Regional state capacity and the optimal degree of fiscal decentralization

Andrés Bellofatto†  Martin Besfamille‡

November 3, 2014

Abstract

This paper presents a model featuring a central government and local authorities, the latter being characterized by the same level of administrative and fiscal capacity. We analyze two fiscal regimes are analyzed. Under partial decentralization, regional governments invest in a local public project. Depending upon their administrative capacity, projects may remain incomplete. Under this institutional regime, regional governments can rely on central bailouts to finalize them, and thus face soft budget constraints. Anticipating this, they may inefficiently over invest. Under full decentralization, regional governments are never rescued, and so face hard budget constraints. This generates the opposite type of inefficiency: regional governments under invest. The first goal of the paper is to assess which regime dominates. The second goal is to investigate how different levels of regional state capacity affect the normative comparison between regimes. As expected, when the regional fiscal capacity is low, partial decentralization dominates; otherwise full decentralization may dominate. But, contrary to the common wisdom, we find that full decentralization is the preferred regime when regional administrative capacity is low.

Keywords: Fiscal federalism - Partial and full fiscal decentralization - Soft and hard budget constraints - State capacity.

JEL Classification: D82 - H77.

*We thank L. Arozamena, P. Dal Bo, A. de Paula, J. Dubra, E. Espino, I. Fainmesser, N. Figueroa, F. Forges, M. González Eiras, B. Guimaraes, C. Hevia, B. Knight, J. C. Hallak, B. Lockwood, A. Manelli, E. Mattos, P. Mongin, A. Neumeyer, E. Ornelas, H. Piffano, F. Piguillem, A. Porto, E. Rezk, C. Ruzzier, R. Serrano, W. Sosa Escudero, M. Tommasi, F. Weinschelbaum, attendants to the Annual Congress of the Association for Public Economic Theory (Seattle, 2014), Latin American Meeting of the Econometric Society (Santiago de Chile, 2011), Jornadas Internacionales de Finanzas Públicas (Córdoba, 2010) and Reunión Anual de la Asociación Argentina de Económica Política (Buenos Aires, 2010), and seminar participants at Brown, EESP-FGV, San Andrés and Pontificia Universidad Católica de Chile for useful comments and suggestions. We are also grateful to S. Briole for his excellent research assistance. The usual disclaimer applies.

†Tepper School of Business, Carnegie Mellon University. E-mail: andresb@cmu.edu
‡Corresponding author. Instituto de Economía, Pontificia Universidad Católica de Chile. E-mail: mbesfamille@uc.cl
1 Introduction

In many countries, tax and expenditure assignments to subnational or regional governments are not balanced. In particular, regional governments are often in charge of delivering local public services, but cannot raise the needed revenues to finance these expenditures. This situation is known as “vertical fiscal imbalance”, and is rather prevalent in advanced and in developing economies. According to Eyraud and Lusinyan (2013), across OECD countries, the average share of subnational government expenditure not financed through own revenues has been 40 percent between 1995 and 2005, with some countries having much higher shares than this (e.g., Belgium and Mexico, with 60 and 83 percent, respectively). Corbacho et al. (2013) document that vertical fiscal imbalances in Latin America are the highest among the developing nations.

In practice, vertical fiscal imbalances are often solved by centrally provided transfers to regional governments (Boadway and Shah (2007)). In such a institutional setting, defined by Brueckner (2009) as “partial decentralization”, regional governments may face soft budget constraints. Wildasin (1997) and Goodspeed (2002) show that the willingness of the central government to bail out regional governments creates a negative externality if, as is usual, the cost of a bailout is met through increases in national taxes. As other regions will partially finance one region’s bailout, this induces excessive expending or borrowing initially. Among others, Pettersson-Lidbom (2010) confirms this theoretical result, by estimating than, between 1979 and 1992, Swedish local governments increase their debt by more than 20 percent when they expected to receive future bailouts.

In the economic and policy literature, it is widely accepted that, because of this common-pool fiscal externality, regional governments have to face hard budget constraints which, according to Weingast (2009), provide “local political officials with incentives for prudent fiscal management of their jurisdiction.” Indeed, much of the literature, as referenced in Rodden et al. (2003) and Oates (2005), is concerned with the design of institutional mechanisms aimed to harden regional government’s budget constraints. One such mechanism that has attracted quite a lot of attention is complete decentralization of taxing powers to subnational governments. Qian and Roland (1998) argue that full fiscal decentralization creates tax competition, which in turn raises the perceived marginal cost of public funds at the regional level. This increase endogenously “hardens” regional governments’ budget constraints, and makes them to be reluctant to bailout inefficient projects undertaken by public enterprises. This formal result is at the basis of the idea that “federalism preserves markets” (see Weingast (1995, 2009), Montinola et al. (1995) and Qian and Weingast (1997), among others), and has also been adopted by international organizations as one of the main arguments to actively promote fiscal decentralization reforms (World Bank, (2000)).

But this conventional wisdom has been challenged on two different grounds. From a theoretical point of view, Besfamille and Lockwood (2008) show that Qian and Roland’s

---

1 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.”
(1998) results were somewhat restrictive. In a model where regional governments exert effort to carry out a local project, they prove that, under more general parameter configurations than those adopted by Qian and Roland, hard budget constraints can be inefficient because they trigger an excessive level of effort, which in turn generates inefficient under provision of projects. This implies that a normative comparison between partial and full decentralization, along the abovementioned trade-offs, deserves to be undertaken.

Other authors, like Prud’homme (1995) and Bardhan and Mookherjee (2006), point out that this pro-decentralization literature ignores relevant local, institutional characteristics. One of these characteristics is the level of state capacity, defined by Besley and Persson (2010) as the “state’s ability to implement a range of policies”. Many qualitative reviews of decentralization reforms (e.g., Bird (1995), Litvack et al. (1998)) and more recent quantitative evaluations of such processes (Loayza et al. (2014)) show that when regional governments lack these abilities, the benefits accruing from a decentralized form of government are far from granted. Moreover, Cabrero and Martínez-Vazquez (2000) even affirm that an adequate administrative capacity at the subnational level is a prerequisite for any successful decentralization reform. These assertions seem to suggest that one cannot compare partial and full decentralization without explicitly incorporating regional state capacities into the model.

To our knowledge, there is no contribution to the local public finance literature that addresses these issues theoretically. The goal of this paper is precisely to build a model that allows to compare, from a normative point of view, a partially and a fully decentralized regime, and to determine how this comparison evolves with the level of regional state capacity.

In our model, regional governments decide whether to provide a discrete, local public good or project. The project’s initial cost is covered by their financial resources. If the project is initiated, it is carried out by regional bureaucracies, which are characterized by a given level of administrative capacity to fulfill this task. This administrative capacity is formalized as the probability that the project’s execution ends in due time and generates a social benefit greater than the initial cost. With the complementary probability, the project’s construction lasts more, and needs a second round of financing to be completed. In this last case, the project also generates a social benefit to the region, but, due to its delay, lower than when the project is carried out on time. The model is symmetric: all governments possess the same level of state capacity, and all projects’ costs and benefits are identical. Despite this, depending upon the project’s execution length, outcomes can be different ex post.

In the partially decentralized regime, no regional government has tax revenues to refinance its incomplete project. But the central government can bailout any region, via a uniform national tax on local capital, which is imperfectly mobile. We show central bailouts do not provide appropriate incentives for efficient initial investment. So, partial decentralization may lead to inefficient overprovision of public projects.

Next, we analyze the fully decentralized regime, where regional governments have to

---

2 Although Rodden and Rose-Ackerman (1997) also argue against the predictions of Montinola et al. (1995), they do not criticize neither the logic nor the results of this model per se, but rather the fact that its assumptions are unrealistic.
refinance incomplete projects through a tax on capital invested in their jurisdiction, in a context of tax competition. In this case, we assume that regional governments can collect only a fraction of their potential tax base. This fraction describes the regional fiscal capacity, a second dimension of state capability. First, we obtain the refinancing and taxing equilibria. When all regions face an incomplete project, they refinance it with the same tax rate but, if it is the case, bearing an implicit cost due to imperfect fiscal capacity. When there is at least one region that does not need to refinance, the other regional governments are compelled to raise taxes that end-up being distortionary because they trigger costly capital outflows. Therefore, for some parameter values of the model, regional governments may decide not to refinance their incomplete project: an endogenous hard budget constraint, as in Qian and Roland (1998). Moving back to the initiation decision, we find the opposite inefficiency than under partial decentralization. Imperfect regional fiscal capacity, distortionary taxation or endogenous hard budget constraints may underincentivize regional governments to initially invest, and thus the fully decentralized regime can generate inefficient underprovision of projects.

Then, for different pairs of administrative and fiscal capacity levels, we compare the expected welfare of both institutional regimes. The main results are the following. When the level of regional fiscal capacity that prevails in the federation is low, refinancing incomplete projects under full decentralization turns out to be too costly. Therefore, regardless of the level of regional administrative capacity, partial decentralization dominates. For higher levels of the regional fiscal capacity, full decentralization dominates, provided the level of regional administrative capability is low. Although this entails that projects remain incomplete almost surely, it also implies that distortionary refinancing under full decentralization is less likely. Thus, full decentralization’s distortions cost less, in welfare terms, than inefficient investments made under partial decentralization.

Finally, we evaluate the model’s robustness in two different ways. First, we undertake some comparative statics exercises. When the regional capital stock (the project’s highest benefit) increases, full (partial) decentralization dominates more often. When the number of regions increases, each regime starts to be preferred in a parameter configuration of the model where the other one used to be the best. Second, we extend the model to include distortionary national taxation. This implies ipso facto that the central government does not bailout all incomplete projects, which corresponds to another endogenous hard budget constraint. We show that this extension modifies significantly some outcomes under partial decentralization, in particular less projects are inefficiently initiated, but inefficient underinvestment may also emerge. All this decreases the expected welfare under partial decentralization, and thus full decentralization is preferred more often.

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 outcomes under partial decentralization, whereas Section 5 studies the equilibrium under full decentralization. Section 6 contains the normative comparison between partial and full decentralization, and some results on its robustness. Section 7 discusses related literature. Finally, Section 8 concludes. The main proofs are shown in the Appendix, and supplementary material appears in an Online Appendix.
2 The model

2.1 Preliminaries

The model has three periods \( t = \{1, 2, 3\} \) and \( L \geq 2 \) regions. Each region \( \ell \in \{1, \ldots, L\} \) has a continuum of measure 1 of risk-neutral residents, each of whom has an endowment \( \kappa \) of capital. In the last period, each resident derives utility from consumption of a numéraire good. This good is produced in every region from the capital input by competitive firms using a constant-returns technology, where 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 one region that exports \( f \) units of capital to other region(s) incurs a mobility cost \( f^2/2 \). As we explain below, residents may also benefit from a discrete 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, i.e., they maximize the sum (or average) of utilities of their jurisdictions’ residents over the two last periods. For simplicity, we assume that there is no discounting of future payoffs.

2.2 Timing

The order of events, and other relevant features of the model in more detail, are as follows.

![Figure 1: Timing of the model](image)

At \( t = 1 \), a political body (e.g., a Congress) chooses the institutional regime that will rule all fiscal interactions between the central and regional governments.

At the beginning of \( t = 2 \), Nature draws projects’ cost \( c \) according to a strictly positive probability density function \( h(c) \) on \([0, b]\). Then, regional governments choose whether to initiate a project in their region. We denote this decision 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, which possesses a level of administrative competence, or capacity, to fulfill this task. Given the level of administrative capacity, a project yields different outcomes. With an exogenous probability $\pi \in [0, 1]$, which we interpret as the regional level of administrative capacity, a project generates, at the end of the current period, a non-financial benefit $B > 0$ for all residents of the region. On the other hand, with probability $(1 - \pi)$, the project remains incomplete and yields no benefit during this period. These realized outcomes are observed by everybody.

