \leq \pi \leq 1$.

(i) $\underline{\pi = 0}$

When $\pi = 0$, $c^*(0) = b/2$, $c^{PD}(0) = \frac{Lb}{L+1}$ and $c_R^{FD}(0) = \frac{\theta b}{1+\theta}$. Hence, expected welfare under each regime is

$$\mathbb{E}W^{PD}(0) = \kappa + \frac{bL}{(1+L)^2} \quad \text{and} \quad \mathbb{E}W^{FD}(0,\theta) = \kappa + \frac{\theta b}{2(1+\theta)}.$$

When $\theta \geq \theta_0$, $\mathbb{E}W^{FD} \geq \mathbb{E}W^{PD}$. Otherwise, partial decentralization dominates.

Last, we analyze the value of $\left.\frac{\partial \mathbb{E}W^{FD}(\pi,\theta)}{\partial \pi}\right|_{\pi=0}$ as a function of the fiscal capacity $\theta$. For any pair $(b,\theta)$, we can always find a real number $\beta$ and write $\kappa = \beta\sqrt{\frac{4b}{1+\theta}}$. Computing the abovementioned derivative, we obtain

$$-\frac{\theta L[-12bB(1+\theta) + 8(L-1)\beta b^2(\beta^2-1)^{\frac{3}{2}} + b^2(9 - 12\beta^2 + 8\beta^4 + 12\theta + L(-8\beta^4 + 2\beta^2 - 3))]}{12(1+\theta)^2}.$$

The value of $L$ that makes this expression equal to zero is

$$\tilde{L}(\theta) = 1 + \frac{3[4B(1+\theta) - 2b(1+2\theta)]}{b[-8\beta^4 + 12\beta^2 + 8\beta(\beta^2-1)^{\frac{3}{2}} - 3]} \geq 1.$$

The fraction's denominator converges fast to zero, i.e., for values of $\beta$ below 1. But recall that in footnote 36 we have imposed $\kappa$ to be sufficiently large; so $\beta$ should be in fact a real number much larger than 1. Thus, $\tilde{L}(\theta)$ turns out to be a big number. Moreover, $\tilde{L}(\theta)$ increases with $\theta$. Hence, if $L \leq \bar{L} \equiv \tilde{L}(0)$

$$\left.\frac{\partial \mathbb{E}W^{FD}(\pi,\theta)}{\partial \pi}\right|_{\pi=0} \geq 0, \forall \theta \in [0,1]$$

and it increases with $\theta$. As we prove in the Online Appendix that $\mathbb{E}W^{FD}(\pi,\theta)$ is convex, we can conclude that $\mathbb{E}W^{FD}(\pi,\theta)$ is an increasing function of the administrative capacity $\pi$.

(ii) $\underline{\pi_2^{PD} = \pi_3^{FD} = b/B \leq \pi \leq 1}$

When $b/B \leq \pi < 1$, partial decentralization replicates the first best outcomes, whereas full decentralization's outcomes are distorted. Hence, partial decentralization dominates. But if we compute

$$\frac{\partial}{\partial \pi}\left[\mathbb{E}W^{FD}(\pi,\theta) - \mathbb{E}W^{PD}(\pi)\right]$$

and we take the limit when $\pi$ converges to one, we obtain

$$\lim_{\pi \to 1} \frac{\partial}{\partial \pi} \left[ \mathbb{E}W^{FD}(\pi, \theta) - \mathbb{E}W^{PD}(\pi) \right] = \frac{1}{b} \left( \int_0^b [b - c] dc - \int_0^{c_1} [b - T(c, \theta)] dc \right).$$

As $c_1 < b$ and $T(c, \theta) > c$, this limit is strictly positive. So, when $\pi$ converges to 1, the full decentralization expected welfare converges, from below, to the first best level. When $\pi = 1$,

$$\mathbb{E}W^{PD}(1) = \mathbb{E}W^{FD}(1, \theta) = \kappa + B - \frac{b}{2}.$$