At $t = 3$, central or regional governments, depending on the institutional regime in place, decide whether to shutdown or continue incomplete projects. In the last case, to be completed, a project requires an additional input of $c$ of the consumption good.

Under partial decentralization ($PD$), the central government decides on refinancing incomplete projects, through a uniform tax $\tau$ on capital, collected by the national tax authority. Uniformity implies that, for (non-modeled) constitutional reasons, the central government can neither set different tax rates contingent on which regional government has asked for additional funds, nor make side-payments to any specific regional government.

Under full decentralization ($FD$), each regional government decides whether to refinance its incomplete project, using a per unit tax on capital invested in its region, at a rate $\tau_\ell$. After taxes are set, capital owners decide to invest in the region(s) with the highest net return(s) and production takes place. Then, regional governments raise their taxes. Due to costly tax enforcement/collection or to regional governments’ inability to compel individuals to pay their due taxes, regional net fiscal revenues end up being a fraction $\theta \in [0, 1]$ of their potential tax base. From now on, the parameter $\theta$ will characterize regions’ fiscal capacity, the second dimension of the state capabilities.

Upon completion, the project generates a non-financial benefit $b$ for all residents of the region. We assume that $B/2 \leq b < B$.

### 2.3 Discussion

Some features of the model deserve further comments. First, we have adopted Mann’s (1984) general definition of state capacity as the infrastructural power of the state to enforce policy within its territory, and followed Snyder (2001) and Ziblatt (2008) to apply this concept to regional governments. Specifically, we focus on the administrative and extractive dimensions of state capacity, as defined by Hanson and Sigman (2013). First, we consider the probability $\pi$ to be an appropriate proxy of the region’s capacity to carry out projects in due time. As $B > b$, the higher is this probability, the higher is the regional administrative capacity.

---

3 Patil et al. (2013) document that public projects’ delays are pervasive in Indian states, and that their main cause is administrative problems that arise during the land acquisition process.

4 The difference between $B$ and $b$ reflects the negative impact of delaying the project’s construction on the regional welfare. For example, an incomplete park may affect pedestrian movement. But the benefit’s decrease due to the delay is lower than 50 percent of its value.

5 This measure of administrative capacity is somewhat related to the “infrastructure reform” ability, quantified by Fortin (2010) in her study of state capacity in post-communist countries.
Regarding the second dimension of the regional state capacity, our characterization of fiscal capability follows Arbetman-Rabinowitz et al. (2012) and Gadenne and Singhal (2013).

Second, we consider a discrete, regional public project, instead of a continuous public good, as is common in most of the literature on tax competition. This has the following consequences. The project’s indivisibility fixes the type of competition between regions. As in Wildasin (1988), here regions compete first in refinancing decisions, and then taxes are set accordingly, in a context of imperfect capital mobility. Moreover, this assumption, combined with the characterization of administrative competence as a probability of completing projects on time, is a simple way to analyze, via refinancing decisions, the interaction between levels of regional state capacity and different intergovernmental fiscal arrangements.

Third, we assume that, across regions, governments and projects are ex ante and interim identical. 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. But projects can be different ex post. Indeed, at the end of period 2, when \( \pi \in (0,1) \), some regions that have initially invested end up with complete projects, while the other face incomplete ones, that might or might not be finalized at the end of period 3. In the Conclusions, we comment more on this important feature of this model.

Finally, we assume that \( c \leq b \). When \( c > b \), both institutional regimes generate the same outcome. Therefore, we prefer to consider a parameter configuration of the model under which the institutional comparison between partial and full decentralization is indeed relevant.

3 First best outcomes

In order to have a benchmark, 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, as individual utilities are linear in income and the planner maximizes the sum of utilities, the refinancing decision in any region is independent of his choice in any other region. 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 again a separable problem between regions. Knowing the cost \( c \), he initiates projects provided their expected benefit

---

6To the best of our knowledge, the unique contribution to the local public finance literature that deals with discrete projects is Cremer et al. (1997). In a two-region model, the authors compare the optimal and the (Nash) equilibrium provision of projects that can be used by residents in both regions. Although they do not assume distortionary local taxation to finance the project (as we do), they find equilibrium projects' under provision, as expected in a model with spillovers.

7Akai and Sato (2008) and Köthenbürger (2011) also analyze models with this timing, but instead consider income or wage taxation.
is higher than their expected cost (which includes a possible second round of financing). Let’s denote by
\[ c^*(\pi) \equiv \frac{\pi B + (1 - \pi)b}{2 - \pi} \]
the cost that makes the expected, net regional welfare from initiating a project equal to zero. Hence, if \( c \leq c^*(\pi) \), initial investment is efficient in any region \( \ell \). When \( c > c^*(\pi) \), not investing is the optimal decision.
For a given configuration of parameters \((b, B)\), efficient investment decisions are depicted in the following figure. There, any point in the \((\pi, c)\) plane represents a project, that costs \( c \), in a region with administrative capacity \( \pi \).

\[ \begin{array}{c}
\text{Figure 2: First best outcomes} \\
\end{array} \]

When \( 0 \leq \pi \leq \pi^* \equiv b/B \), \( 0 < c^*(\pi) \leq b \). Therefore, there exists a non-empty area \( NI^* \), delimited by the thick curve \( c^*(\pi) \), where projects are optimally not initiated. Bellow this curve, in the area \( I^* \), it is efficient to undertake all projects therein. Ceteris paribus, as \( \pi \) increases, the projects’ expected, net benefit increases. Thus \( c^*(\pi) \) increases as well. Therefore, when \( \pi^* < \pi \leq 1 \), \( c^*(\pi) > b \). As regional administrative capacities are relatively high, all projects are efficiently initiated.

### 4 Partial decentralization

Under this institutional framework, again incomplete projects are always refinanced by the central government because \( c \leq b \). Therefore, regional welfares may be a priori interdependent, via the central government’s budget constraint. Given this, there is a simultaneous game between regions at the beginning of the second period, when they decide on initial investment.

At this time, region \( \ell \)’s expected welfare is

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

(1)
where $i_m$ is the profile of investment decisions chosen by regions $m \neq \ell$, $\mathbb{I}_{\{i_i = I\}}$ is an indicator function that takes the value of 1 if region $\ell$ has initiated the project and 0 otherwise, and $\tau^e$ is the expected tax.

What is the value of the expected tax $\tau^e$? At the end of the second period, all projects’ outcomes are realized. Let $\omega$ be a profile of outcomes, and let’s denote by $\#B(\omega)$ 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 $L - \#B(\omega)$ 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

$$\tau_\omega \cdot L\kappa = [L - \#B(\omega)]c.$$  

Therefore, when deciding on initial investment before the realization of projects’ outcomes, each region faces the expected tax $\tau^e$, which satisfies

$$\tau^e \cdot L\kappa = \left( \sum_{\ell=1}^{L} \mathbb{I}_{\{i_i = I\}}(1 - \pi) \right) \cdot c, \quad (2)$$

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

$$\mathbb{E}W^PD_\ell(i_\ell, i_m) = \kappa + \mathbb{I}_{\{i_i = I\}} \left[ \pi B + (1 - \pi)(b - \frac{c}{L}) - c \right] - \left( \sum_{m \neq \ell} \mathbb{I}_{\{i_m = I\}}(1 - \pi) \right) \cdot \frac{c}{L} \quad (3)$$

By inspection of (3), the effect of $i_\ell$ on $\mathbb{E}W^PD_\ell$ - measured by the term in square brackets - is independent $i_m$. 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 expected, net 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**

If $c \leq c^{PD}(\pi)$, initial investment occurs. Otherwise, there is no initial investment in any region.

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, equilibrium outcomes can be inefficient. When this is so, initial investments are made 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 3: Partial decentralization outcomes](image)

When $0 \leq \pi < \pi_1^{PD} \equiv \frac{b}{L(B-b)+b}$, $c^{PD}(\pi) < b$. Therefore, there exists a non-empty area $NI^{PD}$, delimited by the thick curve representing $c^{PD}(\pi)$, where projects are efficiently not initiated in any region. Below this curve, in the area denoted by $I^{PD}$, all projects are initiated. There, in the white area, regional decisions are optimal and so, in equilibrium, each region expects to contribute an amount equal to the cost of refinancing by itself its incomplete project. Thus, interim regional expected welfares coincide with the first best level. But this is not always the case. In the shaded area, when $c \in [c^*(\pi), c^{PD}(\pi)]$, the inefficient investment decision is adopted in equilibrium by all regions.

As $\pi$ increases, $c^{PD}(\pi)$ increases as well. Thus, 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. Still, projects in the shaded area, whose cost satisfies $c^*(\pi) \leq c \leq b$, are inefficiently initiated.

Finally, when $\pi_2^{PD} < \pi \leq 1$, $b < c^*(\pi) < c^{PD}(\pi)$. This implies that 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 case of full decentralization, when regional governments decide not only about initial investment, but also if they refinance their incomplete project, taxing
capital employed in their region, in a context of imperfect capital mobility. Therefore, there is now a three-stage game between regions, where they play simultaneously at each stage. As before, the game is solved backwards.

5.1 Capital flows

Given a profile of tax rates \( \tau = \{\tau_1, \ldots, \tau_L\} \) set by all regions, a household resident in region \( \ell \) decides where to invest its capital endowment. Let’s denote by \( f_{\ell m} \) the amount of capital that this household invests in a region \( m \neq \ell \), and by \( \widetilde{M} \) 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 just 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 has set the minimum tax rate \( \widetilde{\tau}_\ell \), all capital that leaves region \( \ell \) goes there. But if two or more regions have chosen the same tax rate \( \widetilde{\tau}_\ell \), the value of the capital outflow \( f_{\ell \widetilde{m}} \) that goes to each of them is not determined: only the total amount of capital that leaves region \( \ell \) is characterized by the expression shown in the proposition, expression whose intuition is straightforward. Capital leaves region \( \ell \) until the marginal saving in taxes equals 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 \( \ell \): 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, we will see in the next section that these capital inflows will have an important role to play in the determination of the equilibrium tax rates.

5.2 Equilibrium in tax rates

Anticipating households’ portfolio decisions and taking other regional tax rates as given, each regional government sets its tax rate to raise enough resources to refinance its incomplete project, if it has decided to do so.

To obtain the equilibrium tax rates, in the Appendix we derive region \( \ell \)'s reaction function \( \tau_\ell (\tau_m) \), where \( \tau_m \) denotes the profile of tax rates chosen by regions \( m \neq \ell \). We have just seen that, depending upon the whole profile of tax rates, capital may leave region \( \ell \) or may come there, from other regions. Despite these different possibilities, region \( \ell \)'s after-tax, private-good consumption monotonically decreases with \( \tau_\ell \), and so does the regional welfare. Hence, for any profile \( \tau_m \), the tax rate chosen in region \( \ell \) should be the lowest tax rate that,
given its fiscal capacity and the capital invested there, enables the regional government to raise \( c \).

The reaction function \( \tau_\ell (\tau_m) \) is built around the value \( c/\theta \kappa \), which is the tax rate that a regional government would choose in autarchy. Moreover, \( \tau_\ell (\tau_m) \) “decreases” in the following sense.

When the profile \( \tau_m \) is such that all tax rates \( \tau_m < c/\theta \kappa \), region \( \ell \)'s optimal response is to tax strictly above the minimum tax \( \tilde{\tau}_\ell \). 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 in this particular case. Finally, when the profile of tax rates is such that \( \tilde{\tau}_\ell > c/\theta \kappa \) or \( \tilde{\tau}_\ell = c/\theta \kappa \) 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 \( \tilde{\tau}_\ell \). This decision generates an inflow of capital that enables the government to raise sufficient revenues to pursue its incomplete project, moderating the tax burden on its residents.

Another important feature of region \( \ell \)'s reaction function is that it is non-continuous. When \( \tilde{\tau}_\ell \) converges to \( c/\theta \kappa \) from below, \( \tau_\ell (\tau_m) \) converges to this limit from above. But when the distribution of taxes is such that \( \tilde{\tau}_\ell = c/\theta \kappa \) but at least one region \( n \neq \ell \) has chosen \( \tau_n > c/\theta \kappa \), tax undercutting is a profitable deviation in this case. Despite this fact, we can characterize the Nash equilibrium 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 characterized by \( \tilde{\tau}_\ell = c/\theta \kappa \).

When there is at least one region that does not refinance, region \( \ell \) has to set the tax rate

\[
\tau_\ell (0) = \frac{1}{2} \left[ \kappa - \sqrt{\kappa^2 - (4c/\theta)} \right]
\]


to refinance its incomplete project.

Consider the tax competition subgame that emerges when all regions have decided to refinance their incomplete project. The equilibrium tax rate \( \tilde{\tau}_\ell \) increases with the project cost \( c \). 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 \( \tilde{\tau}_\ell \). In this case, in equilibrium, region \( \ell \)'s welfare (net of the initial cost \( c \)) is \( W_\ell^{FD} = \kappa + b - \frac{c}{\theta} \). 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 tax collection’s cost.

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 \( \tau_\ell (0) \) to pursue its ongoing project. Thus, asymmetric taxation emerges as a possible equilibrium, as in
Bucovetsky (1991) and Wilson (1991). The main difference with their result is the following: here, tax differences are not the consequence of an ex ante regional endowment asymmetry, but rather occur because some regions end up with incomplete projects ex post. The tax rate $\tau_\ell(0)$ also decreases with the capital endowment $\kappa$ and the fiscal capacity $\theta$. The intuition is similar in both cases. Ceteris paribus, an increase in $\kappa$ or $\theta$ enlarges the net fiscal revenues, which lowers the tax rate needed to raise the amount $c$. On the other hand, the tax rate increases with $c$ because the fiscal need increases. With this tax rate, the resulting capital outflow is $\sum_{\tilde{m} \in \tilde{M}} \tilde{f}_{\tilde{m} \ell} = \tau_\ell(0)$ and the equilibrium regional welfare (net of the initial cost $c$) ends up being

$$W^{FD}_\ell = (\kappa - \sum_{\tilde{m} \in \tilde{M}} f_{\tilde{m} \ell})(1 - \tau_\ell) + \sum_{\tilde{m} \in \tilde{M}} f_{\tilde{m} \ell} - \frac{1}{2}(\sum_{\tilde{m} \in \tilde{M}} f_{\tilde{m} \ell})^2 + b$$

$$= \kappa + b - T(c, \theta),$$

where

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

measures the total refinancing cost. It comprises the effective refinancing cost $c/\theta$, plus the deadweight loss $[\tau_\ell(0)]^2/2$ of financing project’s continuation 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 $\pi$. This observation will have an important consequence, as we shall see below.

5.3 Refinancing

At the beginning of the third stage, the strategy for region $\ell$ is $r_\ell \in \{R, NR\}$, where $R$ ($NR$) denotes “refinancing” (“not refinancing”). Conditional on $i = (i_1, ..., i_L)$, and given $r_m$, the profile of refinancing decisions chosen by regions $m \neq \ell$, region $\ell$’s welfare is

$$W^{FD}_\ell(r_\ell = R, r_m, i) = \kappa + \mathbb{I}_{\{I, inc\}}\{[b - \mathbb{I}_{\{I, inc,R\}}\frac{c}{\theta} - (1 - \mathbb{I}_{\{I, inc,R\}})T(c, \theta)] - c\}$$

$$W^{FD}_\ell(r_\ell = NR, r_m, i) = \kappa - \mathbb{I}_{\{I, inc\}}c$$

where $\mathbb{I}_{\{I, inc\}}$ is an indicator function that takes the value of 1 if region $\ell$ initiated a project that has remained incomplete at $t = 2$, and 0 otherwise. Also $\mathbb{I}_{\{I, inc,R\}}$ is an indicator function that takes the value of 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 0 otherwise.

Let’s denote by $c_1$ the value of $c$ that makes the total refinancing cost $T(c, \theta)$ equal to the benefit $b$, and by $c_2 > c_1$ the value of $c$ that makes the effective refinancing cost $c/\theta$ equal to the benefit $b$. The following proposition characterizes the Nash equilibria of this subgame.
Proposition 4

When all regions face an incomplete project, they refinance them in equilibrium provided \( c \leq c_2 \). Otherwise, no region refines.

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 there always exists two strictly positive thresholds \( c_1 \) and \( c_2 \) such that \( c \leq c_1 \) (\( c \leq c_2 \)) implies \( T(c, \theta) \leq b \left( c/\theta \leq b \right) \). 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 \( c_1 \leq c \leq c_2 \), the game becomes a standard coordination one, with two Nash equilibria. The first one is trivial: despite the fact that \( c > c_1 \), 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 to not refinance, the other regions also shutdown their project because \( c > c_1 \).

In this case, these regions face an endogenous hard budget constraint due to other regions’ decisions, as in Qian and Roland (1998). As the equilibrium where all regions refinance is strong (Aumann (1959)) and coalition-proof (Berheim et al. (1987)), we select it as the Nash equilibrium. 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 \( \ell \) refines 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. But again, as \( c \leq b \), this is clearly an inefficient outcome.

### 5.4 Project initiation

Anticipating refinancing equilibria, regional governments simultaneously adopt the project initiation decision. Let \( c_{FD}^R(\pi) \left( c_{NAR}^D(\pi) \right) \left[ c_{NR}^D(\pi) \right] \) be the cost that makes the expected, net regional welfare from initiating the project under full decentralization equal to zero, when incomplete projects are always refinanced (when all regions refinance their incomplete project) [when incomplete projects are never refinanced]. In the next proposition, we present the investment equilibria under this regime.

Proposition 5

When \( 0 \leq \pi \leq \pi_1^{FD} \), the symmetric Nash equilibria under full decentralization is as follows. If \( c \leq c_{FD}^R(\pi) \), initial investment occurs in all regions. Otherwise, there is no initial investment in equilibrium in any region.

When \( \pi_1^{FD} < \pi \leq \pi_2^{FD} \), the symmetric Nash equilibria under full decentralization is as follows. If \( c \leq c_{NAR}^D(\pi) \), initial investment occurs in all regions. Otherwise, there is no initial investment in equilibrium in any region.
When $\pi_2^{FD} < \pi \leq \pi_3^{FD}$, the symmetric Nash equilibria under full decentralization is as follows. If $c \leq c_{NR}^{FD}(\pi)$, initial investment occurs in all regions. Otherwise, there is no initial investment in equilibrium in any region.

When $\pi_3^{FD} < \pi \leq 1$, the symmetric Nash equilibria under full decentralization is that investment occurs 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 form of the inefficiencies that emerge under this institutional regime, as follows.

**Corollary 2**

Under full decentralization, equilibrium outcomes can be inefficient. If this is so, either (i) initial investments are not made in equilibrium when it is efficient to do so; (ii) initial investments are made in equilibrium and are efficient, but incomplete projects are not refinanced when it is efficient to do so; or (iii) initial investments are made in equilibrium and are efficient, but incomplete projects are refinanced in a distortionary way.

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

![Figure 4: Full decentralization outcomes](image)

In the non-empty area $NI^{FD}$, delimited from below by the thick curves representing the three abovementioned cost thresholds, projects are never initiated. In the complementary area, denoted by $I^{FD}$, all projects are initiated.

When $0 \leq \pi \leq \pi_1^{FD}$, two different types of inefficiency emerge. First, in the blue 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 $c^F_R(\pi) < c^*(\pi)$, the condition for project initiation is stricter
with full decentralization than for the social planner. Therefore, in the shaded area when $c \in [c^F_R(\pi), c^*(\pi)]$, investments are not initiated in equilibrium, despite the fact that it is efficient
to do so. Underinvestment is due to i) distortionary refinancing, when $c \in [c^F_R(\pi), c_1]$, and
ii) endogenous hard budget constraints, when $c \in [c_1, c^*(\pi)]$. The former finding is analog 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.

When $\pi^F_1 < \pi \leq \pi^F_2$, a new type of inefficiency emerges. In the yellow area, when
$c \in [c_1, c_2]$, projects are initiated but only refinanced when all regions do so. Again, there is
inefficient underinvestment when $c \in [c^F_N(\pi), c^*(\pi)]$, but only caused by endogenous hard
budget constraints.

When $\pi^F_2 < \pi \leq \pi^F_3$, the last type of inefficiency emerges. In the green area when
$c \in [c_2, c^F_{NRI}]$, projects are initiated but never refinanced. There is still a shaded area where,because of endogenous hard budget constraints, projects are inefficiently not initiated. The
figure illustrates that efficient decisions are only adopted in the area where projects
are not initiated. In the remainder areas, either the projects’ initiation and continuation
decisions are downwardly distorted or refinancing is done bearing deadweight losses. Thus,
in these areas, expected welfare is below the first best level.

\section{Regional state capacity and the optimal institutional
regime}

\subsection{The main result}

In the initial period, there is an institutional choice between partial and full decentralization,
made under uncertainty. At this stage, the Congress observes projects’ benefits $b, B$, 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 computes, under each institutional regime $IR \in \{PD, FD\}$, the expected welfare
of a region

$$\mathbb{E}W^{IR} = \int_{0}^{b} \mathbb{E}W^{IR}_\ell h(c)dc$$

and chooses the regime that maximizes it.

\footnote{Figure 4 depicts full decentralization outcomes when $\theta < 1$. If $\theta = 1$, $c_2 = b$ and $\pi^F_2 = \pi^F_3$: the area
where projects are never refinanced vanishes.}
As one of the goals of this paper is to evaluate how this choice is affected by the level of regional state capacity, we first characterize the relation between $E^WR$ and the pair $(\pi, \theta)$. Taking into account equilibrium decisions and outcomes, in the Online Appendix we show that, when $L \leq \bar{L}$ (a threshold defined in the Appendix), $E^FD$ is a continuous, everywhere differentiable, increasing, convex function of the administrative capacity $\pi$. Moreover, we also show that it increases with the fiscal capacity $\theta$, except when $\pi = 1$. Therefore, under full decentralization, an increase in either $\pi$ or $\theta$ increases $E^FD$, as suggested by the decentralization literature mentioned in the Introduction. But this assertion does not necessarily imply that full decentralization dominates for high levels of state capacity. Why not? Because we also prove that $E^{PD}$ is a continuous, increasing, convex function of the administrative capacity $\pi$, with a kink at $\pi = \pi^{PD}_1$. Hence, the comparison between both regimes is not a priori evident. The following proposition presents the result that comes out this normative comparison, for every pair $(\pi, \theta) \in [0, 1]^2$.