**Welfare comparison when $\theta \in [\theta^*, 1]$**

First, assume that $\theta = 1$. When $\pi = 0$,

$$\mathbb{E}W^{FD}(0, 1) = \frac{b}{4},$$

and when $\pi = \pi_1^{PD}$,

$$\mathbb{E}W^{PD}(\pi_1^{PD}) = \frac{b(2B - b)}{2[BL - b(L - 1)]}.$$

If $L \geq \underline{L} \equiv 3 + \frac{B}{B-b}$, $\mathbb{E}W^{FD}(0, 1) \geq \mathbb{E}W^{PD}(\pi_1^{PD})$. Hence, $\mathbb{E}W^{FD}(\pi, \theta)$ lies everywhere above $\mathbb{E}W^{PD}(\pi)$ when $\pi \in [0, \pi_1^{PD}]$ because the former is an increasing function of $\pi$. Therefore, as $\mathbb{E}W^{FD}(\pi, \theta)$ must converge to $\mathbb{E}W^{PD}(\pi)$ from below when $\pi$ converges to one, $\mathbb{E}W^{FD}(\pi, \theta)$ has to cross $\mathbb{E}W^{PD}(\pi)$ in its linear part, from above. Denote by $\hat{\pi}(1)$ the administrative capacity level that corresponds to this intersection.[38] In fact, $\hat{\pi}(1)$ is unique. If this were not the case, $\mathbb{E}W^{FD}(\pi, \theta)$ would cross again $\mathbb{E}W^{PD}(\pi)$, from below. But, if such second intersection occurs, by convexity, $\mathbb{E}W^{FD}(\pi, \theta)$ could not converge to $\mathbb{E}W^{PD}(\pi)$ at $\pi = 1$.

When $\theta$ decreases, $\mathbb{E}W^{FD}(\pi, \theta)$ decreases as well, while $\mathbb{E}W^{PD}(\pi)$ remains constant. By continuity, we know that there exists $\theta^*$, a value of the regional fiscal capacity such that $\mathbb{E}W^{FD}(\pi, \theta^*)$ lies above $\mathbb{E}W^{PD}(\pi)$ everywhere when $\pi \in [0, \pi_1^{PD}]$. Hence, when $\theta \in [\theta^*, 1]$, $\mathbb{E}W^{FD}(\pi, \theta)$ intersects $\mathbb{E}W^{PD}(\pi)$ in its linear part. Using the same geometrical argument as before, we know that both expected welfares cross only once, at $\hat{\pi}(\theta)$. Applying the Implicit Function Theorem, we show that

$$\frac{\partial \hat{\pi}(\theta)}{\partial \theta} = -\frac{\partial \mathbb{E}W^{FD}(\hat{\pi}(\theta), \theta)/\partial \theta}{\partial \mathbb{E}W^{FD}(\hat{\pi}(\theta), \theta)/\partial \pi - \partial \mathbb{E}W^{PD}(\hat{\pi}(\theta))/\partial \theta} > 0$$

because, at $\hat{\pi}(\theta)$, $\mathbb{E}W^{FD}(\pi, \theta)$ crosses $\mathbb{E}W^{PD}(\pi)$ from above. Thus, $\hat{\pi}(\theta)$ increases with $\theta$.

When $\theta < \theta^*$, we cannot a priori ensure that $\mathbb{E}W^{FD}(\pi, \theta)$ crosses $\mathbb{E}W^{PD}(\pi)$ only once. The goal of the following paragraphs is to prove that this is indeed the case.

[38] As partial decentralization dominates when $\pi \geq \pi_2^{PD}$, $\pi_1^{PD} \leq \hat{\pi}(1) < \pi_2^{PD}$.