**Proposition 6**

Assume an intermediate number of regions $L$ and a uniform distribution for the cost $c$.\(^{10}\)

When the regional fiscal capacity $\theta \leq \theta_0 \equiv 2L/(1 + L^2)$, partial decentralization dominates for all values of the regional administrative capacity $\pi \in [0, 1)$. When the regional fiscal capacity $\theta \geq \theta_0$, there exists a unique threshold $\tilde{\pi}(\theta)$ such that, when the regional administrative capacity $\pi \leq \tilde{\pi}(\theta)$, full decentralization dominates. Otherwise, partial decentralization dominates. When the regional administrative capacity $\pi$ is equal to one, both regimes are efficient.

The following figure illustrates this result. There, each point in the $(\theta, \pi)$ plane represents the regional state capacity that prevails in the federation. We point out the area where each institutional regime dominates.

---

\(^9\)When $\pi \geq \pi^{PD}_1$, $E^{PD}$ increases linearly with the administrative capacity $\pi$.

\(^{10}\)The technical conditions on $L$ are given in the Appendix.
When the regional fiscal capacity \( \theta \leq \theta_0 \), refinancing incomplete projects under full decentralization is too costly. Therefore, regardless the level of regional administrative capacity \( \pi \in [0, 1) \), partial decentralization dominates. This result supports Gadenne and Singhal’s (2014) empirical finding that there is relatively less fiscal decentralization and larger fiscal imbalances in developing countries, most of them characterized by low levels of regional fiscal capacity. But this partial decentralization’s dominance does not hold for all other values of \( \theta \). As expected, full decentralization may dominate when the regional fiscal capacity is relatively high. What is more surprising is that this indeed occurs\(^{11}\), but for relatively low levels of the regional administrative capacity, i.e., when \( \pi \leq \hat{\pi}(\theta) \). The intuition for this result is the following. When \( \pi \) is relatively low, it is very likely that projects will remain incomplete, and thus refinancing them is the issue to deal with. When is refinancing a problem under full decentralization? When there is at least one region that has carried out its project in due time, and thus the refinancing region needs to tax capital in a distortionary way. But when \( \pi \) is low, the likelihood of this event is also low. Thus, full decentralization’s distortions cost less than inefficient investments made under partial decentralization. In the Appendix, we show that \( \hat{\pi}(\theta) \) is unique and increases with \( \theta \).\(^{12}\) Indeed, the higher the fiscal capacity \( \theta \), the higher the administrative capacity \( \pi \) for which this abovementioned argument holds.

\(^{11}\)As \( L \geq L, \theta_0 \leq 0.6 \). So full decentralization dominates in a non negligible area of the figure. This result is in some way unexpected because, in this model, the fully decentralized regime generates qualitatively more distortions than partial decentralization.

\(^{12}\)In order to prove \( \hat{\pi}(\theta) \)’s unicity, we have imposed some conditions on the number of regions. When these sufficient conditions do not hold, the model becomes analytically untractable and thus \( \hat{\pi}(\theta) \)’s unicity 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 5. All simulations confirmed the results presented in Proposition 6, and are available upon request from the authors.
When \( \pi \) is above \( b/B \), the model is biased towards project initiation. Under partial decentralization, outcomes are efficient. On the other hand, under full decentralization, the likelihood of facing capital mobility costs or the project’s shutdown is very high. 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, we confirm that a high level of fiscal capacity is a necessary condition for full decentralization to be the optimal institutional regime. But, on the other hand, we also caution against the position, held by some contributors to the literature on decentralization, that affirm the necessity of high levels of regional state capacity 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, low levels of administrative capacity do not prevent full decentralization’s dominance. Under this last circumstance, the outside option (or the status quo regime) may generate more important distortions.

### 6.2 Robustness

In this section we undertake some sensitivity analysis of the model, and then we extend it, to evaluate its robustness.

#### 6.2.1 Comparative statics

Next paragraphs explain in detail, with reference to their corresponding figures, how changes in some parameters of the model modify the way by which regional state capacity affects the choice between partial and full decentralization.

![Figure 6a: Increase in capital endowment \( \kappa \)](image)

![Figure 6b: Increase in project’s benefit \( B \)](image)

![Figure 6c: Increase in the number of regions \( L \)](image)
Change in the capital endowment $\kappa$ Figure 6a shows that an increase in $\kappa$ favors full decentralization, implying that this regime dominates in 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 $\tau_\ell (0)$ decreases with $\kappa$, so does capital mobility costs.

Change in the highest project’s benefit $B$ Figure 6b depicts that an increase in $B$ always favors partial decentralization, implying that this regime dominates in a larger area of the plane $(\theta, \pi)$. The reason is as follows: when $B$ increases, the welfare loss due to overinvestment (underinvestment) under partial (full) decentralization decreases (increases).

Change in the number of regions $L$ Contrary to the previous comparative static exercises, Figure 6c 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. On the other hand, for relatively low values of the regional administrative capacity, full decentralization dominates in an area where partial decentralization used to be the optimal regime. The intuition for this result hinges on the relative costs of the consequences of $L$’s increase. On the one hand, the common-pool fiscal externality under partial decentralization gets larger. On the other hand, under full decentralization, the likelihood that a given region has to bear deadweight losses due to capital mobility or to shutdown its project 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 $\pi$.

6.2.2 An extension: distortionary national taxation.

So far, the tax set by the central government to bailout regions under partial decentralization was non-distortionary. This is of course a strong assumption: usually national taxes also generate deadweight losses. If so, assuming (as we did) non-distortionary national taxation underestimates the cost of the common-pool fiscal externality. To address this issue, let us model distortions in national taxation in the simplest possible way, by assuming that the cost of bailing out an incomplete project under partial decentralization is $c + \lambda$, where $\lambda > 0$.\textsuperscript{13} Although this change does not affect the fully decentralized regime, outcomes in the other regime are substantially modified, as shown by the following figure, which depicts equilibrium outcomes that emerge under partial decentralization with distortionary national taxation.\textsuperscript{14}

\textsuperscript{13}This assumption is a shortcut that captures, for example, cost differences in complying with central and regional tax systems. As stressed by Slemrod and Venkatesch (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.

\textsuperscript{14}Proofs that correspond to this section are in the Online Appendix.
Now, the central government refinances less projects than before. Indeed, only when \( c \leq b - \lambda < b \), incomplete projects are bailout; otherwise, the central government faces an endogenous hard budget constraint. When \( 0 \leq \pi \leq \pi_1^{PD,\lambda} \equiv \frac{b-\lambda}{L(B-b)+b} \), despite the fact that the central government bailouts all projects with costs lower than \( b-\lambda \), only those for which

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

are initiated and refinanced. Otherwise, projects are not undertaken, because refinancing them is too costly. As \( c^{PD,\lambda}(\pi) < c^{PD}(\pi) \), the condition for project initiation is stricter than under the same regime with non-distortionary national taxation. Hence, in the shaded area, less projects are inefficiently initiated than in Section 4. As \( \pi \) increases, the cost interval where these inefficient projects are initiated when the central government refinances vanishes.

And when the central government does not bailout regions, no project is undertaken in equilibrium, as shown when \( \frac{b-\lambda}{L(B-b)+b} < \pi \leq \pi_2^{PD,\lambda} \equiv \frac{b-\lambda}{B} \). Observe that, for \( \pi \in [\alpha, \pi_2^{PD,\lambda}] \), \( c^{PD,\lambda}(\pi) < c^*(\pi) \). This implies the opposite project’s distortion: in the shaded area, regions inefficiently underinvest. Then, as \( \frac{b-\lambda}{B} \leq \pi \leq \pi_3^{PD,\lambda} \equiv b/B \), 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. 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.

How this new regime compares with full decentralization? The following proposition presents the result.

**Proposition 7**

When the value of the deadweight loss \( \lambda \) increases, full decentralization dominates in a larger area of the plane \((\theta, \pi)\). relative to a situation with a non-distorted national tax system.
Clearly, departing from $\lambda = 0$, increasing $\lambda$ favors full decentralization. Although expected, the intuition of this result should not be based on the mere assertion that, by increasing the bailout cost, expected welfare under partial decentralization should automatically decrease. In fact, when $\lambda$ increases, this institutional regime does generate bailouts more costly 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 effects, we can show that the last one is dominated by the former three, and thus expected welfare under this regime decreases with the deadweight loss $\lambda$.

7 Related Literature

Related literature is as follows. First, other contributions have analyzed the trade-off between partial and full decentralization. Brueckner (2009) presents a Tiebout-type model, where local governments exert effort to reduce the cost of a local public good, private developers build houses and heterogeneous consumers decide on their location. Under partial decentralization, the federal government taxes the population and transfers the tax collection to jurisdictions, on an equally per capita basis. Then, local governments choose the level of effort and local public good provision, and finally Tiebout sorting occurs. Under full decentralization, the unique difference is that each local government sets a non-distortionary property tax to cover the public good’s cost. Brueckner (2009) finds that full decentralization always dominates, because the transfer’s uniformity under partial decentralization generates less variety of local public goods, and thus a worse preference matching. Partial decentralization can be optimal, provided local governments are of a Leviathan type. Here, in our model, we do not find a complete regime dominance. Peralta (2011) compares both regimes, focusing on accountability issues at the local level. She presents a political economy model, with benevolent and self-interested, rent-seeking local politicians, who tax their jurisdiction and provide local public goods, in a context of double asymmetric information because voters observe neither the politician’s type nor the local public good’s cost. Two equilibria emerge: a pooling equilibrium, when the rent-seeker politician mimics a benevolent one, and a separating equilibrium, when he extracts rents, and then he is replaced in the following election. Peralta (2011) shows 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 benevolents). This last regime dominates when the proportion of rent seekers is low. Other authors have adopted different definitions of partial decentralization. Janeba and Wilson (2011) and Hatfield and Padró i Miquel (2012) define partial decentralization when a subset of public goods are exclusively funded and provided by local governments. These authors find that devolving some public goods to local authorities is always optimal. Janeba and Wilson (2011) trade-off inefficient provision decided by a minimum winning coalition at the central level against distortionary taxation and low public good provision in a context of capital tax competition under full decentralization. In a voting model, Hatfield and Padró i Miquel (2012) obtain that devolution serves as a commitment
device against excessive capital taxation chosen when individuals vote on the central provision of public goods. Joannis (2014) defines partial decentralization as “shared responsibility”, an institutional regime where both the central and the regional government participate in the funding of a given public good. In his model, central and local, rent-seeking, politicians simultaneously decide how much fiscal resources invest in the provision of this public good. Despite their preferences for rents, they also invest in such provision because their aim is to manipulate their reelection probability, in a context of weak accountability, where the electorate is unable to assess the contribution of each level of government to the public good provision. Partial decentralization obtains, balancing losses in productive complementarities between both levels of government and the lack of accountability at the local level. The main differences between our paper and these contributions hinges on the fact that we study a different trade-off between partial and full decentralization, namely inefficient bailouts and projects overprovision vs. capital tax competition and projects underprovision. Moreover, these articles do not incorporate regional state capacity into the analysis, and thus they do not consider it as a relevant factor affecting the trade-off.

The paper is also related to an important set of contributions that analyze the pros and cons of different types of regional budget constraints in federations. On the one hand, 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 this paper are the following. First, they do not describe the institutional regime that hardens regional government’s budget constraints; they simply assume that the federal government is able to impose, exogenously and at no cost, a hard budget constraint to local governments. Here, we analyze full decentralization as the institutional regime that hardens regional governments’ budget constraint, but endogenously and costly. Second, the authors compare, from a normative point of view, soft and hard budget constraints at the interim stage (in other words, project by project) whereas we take an ex ante perspective. Finally, 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 concerns. 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. Finally, Silva and Caplan (1997), Caplan et al. (2000) and Köthenbürger (2004) claim that, under some conditions, 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 local public goods or tax spillovers to hold.

Finally, the paper is related to a recent literature that studies empirically how, in contexts of decentralized regimes, local state capacity impacts 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. She finds evidence that both household consumption and school enrollment are positively related with the level of capacity of district governments. Loayza et
al. (2014) evaluate how budget size and allocation process, local capacity, local needs, and political economy considerations (four factors that are usually considered important features that affect the effectiveness of decentralization reforms) affect municipal budget execution rate in Peru. The authors find that budget size and local capacity are the statistically most important constraints that explain municipal budget execution rates. Bandyopadhyay and Green (2012) show that higher percentage of residents from centralized ethnic groups imply higher development indicators at the local level in Uganda. Finally, Acemoglu et al. (2014) study the impact of municipal state capacity on public goods provision in Colombia. An important feature of their paper is the consideration of spillovers: when a municipality invests in its state capacity, it also generates positive effects on neighboring municipalities. They empirically confirm that state capacity decisions are indeed strategic between municipalities, and with large effects on local prosperity. All these papers take the intergovernmental institutional setting as given, and thus do not analyze different regimes, as we do.

8 Conclusions

This paper presents a model featuring a central government and regional authorities. The latter are characterized by the same level of administrative and fiscal capacity. We analyze two fiscal regimes. Under partial decentralization, regional governments invest in a local public project but, if it remains incomplete, it is the central government that bailouts the regions. Thus, regions face soft budget constraints and inefficiently overinvest in local public projects. On the other hand, under full decentralization, regional governments also decide on initial investment but cannot rely on central bailouts. Thus, regional governments face hard budget constraints because capital tax competition increases the marginal cost of public funds. This implies the opposite type of inefficiency: they inefficiently underinvest.

The first goal of the paper is to evaluate the normative comparison between these distortions, and to assess which regime dominates. The second goal is to investigate how different levels of regional state capacity affect this comparison. As expected, when the regional fiscal capacity is low, partial decentralization dominates; otherwise full decentralization may dominate. But, contrary to the common wisdom, we find that full decentralization should be the preferred regime when the regional administrative capacity is low.

The model can be generalized and extended in several directions. In particular, we would like to consider ex ante asymmetries between regions, either in capital endowments or in state capacities. This is an essential “input” if one would like to analyze how regional governments invest in state capacity under different fiscal regimes. Also, performing this extension would be crucial for the study of the long term trade-off between partial and full decentralization.

References


Appendix

9.1 Proof of Proposition 1

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

$$\kappa + \mathbb{I}_{\{i_1=1\}} \left[ \pi B + (1 - \pi) \left( b - \frac{c}{L} \right) - c \right] - \sum_{m \neq \ell} \mathbb{I}_{\{i_m=1\}} (1 - \pi) \cdot \frac{c}{L},$$

(4)
whereas its expected, net welfare from not investing is
\[ \kappa - \sum_{m \neq \ell} \mathbb{I}_{\{i_m = \ell\}} (1 - \pi) \cdot \frac{c}{L}. \]

So, for any region \( \ell \), initiating the project is a dominant strategy if
\[ c \leq c^{PD}(\pi) \equiv \frac{L[\pi B + (1 - \pi)b]}{L + 1 - \pi}. \]

It is straightforward to show that \( c^{PD}(\pi) > c^*(\pi) \), \( \frac{\partial}{\partial \pi} c^{PD}(\pi) = \frac{L(B - b) + B}{(L + 1 - \pi)^2} > 0 \), and that \( c^{PD} = b \) when \( \pi = \frac{b}{L(B - b) + B} \). 

### 9.2 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
\[ f_{\ell m} \geq 0 \quad \forall m \neq \ell, \]

where \( h_{\ell} \) is capital invested in region \( \ell \), and \( f_{\ell m} \) is capital invested in another region \( m \neq \ell \).

Let’s 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
\[
\begin{align*}
\tau_{\ell} - \tau_m + \lambda_{\ell m} &= \sum_{m \neq \ell} f_{\ell m} \\
\lambda_{\ell m} f_{\ell m} &= 0 \quad \lambda_{\ell m} \geq 0 \quad \forall m \neq \ell.
\end{align*}
\]

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

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

**Proof.**

i) First, assume that \( \text{Card}\{\tilde{M}\} = 1 \) and consider the first-order condition

\[
\tau_\ell - \tilde{\tau}_\ell + \lambda_{\ell \tilde{m}} = f_{\ell \tilde{m}}. \tag{5}
\]

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

\[
\lambda_{\ell \tilde{m}} = f_{\ell \tilde{m}}.
\]

If \( \lambda_{\ell \tilde{m}} > 0 \), then \( f_{\ell \tilde{m}} > 0 \), which implies, by the complementary slackness condition, that \( \lambda_{\ell \tilde{m}} = 0 \), which is a contradiction. Hence \( \lambda_{\ell \tilde{m}} = 0 \).

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

\[
\tau_\ell - \tilde{\tau}_\ell + \lambda_{\ell \tilde{m}} = \sum_{\tilde{m} \in \tilde{M}} f_{\ell \tilde{m}}.
\]

In order to satisfy these first-order conditions, all multipliers \( \lambda_{\ell \tilde{m}} \) should have the same value. If they were all strictly positive, then, by their corresponding complementary slackness condition, all outflows \( f_{\ell \tilde{m}} \) should be equal to 0, implying that \( \sum_{\tilde{m} \in \tilde{M}} f_{\ell \tilde{m}} = 0 \). But, as \( \tau_\ell \geq \tilde{\tau}_\ell \), all first-order conditions would yield a contradiction. Hence, all multipliers \( \lambda_{\ell \tilde{m}} = 0 \).

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

\[
\sum_{\tilde{m} \in \tilde{M}} f_{\ell \tilde{m}} = \tau_\ell - \tilde{\tau}_\ell \tag{6}
\]

### 9.3 Proof of Proposition 3

Due to specific features of this model, we cannot apply Wildasin’s (1988) methodology. So, to obtain the equilibrium tax rates, first we derive region \( \ell \)'s reaction function.

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

Let’s denote by \( \tau_m \) the profile of tax rates chosen by regions \( m \neq \ell \). For any profile \( \tau_m \), we need to consider three cases.
1. If the regional government of $\ell$ plans to set its tax rate strictly above $\tilde{\tau}_\ell$, there will be capital outflows to regions $\tilde{m} \in \tilde{M}$. Hence, the regional welfare would be

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

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

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

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

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

Throughout the paper, we assume that $\kappa$ is sufficiently large so that this square root always exists (see footnote 15).

2. If the regional government of $\ell$ plans to set its tax rate strictly below $\tilde{\tau}_\ell$, there will be no capital outflows to regions $m \neq \ell$. Thus, the 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, the regional government of $\ell$ should choose the lowest tax rate that satisfies its budget constraint

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

where, by a symmetric application of (6),

$$f_{m\ell} = \tau_m - \tau_\ell \quad (10)$$

represents a capital inflow from a region $m \neq \ell$ to region $\ell$. Rearranging terms, the smallest root of (9) is given by

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

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

$$W_{\ell}^{FD} = \kappa (1 - \tilde{\tau}_\ell) + \mathbb{1}_{\{ETC \geq c\}} b,$$
where $ETC$ denotes the effective tax collection

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

and $\alpha_\ell$, $0 < \alpha_\ell < 1$, represents the fraction of capital outflows that leave regions $m \notin \tilde{M}$ and move to region $\ell$.

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

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

$$\frac{\partial \tau_\ell}{\partial \tau_m} < 0.$$

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

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

We can show that

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

and

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

If $\widetilde{\tau}_\ell \leq c/\theta \kappa$, $\tau_\ell(\widetilde{\tau}_\ell) \geq \widetilde{\tau}_\ell$, contradicting the definition of $\tau_\ell(\widetilde{\tau}_\ell)$ as the highest tax rate that region $\ell$ can set strictly below $\widetilde{\tau}_\ell$. Therefore, if $\widetilde{\tau}_\ell \leq c/\theta \kappa$, the government in region $\ell$ cannot set a tax rate lower than $\widetilde{\tau}_\ell$ while, at the same time, satisfying its budget constraint.

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

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

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

and

$$\lim_{\widetilde{\tau}_\ell \to c/\theta \kappa} \tau_\ell = c/\theta \kappa.$$
If $\tau_\ell \geq c/\theta\kappa$, $\tau_\ell(\tau_\ell) \leq \tau_\ell$, which again contradicts the definition of $\tau_\ell$ as the lowest tax rate that region $\ell$ can set strictly above $\tau_\ell$. Hence, if $\tau_\ell \geq c/\theta\kappa$, the government in region $\ell$ cannot set a tax higher than $\tau_\ell$ while, at the same time, satisfying its budget constraint.

Two more cases remain to be analyzed. First, consider a profile of tax rates $\tau_m$ such that $\tau_\ell = c/\theta\kappa$ but when at least one region $n \neq \ell$ has set $\tau_n > c/\theta\kappa$. We have just proved that region $\ell$ cannot tax strictly above $c/\theta\kappa$. If region $\ell$ chooses $\tau_\ell = c/\theta\kappa$, there are no outflows $f_{\ell m}$. Thus, its welfare amounts to $W_\ell^{FD} = \kappa - c$. But region $\ell$ receives a capital inflow $\alpha_\ell \sum_{m \notin M} \sum_{\tilde{m} \in M} f_{\ell \tilde{m}}$, and so its tax collection is

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

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

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

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

$$\tau_\ell (\tau_m) =\begin{cases} 
\frac{1}{2} \left[ \kappa + \tau_\ell - \sqrt{(\kappa + \tau_\ell)^2 - \frac{4c}{\theta}} \right] & \text{if } \tau_m < c/\theta\kappa \forall m \neq \ell \\
c/\theta\kappa & \text{if } \tau_m = c/\theta\kappa \forall m \neq \ell \\
\frac{1}{2} \left[ \frac{\kappa + \sum_{m \neq \ell} \tau_m}{L-1} - \sqrt{\left( \frac{\kappa + \sum_{m \neq \ell} \tau_m}{L-1} \right)^2 - \frac{4c}{\theta(L-1)}} \right] & \text{if } \tau_m \geq c/\theta\kappa \forall m \neq \ell \\
\text{and } \exists n \neq \ell : \tau_n > c/\theta\kappa. 
\end{cases} \quad (12)$$

The reaction function is build around the value $c/\theta\kappa$. When the profile of tax rates $\tau_m$ is such that all tax rates $\tau_m < c/\theta\kappa$, region $\ell$'s optimal response is to tax strictly above the minimum tax $\tau_{\ell}$, which again contradicts the definition of $\tau_{\ell}$ as the lowest tax rate that region $\ell$ can set strictly above $\tau_{\ell}$. In this case, $\tau_{\ell} (\tau_m)$ depends only upon $\tau_{\ell}$, the unique pertinent tax rate that defines the size of the capital outflow. When the profile of tax rates changes, so that the minimum tax rate $\tau_{\ell}$ increases and converges to $c/\theta\kappa$ from below, $\tau_{\ell} (\tau_m)$ decreases and converges to $c/\theta\kappa$ from above.

When all tax rates $\tau_m$ are equal to $c/\theta\kappa$, region $\ell$'s optimal response is to replicate this level. Finally, when the profile of tax rates is such that $\tau_m \geq c/\theta\kappa$ 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 $\tau_{\ell}$. In this case, $\tau_{\ell} (\tau_m)$ depends upon the whole profile $\tau_m$, because each tax rate $\tau_m$ determines the size of the capital inflow that moves from a region $m \neq \ell$ to region $\ell$.