**Welfare comparison when $\theta \leq \theta_0$**

When $\theta = \theta_0$,

$$\left.\frac{\partial \mathbb{E}W^{PD}(\pi)}{\partial \pi}\right|_{\pi=0} = \frac{L[4(L+1)B + (L^2 - 5L - 2)b]}{2(L+1)^3} > 0$$

because $L^2 - 5L - 2 > -8.25$ and $L \geq \underline{L} \geq 3$. Also

$$\begin{aligned}
\left.\frac{\partial \mathbb{E}W^{FD}(\pi, \theta)}{\partial \pi}\right|_{\pi=0} &= \frac{1}{b}\int_0^{\frac{\theta_0 b}{1+\theta_0}} \left[B - b + L\frac{c}{\theta_0} + (1-L)T(c,\theta_0)\right] dc \\
&< \frac{1}{b}\left\{\int_0^{\frac{\theta_0 b}{1+\theta_0}} [B-b]dc + \int_0^{\frac{\theta_0 b}{1+\theta_0}} T(c,\theta_0)dc\right\} \\
&< \frac{1}{b}\int_0^{\frac{\theta_0 b}{1+\theta_0}} [B-b]dc + \frac{\theta_0 b}{2(1+\theta_0)} = \frac{(2B-b)L}{(1+L)^2},
\end{aligned}$$

because $T(c, \theta_0) \leq b$. Hence,

$$\left.\left(\frac{\partial \mathbb{E}W^{PD}(\pi)}{\partial \pi} - \frac{\partial \mathbb{E}W^{FD}(\pi, \theta_0)}{\partial \pi}\right)\right|_{\pi=0} > \frac{bL^2(L-3)}{2(L+1)^3} > 0.$$

Next, we prove that, when $\pi \in [0, \pi_1^{PD}]$, $\mathbb{E}W^{PD}(\pi) > \mathbb{E}W^{FD}(\pi, \theta_0)$. To do so, we first find an upper-bound for the expected welfare under full decentralization. If $\pi_1^{PD} \geq \pi_2^{FD}$,

$$\begin{aligned}
\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0) = \kappa + \frac{1}{b}\Bigg\{&\int_0^{c_{NR}^{FD}(\pi)} \left[\pi_1^{PD}B - c\right] dc + \int_0^{c_2} \left[\left(1-\pi_1^{PD}\right)^L \left(b - \frac{c}{\theta_0}\right)\right] dc \\
&+ \int_0^{c_1} \left[\left(1-\pi_1^{PD}\right)\left(1 - \left(1-\pi_1^{PD}\right)^{L-1}\right)(b - T(c,\theta_0))\right] dc\Bigg\}
\end{aligned}$$

where $c_{NR}^{FD} = \pi_1^{PD} B$ and $c_2 = \theta_0 b$. So

$$\begin{aligned}
\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0) &= \kappa + \frac{bB^2}{2[b + L(B-b)]^2} + \frac{bL(1 - \frac{b}{b+L(B-b)})^L}{1+L^2} \\
&\quad + \frac{1}{b}\int_0^{c_1}\left[\left(1-\pi_1^{PD}\right)\left(1-\left(1-\pi_1^{PD}\right)^{L-1}\right)(b - T(c,\theta_0))\right]dc \\
&< \kappa + \frac{bB^2}{2[b + L(B-b)]^2} + \frac{bL(1 - \frac{b}{b+L(B-b)})^L}{1+L^2} \\
&\quad + \frac{1}{b}\int_0^{c_1}\left[\left(1-\pi_1^{PD}\right)\left(1-\left(1-\pi_1^{PD}\right)^{L-1}\right)\left(b - \frac{c}{\theta_0}\right)\right]dc \\
&= \kappa + \frac{bB^2}{2[b + L(B-b)]^2} + \frac{bL\left(1 - \frac{b}{b+L(B-b)}\right)}{1+L^2} \equiv A.
\end{aligned}$$