Region $\ell$'s reaction function is non-continuous. When $\tau_{\ell}$ converges to $c/\theta\kappa$ from below, $\tau_{\ell} (\tau_m)$ converges to this limit from above. But when the distribution of taxes is such that $\tau_{\ell} = c/\theta\kappa$ but at least one region $n \neq \ell$ has set $\tau_n > c/\theta\kappa$, the optimal response is to set $\tau_{\ell} (\tau_m) < c/\theta\kappa$.
Clearly, \( \tau_1 = ... = \tau_L = c/\theta \kappa \) is a Nash equilibrium of this subgame because it is a fixed point of the best reply correspondence. To prove uniqueness, we proceed in two steps. First, by a simple inspection of (12), it is immediate to notice that no asymmetric choice of taxes can be a Nash equilibrium. Second, we show that there cannot be another symmetric equilibrium. Assume the contrary: let \( \tau'_\ell = \tau'_m \) and, without any loss of generality,

\[
\tau'_\ell = \tau'_m = \frac{c}{\theta \kappa} + \varepsilon,
\]

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

9.3.2 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 \( \tau_\ell(0) \). This tax rate is obtained replacing \( \tau'_\ell \) by 0 in the definition of \( \tau'_\ell \).

9.4 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

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

Hence, by Bolzano’s Theorem, there exists a threshold \( 0 < c_1 < b \) such that, when \( c \leq c_1 \), \( b - T(c, \theta) \geq 0 \). Replacing \( \tau_\ell(0) \) in \( T(c, \theta) \), and then equaling to \( b \), we obtain

\[
c_1 = \theta(2b - H),
\]

where \( H = \kappa^2 - \sqrt{\kappa^2(\kappa^2 - 2b)} \). Also, as \( c/\theta > c \), there trivially exists a threshold \( c_2 \) such that \( c_2/\theta = b \). Moreover, as \( c/\theta < T(c, \theta) \), \( c_1 < c_2 \).

Then, consider the first case, when all regions face an incomplete project. When \( 0 \leq c \leq c_1 \),

\[
W^{FD}_\ell(\ell = R, m, i) \geq W^{FD}_\ell(\ell = NR, m, i).
\]

Refinancing is a dominant strategy, and so all regions refinance in equilibrium. When \( c_1 < c \leq c_2 \),

\[
W^{FD}_\ell(\ell = R, m = R, i) = b - c/\theta \geq W^{FD}_\ell(\ell = NR, m = R, i) = 0 \quad \forall m \neq \ell
\]

\[15\] A sufficient condition for the existence of the square root in (8) and to ensure that \( \tau'_\ell < 1 \) is \( \kappa \geq \text{Max}\{2\sqrt{b/\theta}, b/\theta\} \).
but
\[ W^{FD}_\ell(r_\ell = R, r_m = NR, i) = b - T(c, \theta) \leq W^{FD}_\ell(r_\ell = NR, r_m = NR, i) = 0 \quad \forall m \neq \ell. \]

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 refines.

In fact, we prove that the first equilibrium is both the strong (Aumann, 1959) and the coalition-proof Nash equilibrium (Berheim et al., 1987), as follows.

1. As both equilibria are Nash equilibria, no region can do better by unilaterally changing its equilibrium strategy. Then consider \( L \)-regions coalitions, with \( 1 < L < L \). If the other regions refinance (not refinance), the \( L \) 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 refines, 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 \( L \)-regions coalitions, with \( 1 < L < L \). If the other regions refinance, the coalition of \( L \) regions can jointly decide to not 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. But 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 refines, 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 select the equilibrium where all regions refinance as the Nash equilibrium of this subgame. Finally, when \( c_2 < c \leq b \),
\[ W^{FD}_\ell(r_\ell = R, r_m, i) \leq W^{FD}_\ell(r_\ell = NR, r_m, i). \]

Not refinancing is a dominant strategy, and thus no region refinances in equilibrium.

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

9.5 Proof of Proposition 5

First, we proceed to evaluate expected, net regional welfares under different parameter conditions. If \( c \leq c_1 \), region \( \ell \)'s expected, net welfare is:
\[
\begin{align*}
EW^{FD}_\ell(i_\ell = I, i_m = I) &= \kappa - c + \pi B + (1 - \pi)L[b - \frac{c}{\theta}] \\
&\quad - (1 - \pi)(1 - (1 - \pi)^{L-1})[b - T(c, \theta)] \quad \text{if } i_m = I \forall m \neq \ell \\
EW^{FD}_\ell(i_\ell = I, i_m = NI) &= \kappa - c + \pi B + (1 - \pi)[b - T(c, \theta)] \quad \text{if } \exists m \neq \ell : i_m = NI \ell \\
EW^{FD}_\ell(i_\ell = NI, i_m) &= \kappa \quad \forall i_m \in \{I, NI\}
\end{align*}
\]
If \( c_1 < c \leq c_2 \), region \( \ell \)'s expected, net welfare is:

\[
\begin{align*}
\mathbb{E}W^FD_\ell(i_\ell = I, i_m = NI) &= \kappa - c + \pi B + (1 - \pi)L[b - \frac{c}{\theta}] \quad \text{if } i_m = I \forall m \neq \ell \\
\mathbb{E}W^FD_\ell(i_\ell = I, i_m = NI) &= \kappa - c + \pi B \quad \text{if } \exists m \neq \ell : i_m = NI_\ell \\
\mathbb{E}W^FD_\ell(i_\ell = NI, i_m) &= \kappa \quad \forall i_m \in \{I, NI\}
\end{align*}
\]

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

\[
\begin{align*}
\mathbb{E}W^FD_\ell(i_\ell = I, i_m) &= \kappa - c + \pi B \quad \forall i_m \in \{I, NI\} \\
\mathbb{E}W^FD_\ell(i_\ell = NI, i_m) &= \kappa \quad \forall i_m \in \{I, NI\}
\end{align*}
\]

In order to characterize the Nash equilibria, we need to evaluate these different levels of expected, net regional welfare. As their comparison is cumbersome, we proceed in an indirect way, as follows. First, let’s define

\[
\Delta^*(c) \equiv (1 - \pi)(B - (b - c))
\]

it 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. \( \mathbb{E}W^FD_\ell(i_\ell = I, i_m = I) \geq \mathbb{E}W^FD_\ell(i_\ell = NI, i_m) \iff \Delta_R^FD(c) \leq B - c \), where

\[
\Delta_R^FD(c) \equiv (1 - \pi)\bigg[B - (1 - \pi)L^{-1}(b - \frac{c}{\theta}) - (1 - (1 - \pi)L^{-1})(b - T(c, \theta))\bigg]
\]

is a continuous, increasing, convex function of \( c \) that satisfies

\[
\lim_{c \to 0} \Delta_R^FD(c) \equiv \Delta_R^FD(0) = (1 - \pi)[B - b] = \Delta^*(0), \quad \Delta_R^FD(c) > \Delta^*(c)
\]

and

\[
\lim_{c \to c_1} \Delta_R^FD(c) = \Delta_R^FD(c_1) \equiv (1 - \pi)\bigg[B - (1 - \pi)L^{-1}(b - \frac{c_1}{\theta})\bigg] < \Delta^*(b).
\]

2. \( \mathbb{E}W^FD_\ell(i_\ell = I, i_m = NI) \geq \mathbb{E}W^FD_\ell(i_\ell = NI, i_m) \iff \Delta'(c) \leq B - c \), where

\[
\Delta'(c) \equiv (1 - \pi)[B - (b - T(c, \theta))]
\]

is another continuous, increasing, convex function of \( c \) that satisfies

\[
\lim_{c \to 0} \Delta'(c) = (1 - \pi)[B - b] = \Delta^*(0), \quad \Delta'(c) > \Delta_R^FD(c) > \Delta^*(c)
\]

and

\[
\lim_{c \to c_1} \Delta'(c) = \Delta'(c_1) \equiv (1 - \pi)B = \Delta^*(b).
\]

36
When $c_1 < c \leq c_2$, we can also show that:

1. $\mathbb{E}W^F_D(i_\ell = I, i_m = I) \geq \mathbb{E}W^F_D(i_\ell = NI, i_m) \Leftrightarrow \Delta^{FD}_{NAR}(c) \leq B - c$, where

   \[ \Delta^{FD}_{NAR}(c) \equiv (1 - \pi) \left[ B - (1 - \pi)L^{-1}(b - \frac{c}{\theta}) \right] \]

   is a continuous, increasing, linear function of $c$ that satisfies

   \[ \lim_{c \to c_1} \Delta^{FD}_{NAR}(c) = \Delta^R(c_1), \quad \Delta^{FD}_{NAR}(c) > \Delta^*(c) \quad \text{and} \quad \lim_{c \to c_2} \Delta^{FD}_{NAR}(c) = \Delta^{FD}_{NAR}(c_2) \equiv (1 - \pi)B = \Delta^*(b). \]

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

   \[ \Delta''(c) \equiv (1 - \pi)B \]

   satisfies

   \[ \Delta''(c) = \Delta^{FD}_{NAR}(c_2) = \Delta'(c_1) = \Delta^*(b). \]

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

\[ \Delta^{FD}_{NR}(c) \equiv (1 - \pi)B = \Delta''(c). \]

These functions $\Delta^*(c), \Delta^R(c), \Delta'(c), \Delta^{FD}_{NAR}(c)$ and $\Delta^{FD}_{NR}(c)$ are (weakly) increasing in $c$, while $B - c$ decreases with $c$. Therefore, at some point, these five functions intersect $B - c$. We characterize these intersections as cost thresholds. Then, we divide the cost range $[0, b]$ in sub-intervals, according to these thresholds, and we find the Nash equilibria in each corresponding sub-interval. Moreover, as these five functions depend upon $\pi$, the cost thresholds also depend upon this parameter.

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

   Let $\pi_1^{FD}$ be implicitly defined by $\Delta^R(c_1) = B - c_1$. As $\Delta'(c) \geq \Delta^R(c) \geq \Delta^*(c)$, the intersection between $\Delta'(c)$ and $B - c$ defines a threshold $c'(\pi) \leq c^*(\pi)$, whereas the intersection between $\Delta^R(c)$ 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:\footnote{We present here the complete proof for this case. We omit those for the remaining cases because they are identical.}

   (a) When $0 \leq c \leq c'(\pi), \Delta^R(c) \leq \Delta'(c) \leq B - c$, which implies that

   \[ \mathbb{E}W^F_D(i_\ell = I, i_m = I) \geq \mathbb{E}W^F_D(i_\ell = NI, i_m) \]

   and

   \[ \mathbb{E}W^F_D(i_\ell = I, i_m = NI) \geq \mathbb{E}W^F_D(i_\ell = NI, i_m). \]

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

37
(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
\[
\mathbb{E}W_{\ell}^{FD}(i_{\ell} = I, i_{m} = NI) \leq \mathbb{E}W_{\ell}^{FD}(i_{\ell} = NI, i_{m}).
\]
So 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. As the first equilibrium is strong and coalition-proof, we select it as the Nash equilibrium.

(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. $\pi^FD_3 < \pi \leq 1$

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.

9.6 Proof of Proposition 6

The proof of this proposition uses the expected welfare expressions derived in the Online Appendix, and many intermediate results.