If $\pi_1^{PD} < \pi_2^{FD}$,

$$\begin{aligned}
\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0) = \kappa + \frac{1}{b}\Bigg\{ & \int_0^{c_{NAR}^{FD}(\pi)} \left[\pi_1^{PD} B - c + \left(1-\pi_1^{PD}\right)^L \left(b - \frac{c}{\theta_0}\right)\right] dc \\
& + \int_0^{c_1} \left[\left(1-\pi_1^{PD}\right)\left(1-\left(1-\pi_1^{PD}\right)^{L-1}\right)\left(b - T(c,\theta_0)\right)\right] dc \Bigg\} \\
& < \kappa + \frac{1}{b}\Bigg\{ \int_0^{c_2} \left[\pi_1^{PD} B - c + \left(1-\pi_1^{PD}\right)^L \left(b - \frac{c}{\theta_0}\right)\right] dc \\
& \int_0^{c_1} \left[\left(1-\pi_1^{PD}\right)\left(1-\left(1-\pi_1^{PD}\right)^{L-1}\right)\left(b - T(c,\theta_0)\right)\right] dc \Bigg\} < A.
\end{aligned}$$

Hence $A$ is one upper bound for $\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0)$. Now, define

$$\Gamma_{\theta_0}(\pi_1^{PD}) \quad \equiv \mathbb{E}W^{PD}(0) + \left.\frac{\partial \mathbb{E}W^{PD}(\pi)}{\partial \pi}\right|_{\pi=0} .\pi_1^{PD} \quad = \kappa - \frac{bL[bL(L+5) - 2B(L+1)(L+2)]}{2(1+L)^3[BL + b(L-1)]}.$$

Finally, we evaluate

$$\Gamma_{\theta_0}(\pi_1^{PD}) - A = H(b,B,L) \equiv \ \tfrac{1}{2}b\left[\tfrac{-B^2}{[b+L(B-b)]^2} - \tfrac{L[bL(L+5)-2B(L+1)(L+2)]}{(1+L)^3[BL+b(L-1)]} - \tfrac{2L\left(1-\frac{b}{b+L(B-b)}\right)}{1+L^2}\right].$$

As $L \geq \underline{L}$, we can write $L = 2 + n + \frac{B}{B-b}$, for $n \geq 1$. The solution to the equation $H(b,B,L) = 0$ in $B$ can be expressed in the form $B = \alpha(n).b$. Although it is not possible to obtain a close form expression for $\alpha(n)$, the following figure depicts the curve $\alpha(n)$, when $n \in \{1, ..., 100\}$.

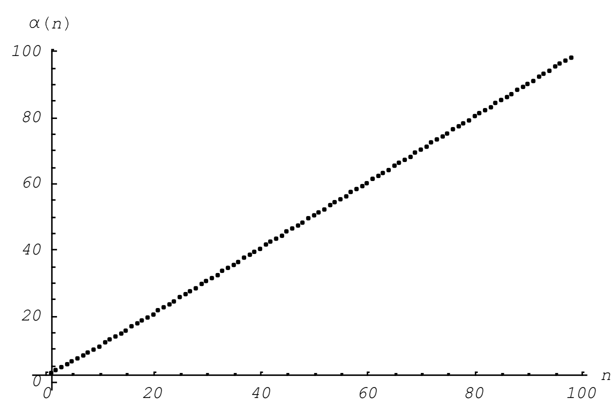


Figure 9: The curve $\alpha(n)$

As $\alpha(1) = 2.6817$ and $B/2 < b$, $H(b, B, L) > 0$ for any $n \in \{1, ..., 100\}$.[39] Therefore, we conclude that $\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0) < \Gamma_0(\pi_1^{PD})$.

These results enable us to assert that, when $\pi \in [0, \pi_1^{PD}]$, $\mathbb{E}W^{FD}(\pi, \theta_0)$ lies below $\Gamma_{\theta_0}(\pi)$, a straight line that takes the values $\mathbb{E}W^{PD}(0)$ and $\Gamma_{\theta_0}(\pi_1^{PD})$ when $\pi = 0$ and $\pi_1^{PD}$, respectively. In fact, $\Gamma_{\theta_0}(\pi)$ is the tangent line to $\mathbb{E}W^{PD}(\pi, )$ at $\pi = 0$. Hence, as $\mathbb{E}W^{PD}(\pi)$ is convex, $\mathbb{E}W^{FD}(\pi, \theta_0) < \mathbb{E}W^{PD}(\pi)$ everywhere on $\pi \in [0, \pi_1^{PD}]$. We depict this result in the following figure.