9.6.1 Analysis of expected welfares at the extreme values of $\pi$

$\pi=0$

When $\pi = 0$, $c^*(0) = b/2$, $c^{PD}(0) = \frac{Lb}{L+1}$ and $c^{FD}(0) = \frac{\theta b}{1+\theta}$. Hence, expected welfares are

$$E_{PD}(\pi, \theta) = \kappa + \frac{bL}{(1 + L)^2}$$

and

$$E_{FD}(\pi, \theta) = \kappa + \frac{\theta b}{2(1 + \theta)}.$$

Let $\theta_0 \equiv \frac{2L}{1+L}$ be the regional fiscal capacity that equalizes these expected welfares. Hence, when $\theta \geq \theta_0$, $E_{FD} \geq E_{PD}$. Otherwise, partial decentralization dominates.

Last, we analyze the value of $\frac{\partial E_{FD}(\pi, \theta)}{\partial \pi} \bigg|_{\pi=0}$ 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

$$\frac{\partial E_{FD}(\pi, \theta)}{\partial \pi} \bigg|_{\pi=0} = \frac{-\theta L[-12bB(1 + \theta) + 8(L - 1)b^2(\beta^2 - 1)^{\frac{3}{2}} + b^2(9 - 12\beta^2 + 8\beta^4 + 12\theta + L(-8\beta^4 + 2\beta^2 - 3))]}{12(1 + \theta)^2}.$$

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

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

The fraction’s denominator converges fast to zero, i.e., for values of $\beta$ below 1. But recall that in footnote 15 we have imposed $\kappa$ to be sufficiently high; so $\beta$ should be in fact a real
number higher than 1. Thus, $\tilde{L}(\theta)$ turns out to be a big number. Moreover, $\tilde{L}(\theta)$ increases with $\theta$. Hence, if $L \leq \tilde{L} \equiv \tilde{L}(0)$

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

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 $\pi$.

$\pi^{PD}_2 = \pi^{PD}_3 = b/B \leq \pi \leq 1$

When $b/B \leq \pi < 1$, partial decentralization replicates the first best outcomes, whereas full decentralization’s outcomes are highly 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, 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.

9.6.2 Comparison of expected welfares when $\theta \in [\theta^*, 1]$

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

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

and when $\pi = \pi^{PD}_1$,

$$\mathbb{E}W^{PD}(\pi^{PD}_1) = \frac{bL(2B - b)}{2[BL - b(L - 1)]}.$$

If $L \geq \tilde{L} \equiv 3 + \frac{B}{B-b}$, $\mathbb{E}W^{FD}(0, 1) \geq \mathbb{E}W^{PD}(\pi^{PD}_1)$. Hence, $\mathbb{E}W^{FD}(\pi, \theta)$ lies everywhere above $\mathbb{E}W^{PD}(\pi)$ when $\pi \in [0, \pi^{PD}_1]$ 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. Let’s denote by $\hat{\pi}(1)$ the administrative capacity level that corresponds to this intersection.\(^\text{17}\) 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

\(^\text{17}\) As partial decentralization dominates when $\pi \geq \pi^{PD}_2$, $\pi^{PD}_1 \leq \hat{\pi}(1) < \pi^{PD}_2$. 

40
$E_{W}^{FD}(\pi, \theta^*)$ lies above $E_{W}^{PD}(\pi)$ everywhere when $\pi \in [0, \pi_1^{PD}]$. Hence, when $\theta \in [\theta^*, 1]$, $E_{W}^{FD}(\pi, \theta)$ intersects $E_{W}^{PD}(\pi)$ in its linear part. Using the same geometrical argument as we did above, we know that both expected welfares cross only once, at $\hat{\pi}(\theta)$. Applying the Implicit Function Theorem, we show that

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

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

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

9.6.3 Comparison of expected welfares when $\theta \leq \theta_0$

When $\theta = \theta_0$,

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

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

$$\frac{\partial E_{W}^{FD}(\pi, \theta)}{\partial \pi} \bigg|_{\pi=0} = \frac{1}{b} \int_0^{\theta_0 b} [B - b + L \frac{c}{\theta_0} + (1 - L)T(c, \theta_0)] dc$$

$$< \frac{1}{b} \left\{ \int_0^{\theta_0 b} [B - b] dc + \int_0^{\theta_0 b} T(c, \theta_0) dc \right\}$$

$$< \frac{1}{b} \int_0^{\theta_0 b} [B - b] dc + \frac{\theta_0 b}{2(1 + \theta_0)} = \frac{(2b - b)L}{(1 + L)^2},$$

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

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

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

$$E_{W}^{FD}(\pi_1^{PD}, \theta_0) = \kappa + \frac{1}{b} \left\{ \int_0^{\pi_1^{PD}} [\pi_1^{PD}B - c] dc + \int_0^{\pi_1^{PD}} (1 - \pi_1^{PD})L(b - \frac{c}{\theta_0}) dc \right\}$$

$$+ \int_0^{\pi_1^{PD}} (1 - \pi_1^{PD})(1 - (1 - \pi_1^{PD})^{L - 1})(b - T(c, \theta_0)) dc \right\}$$

$$+ \int_0^{\pi_1^{PD}} (1 - \pi_1^{PD})(1 - (1 - \pi_1^{PD})^{L - 1})(b - T(c, \theta_0)) dc \right\}$$
\[
\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0) = \kappa + \frac{bb^2}{2(b+L(b-b))^2} + \frac{bL(1-\frac{b}{b+L(b-b)})}{1+L^2} + \frac{1}{b} \int_0^c (1 - \pi_1^{PD})(1 - (1 - \pi_1^{PD})^{L-1}')(b - T(c, \theta_0))dc
\]

\[
< \kappa + \frac{bb^2}{2(b+L(b-b))^2} + \frac{bL(1-\frac{b}{b+L(b-b)})}{1+L^2} + \frac{1}{b} \int_0^c (1 - \pi_1^{PD})(1 - (1 - \pi_1^{PD})^{L-1}')(b - \frac{c}{\theta_0})dc
\]

\[
= \kappa + \frac{bb^2}{2(b+L(b-b))^2} + \frac{bL(1-\frac{b}{b+L(b-b)})}{1+L^2} \equiv A.
\]

If \(\pi_1^{PD} < \pi_2^{FD}\),

\[
\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0) = \kappa + \frac{1}{b} \left\{ \int_0^{c^{PD}} \frac{[\pi_1^{PD} - (1 - \pi_1^{PD})^{L-1}']}{\pi_1^{PD}} \mathbb{P}(\pi_1^{PD} = 1) \right\} dc
\]

\[
+ \int_0^c (1 - \pi_1^{PD})(1 - (1 - \pi_1^{PD})^{L-1}')(b - T(c, \theta_0))dc
\]

\[
\kappa + \frac{1}{b} \left\{ \int_0^c \left[ \pi_1^{PD} B - c + (1 - \pi_1^{PD})L(b - \frac{c}{\theta_0}) \right] dc
\]

\[
\int_0^c (1 - \pi_1^{PD})(1 - (1 - \pi_1^{PD})^{L-1}')(b - T(c, \theta_0))dc < A.
\]

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

\[
\Gamma_{\theta_0}(\pi_1^{PD}) \equiv \mathbb{E}W^{PD}(0) + \left. \frac{\partial \mathbb{E}W^{PD}(\pi_1^{PD})}{\partial \pi_1^{PD}} \right|_{\pi_1^{PD} = 0}
\]

\[
= \kappa - \frac{bL[bL(L+5) - 2B(L+1)(L+2)]}{2(1+L)^3[B+b(L-1)]}.
\]

Finally, we evaluate

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

As \(L \geq L\), we can write \(L = 2 + n + \frac{B}{B-6}\), 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 \(n\) is not possible to obtain a close form expression for \(\alpha(n)\), the following figure depicts the curve \(\alpha(n)\), when \(n \in \{0, \ldots, 100\}\).  

![Figure 8: The curve \(\alpha(n)\)](image_url)
As \( \alpha(1) = 2.6817 \) and \( B/2 < b, \) \( H(b, B, L) > 0 \) for any \( n \in \{1, ..., 100\}. \) 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 \( \Gamma_0(\pi) \), a straight line that takes the values \( \mathbb{E}W^{PD}(0) \) and \( \Gamma_0(\pi_1^{PD}) \) when \( \pi = 0 \) and \( \pi_1^{PD} \), respectively. In fact, \( \Gamma_0(\pi) \) is the tangent of \( \mathbb{E}W^{PD}(\pi) \) 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 9: Expected welfares when \( \theta = \theta_0 \) and \( \pi \in [0, \pi_1^{PD}] \)](image)

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 \( \theta \leq \theta_0 \).

### 9.6.4 Comparison of expected welfares 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

\[
\mathbb{E}W^{FD}(0, \theta_0 + \epsilon) > \mathbb{E}W^{PD}(0) \quad \text{and} \quad \mathbb{E}W^{FD}(\pi_1^{PD}, \theta_0 + \epsilon) < \Gamma_0(\pi_1^{PD}) < \mathbb{E}W^{PD}(\pi_1^{PD}).
\]

Clearly, \( \mathbb{E}W^{FD}(\pi, \theta_0 + \epsilon) \) intersects the line \( \Gamma_0(\pi) \) 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

\[
\mathbb{E}W^{FD}(\pi_1^{PD}, \theta_1) = \Gamma_0(\pi_1^{PD}).
\]

\footnote{This result also holds for \( n > 100 \).}
When $\theta = \theta_1$, there can be two possibilities. When $\pi$ converges to $\pi_1^{PD}$, either 1) $EW^{FD}(\pi, \theta_1)$ converges to $\Gamma_{\theta_0}(\pi_1^{PD})$ from below, in which case the previous intersection argument applies, or 2) this convergence is from above. This means that

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

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

Now let’s construct $\Gamma_{\theta_1}(\pi)$: it’s a line that has the same slope than $\Gamma_{\theta_0}(\pi)$ and is defined by $\Gamma_{\theta_1}(0) = EW^{FD}(0, \theta_1)$. By construction, $\Gamma_{\theta_1}(\pi_1^{PD}) > EW^{FD}(\pi_1^{PD}, \theta_1)$. Now assume that we can find $\theta = \theta_1 + \epsilon$ such that $EW^{FD}(0, \theta_1 + \epsilon') > \Gamma_{\theta_1}(0)$ and $EW^{FD}(\pi_1^{PD}, \theta_1 + \epsilon') < \Gamma_{\theta_1}(\pi_1^{PD})$. Clearly, $EW^{FD}(\pi, \theta_1 + \epsilon')$ crosses only once $\Gamma_{\theta_1}(\pi)$, from above. Hence, as $EW^{FD}(\pi, \theta_1 + \epsilon)$ and $EW^{PD}(\pi)$ 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

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

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

$$\begin{cases} 
EW^{FD}(0, \theta_{n+1}) > \Gamma_{\theta_n}(0) \\
EW^{FD}(\pi_1^{PD}, \theta_{n+1}) = \Gamma_{\theta_n}(\pi_1^{PD}).
\end{cases}$$

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 $EW^{FD}(\pi, \theta)$ crosses $EW^{PD}(\pi)$ only once, from above, when $\pi \in [0, \pi_1^{PD}]$. Let’s we 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. Let’s denote this limit by $\overline{\theta}$. As the functions that define the sequence are continuous, they also converge towards $EW^{FD}(\pi, \overline{\theta})$ and $\Gamma_{\overline{\theta}}(\pi)$. These limit functions have to satisfy

$$\begin{cases} 
EW^{FD}(0, \overline{\theta}) = \Gamma_{\overline{\theta}}(0) \\
EW^{FD}(\pi_1^{PD}, \overline{\theta}) = \Gamma_{\overline{\theta}}(\pi_1^{PD}).
\end{cases}$$

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

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

44
9.7 Comparative statics

9.7.1 Increase in the capital endowment $\kappa$

Applying the Implicit Function Theorem, we can compute

$$\frac{\partial \hat{\pi}(\theta)}{\partial \kappa} = \frac{\partial EW^{FD}(\hat{\pi}(\theta), \theta)}{\partial \kappa} - \frac{\partial EW^{PD}(\hat{\pi}(\theta))}{\partial \pi} - \frac{\partial EW^{FD}(\hat{\pi}(\theta), \theta)}{\partial \pi}.$$

This derivative has the sign of the numerator because the denominator is always positive at $\hat{\pi}(\theta)$. Therefore, as

$$\frac{\partial EW^{PD}(\hat{\pi}(\theta))}{\partial \kappa} = 1$$

and

$$\frac{\partial EW^{FD}(\hat{\pi}(\theta), \theta)}{\partial \kappa} = \begin{cases} 1 - \frac{(1-\pi)(1-(1-\pi)L^{-1})}{b} \int_{0}^{\hat{c}(\pi)} \frac{\partial T(c, \theta)}{\partial \kappa} dc & \text{when } 0 \leq \pi \leq \pi_1^{FD} \\
1 - \frac{(1-\pi)(1-(1-\pi)L^{-1})}{b} \int_{0}^{c_1} \frac{\partial T(c, \theta)}{\partial \kappa} dc & \text{when } \pi_1^{FD} \leq \pi \leq 1, \end{cases}$$

the numerator is strictly positive, except at $\hat{\pi}(\theta_0) = 0$ when it is equal to zero. Hence, $\partial \hat{\pi}(\theta)/\partial \kappa \geq 0$.