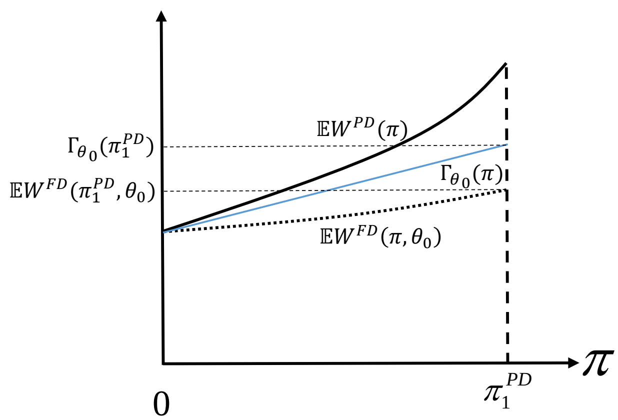


**Figure 10: Expected welfare when $\theta = \theta_0$ and $\pi \in [0, \pi_1^{PD}]$**

Moreover, when $\pi \in (\pi_1^{PD}, 1]$, $\mathbb{E}W^{FD}(\pi, \theta_0)$ cannot cross $\mathbb{E}W^{PD}(\pi)$. If this were the case, the former curve would intersect the latter from below, in its linear part. But then, by convexity, it could not converge towards $\mathbb{E}W^{PD}(\pi)$ from below at $\pi = 1$.

Hence, when $\theta = \theta_0$, $\mathbb{E}W^{FD}(\pi, \theta_0) < \mathbb{E}W^{PD}(\pi)$, except at $\pi = 0$ when they coincide. Therefore, as $\mathbb{E}W^{FD}(\pi, \theta)$ increases with $\theta$, $\mathbb{E}W^{FD}(\pi, \theta) < \mathbb{E}W^{PD}(\pi)$ for all values of $\pi$ whenever $\theta \leq \theta_0$.

### Welfare comparison when $\theta \in [\theta_0, \theta^*]$

We will prove, using a series of geometrical arguments, that when $\theta \in [\theta_0, \theta^*]$, $\mathbb{E}W^{FD}(\pi, \theta)$ intersects $\mathbb{E}W^{PD}(\pi)$ exactly once.

Assume that we can find $\theta = \theta_0 + \epsilon$ such that

$$\mathbb{E}W^{FD}(0, \theta_0 + \epsilon) > \mathbb{E}W^{PD}(0) \text{ and } \mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0 + \epsilon) < \Gamma_{\theta_0}(\pi_1^{PD}) < \mathbb{E}W^{PD}(\pi_1^{PD}).$$

Clearly, $\mathbb{E}W^{FD}(\pi, \theta_0+\epsilon)$ intersects the line $\Gamma_{\theta_0}(\pi)$ once, from above. Hence, as $\mathbb{E}W^{FD}(\pi, \theta_0+\epsilon)$ and $\mathbb{E}W^{PD}(\pi)$ are both convex in $\pi$, the former has to cross the latter only once, from above. This argument can be replicated until $\theta = \theta_1 > \theta_0$, which is implicitly defined by

$$\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_1) = \Gamma_{\theta_0}(\pi_1^{PD}).$$

[39]This result also holds for $n > 100$.

When $\theta = \theta_1$, there can be two possibilities. When $\pi$ converges to $\pi_1^{PD}$, either (i) $\mathbb{E}W^{FD}(\pi, \theta_1)$ converges to $\Gamma_{\theta_0}(\pi_1^{PD})$ from below, in which case the previous intersection argument applies, or (ii) this convergence is from above. This means that

$$\left.\frac{\partial \mathbb{E}W^{FD}(\pi, \theta_1)}{\partial \pi}\right|_{\pi=\pi_1^{PD}} < \text{slope}\, \Gamma_{\theta_0}(\pi).$$

Hence, $\mathbb{E}W^{FD}(\pi, \theta_1)$ has to cross $\mathbb{E}W^{PD}(\pi)$ from above, only once. So, for any $\theta \in [\theta_0, \theta_1]$, $\mathbb{E}W^{FD}(\pi, \theta)$ intersects $\mathbb{E}W^{PD}(\pi)$ exactly once.

Now, let's define $\Gamma_{\theta_1}(\pi)$, a new line that has the same slope than $\Gamma_{\theta_0}(\pi)$ and is characterized by $\Gamma_{\theta_1}(0) = \mathbb{E}W^{FD}(0, \theta_1)$. By construction, $\Gamma_{\theta_1}(\pi_1^{PD}) > \mathbb{E}W^{FD}(\pi_1^{PD}, \theta_1)$. Now assume that we can find $\theta = \theta_1 + \epsilon'$ such that $\mathbb{E}W^{FD}(0, \theta_1 + \epsilon') > \Gamma_{\theta_1}(0)$ and $\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_1 + \epsilon') < \Gamma_{\theta_1}(\pi_1^{PD})$. Clearly, $\mathbb{E}W^{FD}(\pi, \theta_1 + \epsilon')$ crosses $\Gamma_{\theta_1}(\pi)$ only once, from above. Hence, as $\mathbb{E}W^{FD}(\pi, \theta_1 + \epsilon)$ and $\mathbb{E}W^{PD}(\pi)$ are both convex in $\pi$, the former has to cross the latter only once, from above. This argument can be replicated until $\theta = \theta_2 > \theta_1$, which is implicitly defined by

$$\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_2) = \Gamma_{\theta_1}(\pi_1^{PD}).$$

Now construct an increasing sequence $\theta_n$, $\forall n \geq 0$ defined by

$$\mathbb{E}W^{FD}(0, \theta_{n+1}) > \Gamma_{\theta_n}(0),$$

and

$$\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_{n+1}) = \Gamma_{\theta_n}(\pi_1^{PD}).$$

This sequence can behave in two different ways: either there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $\theta_n \geq \theta^*$; or $\theta_n < \theta^*$ for all $n \in \mathbb{N}$. In the first case, the previous geometric arguments apply, and thus we can assert that $\mathbb{E}W^{FD}(\pi, \theta)$ crosses $\mathbb{E}W^{PD}(\pi)$ only once, from above, when $\pi \in [0, \pi_1^{PD}]$.

Let's prove that the second case can be ruled out. Assume that $\theta_n < \theta^*$ for all $n \in \mathbb{N}$. As the sequence $\theta_n$ is increasing and bounded (by $\theta = 1$), it must converge. Denote this limit by $\overline{\theta}$. As the functions that define the sequence are continuous, they also converge towards $\mathbb{E}W^{FD}(\pi, \overline{\theta})$ and $\Gamma_{\overline{\theta}}(\pi)$. These limit functions have to satisfy

$$\mathbb{E}W^{FD}(0, \overline{\theta}) = \Gamma_{\overline{\theta}}(0),$$

and

$$\mathbb{E}W^{FD}(\pi_1^{PD}, \overline{\theta}) = \Gamma_{\overline{\theta}}(\pi_1^{PD}).$$

Define $\overline{\overline{\theta}} = \overline{\theta} + \epsilon$. Clearly, as $\mathbb{E}W^{FD}(\pi, \theta)$ increases with $\theta$, $\mathbb{E}W^{FD}(0, \overline{\overline{\theta}}) > \Gamma_{\overline{\theta}}(0)$. But if this were the case, we can construct $\Gamma_{\overline{\overline{\theta}}}(\pi) > \Gamma_{\overline{\theta}}(\pi)$, which is a contradiction. Hence, the sequence $\theta_n$ does not converge to a value $\overline{\theta} < \theta^*$.

We conclude that, when $\theta \in [\theta_0, \theta^*]$, $\mathbb{E}W^{FD}(\pi, \theta)$ intersects $\mathbb{E}W^{PD}(\pi)$ exactly once, from above, at $\hat{\pi}(\theta)$. Applying the Implicit Function Theorem, we can show that $\partial \hat{\pi}(\theta)/\partial \theta > 0$. ■