9.7.2 Increase in the project’s benefit $B$

Applying the Implicit Function Theorem, we can compute

$$\frac{\partial \hat{\pi}(\theta)}{\partial B} = \frac{\partial EW^{FD}(\hat{\pi}(\theta), \theta)}{\partial B} - \frac{\partial EW^{PD}(\hat{\pi}(\theta))}{\partial \pi} - \frac{\partial EW^{FD}(\hat{\pi}(\theta), \theta)}{\partial \pi}.$$

Again, this derivative has the sign of the numerator. Let’s compute

$$\frac{\partial EW^{PD}(\hat{\pi}(\theta))}{\partial B} = \begin{cases} \frac{\pi L^{PD}(\pi)}{b} \Phi(\pi) & \text{when } 0 \leq \pi \leq \pi_1^{PD} \\
\pi & \text{when } \pi_1^{PD} \leq \pi \leq 1 \end{cases}$$

where

$$\Phi(\pi) \equiv \frac{2 + (L - 2)\pi}{L + 1 - \pi}.$$
and

\[
\partial E W^{FD}(\hat{\pi}(\theta), \theta) = \begin{cases}
\frac{\pi c_{FD}^{PD}(\pi)}{b} & \text{when } 0 \leq \pi \leq \pi_{1}^{FD} \\
\frac{\pi c_{NAR}^{FD}(\pi)}{b} & \text{when } \pi_{1}^{FD} \leq \pi \leq \pi_{2}^{FD} \\
\frac{\pi c_{NAR}^{FD}(\pi)}{b} & \text{when } \pi_{2}^{FD} \leq \pi \leq \pi_{3}^{FD} \\
\pi & \text{when } \pi_{3}^{FD} \leq \pi \leq 1.
\end{cases}
\]

When \(\pi_{1}^{PD} \leq \hat{\pi}(\theta) \leq \pi_{2}^{PD} \equiv b/B\),

\[
\frac{\partial E W^{FD}(\hat{\pi}(\theta), \theta)}{\partial B} < \frac{\partial E W^{PD}(\hat{\pi}(\theta))}{\partial B}.
\]

So \(\frac{\partial \hat{\pi}(\theta)}{\partial B} < 0\).

When \(\hat{\pi}(\theta) = 0\),

\[
\frac{\partial E W^{FD}(\hat{\pi}(\theta), \theta)}{\partial B} = \frac{\partial E W^{PD}(\hat{\pi}(\theta))}{\partial B} = 0.
\]

So \(\frac{\partial \hat{\pi}(\theta)}{\partial B} = 0\).

To show what happens when \(\hat{\pi}(\theta)\) is close to zero, we need to analyze how the difference

\[
c(\pi) - c_{R}^{PD}(\pi)\Phi(\pi) \quad (14)
\]

evolves on the locus \((\hat{\pi}(\theta), \theta)\) when \(\theta_{0}\) is close to zero. Let’s study the sign of

\[
\frac{\partial}{\partial \theta} \left[ c_{R}^{FD}(\hat{\pi}(\theta)) - c_{R}^{PD}(\hat{\pi}(\theta))\Phi(\hat{\pi}(\theta)) \right]_{\theta = \theta_{0}}
\]

\[
= \left[ \frac{\partial c_{R}^{FD}}{\partial \pi} - \left( \frac{\partial c_{PD}}{\partial \pi}\Phi(\pi) + c_{PD}\frac{\partial \Phi(\pi)}{\partial \pi} \right) \right] \frac{\partial \hat{\pi}(\theta)}{\partial \theta} + \frac{\partial c_{R}^{PD}}{\partial \theta}.
\]

As

\[
\frac{\partial c_{R}^{FD}}{\partial \pi} \bigg|_{\theta = \theta_{0}} = \frac{B - b + L^{c_{R}^{FD}(\hat{\pi}(\theta))}_{\theta_{0}} - (L - 1)T(c_{R}^{FD}(\hat{\pi}(\theta)), \theta_{0})}{1 + \frac{1}{\theta_{0}}} < \frac{2L [(L + 1)^{2}(B - b) + (L^{2} + 1)b]}{(L + 1)^{4}}
\]

because \(T(c_{R}^{FD}(\hat{\pi}(\theta)), \theta_{0}) > c_{R}^{FD}(\hat{\pi}(\theta))/\theta_{0}\),

\[
\left[ \frac{\partial c_{R}^{FD}}{\partial \pi} - \left( \frac{\partial c_{PD}}{\partial \pi}\Phi(\pi) + c_{PD}\frac{\partial \Phi(\pi)}{\partial \pi} \right) \right] < -\frac{b(L + 1)^{2}L^{2}}{(L + 1)^{4}} < 0.
\]

Now, we analyze

\[
\frac{\partial \hat{\pi}(\theta)}{\partial \theta} \bigg|_{\theta = \theta_{0}} = \frac{\partial E W^{FD}(\hat{\pi}(\theta), \theta)}{\partial \theta} \bigg|_{\theta = \theta_{0}} - \frac{\partial E W^{FD}(\hat{\pi}(\theta), \theta)}{\partial \pi} \bigg|_{\theta = \theta_{0}}.
\]
The derivatives are

\[
\frac{\partial E^{FD}(\hat{\pi}(\theta), \theta)}{\partial \theta} \bigg|_{\theta = \theta_0} = \frac{b(L^2 + 1)^2}{2(L + 1)^4},
\]

and

\[
\frac{\partial E^{PD}(\hat{\pi}(\theta))}{\partial \pi} \bigg|_{\theta = \theta_0} = \frac{L[4(L + 1)B - (5L + 2 - L^2)b]}{2(L + 1)^3}
\]

and

\[
\frac{\partial E^{FD}(\hat{\pi}(\theta), \theta)}{\partial \pi} \bigg|_{\theta = \theta_0} = \frac{1}{b} \left\{ \int_{0}^{\pi_0} b \left[ B - b + T(c, \theta_0) - L(T(c, \theta_0) - \frac{\pi}{\theta_0}) \right] dc \right\}
\]

As \( c/\theta_0 < T(c, \theta_0) \leq b \),

\[
\frac{\partial E^{FD}(\hat{\pi}(\theta), \theta)}{\partial \pi} \bigg|_{\theta = \theta_0} \geq \frac{2L(B - b) + Lb(L^2 + 1)}{(L + 1)^2} - \frac{2L^3b}{(L + 1)^4}
\]

and thus

\[
\frac{\partial \hat{\pi}(\theta)}{\partial \theta} \bigg|_{\theta = \theta_0} \geq \frac{(L^2 + 1)^2}{L^2(L + 1)^2}.
\]

As

\[
\frac{\partial c^{FD}}{\partial \theta} \bigg|_{\theta = \theta_0} = \frac{b(L + 1)^2}{(L + 1)^4},
\]

we finally obtain

\[
\frac{\partial}{\partial \theta} \left[ c^{FD}_R(\hat{\pi}(\theta), \theta) - c^{PD}_R(\hat{\pi}(\theta)) \Phi(\hat{\pi}(\theta)) \right] \bigg|_{\theta = \theta_0} < -\frac{b(L + 1)^2L^2}{(L + 1)^4} \cdot \frac{(L^2 + 1)^2}{L^2(L + 1)^2} + \frac{b(L + 1)^2}{(L + 1)^4} = 0.
\]

So, at \( \hat{\pi}(\theta_0) = 0 \), the derivative of (14) is negative, implying that when \( \theta \) is slightly above \( \theta_0 \), \( \frac{\partial \hat{\pi}(\theta)}{\partial B} < 0 \).

### 9.7.3 Increase in the number of regions \( L \)

Applying the Implicit Function Theorem, we can compute

\[
\frac{\partial \hat{\pi}(\theta)}{\partial L} = \frac{\partial E^{FD}(\hat{\pi}(\theta), \theta)/\partial L - \partial E^{PD}(\hat{\pi}(\theta))/\partial L}{\partial E^{PD}(\hat{\pi}(\theta))/\partial \pi - \partial E^{FD}(\hat{\pi}(\theta), \theta)/\partial \pi}.
\]

Again, this derivative has the sign of the numerator. Let’s compute

\[
\frac{\partial E^{PD}(\hat{\pi}(\theta))}{\partial L} = \begin{cases} \frac{(L-1)(1-\pi)^2(\pi B + (1-\pi)b)^2}{b(L + 1 - \pi)^4} & \text{when } 0 \leq \pi \leq \pi_1^{PD} \\ 0 & \text{when } \pi_1^{PD} \leq \pi \leq 1 \end{cases}
\]
and

$$\frac{\partial E W^{FD}(\hat{\pi}(\theta), \theta)}{\partial L} = \begin{cases} 
\frac{1}{b}(1 - \pi)^L \ln(1 - \pi) \int_0^{c_{RD}(\pi)} (T(c, \theta_0) - \frac{c}{\theta_0}) dc & \text{when } 0 \leq \pi \leq \pi_1^{FD} \\
\frac{1}{b}(1 - \pi)^L \ln(1 - \pi) \left[ \int_0^{c_{RNAR}(\pi)} (b - \frac{c}{\theta_0}) dc - \int_0^{c_1} (b - T(c, \theta_0)) dc \right] & \text{when } \pi_1^{FD} < \pi \leq \pi_2^{FD} \\
\frac{1}{b}(1 - \pi)^L \ln(1 - \pi) \left[ \int_0^{c_2} (b - \frac{c}{\theta_0}) dc - \int_0^{c_1} (b - T(c, \theta_0)) dc \right] & \text{when } \pi_2^{FD} < \pi \leq 1. 
\end{cases}$$

So, when $\hat{\pi}(\theta) = 0$, $\frac{\partial \hat{\pi}(\theta)}{\partial L} > 0$. On the other hand, when $\pi_1^{FD} \leq \hat{\pi}(\theta) \leq 1$, $\frac{\partial \hat{\pi}(\theta)}{\partial L} < 0$. \blacksquare