Partial fiscal decentralisation, local elections, and accountability∗

Susana Peralta†

December 8, 2011

Abstract

Under partial fiscal decentralization, expenditure competencies are decentralized while revenue is not (Brueckner, 2009). We compare this system to the benchmark of full decentralisation, from the viewpoint of political accountability. Local politicians may be public or self interested. They provide a public good whose provision cost is unknown to both voters and the central government. Elections at the end of the first period in office allow the voters to oust the incumbent. We show that selection (i.e., voting out bad incumbents) is improved under partial decentralisation, while discipline (i.e., giving incentives to the bad incumbent) is improved under full decentralisation. Full decentralisation delivers higher voter welfare when the proportion of self interested politicians is low. This result is robust to allowing the central government to set its transfer to the local government optimally. In addition, the introduction of local elections tends to favour full decentralisation, welfare-wise.

∗This paper is the most recent version of “Budget Setting Autonomy and Political Accountability”, written for presentation at the conference “Desenvolvimento Económico Português no Espaço Europeu”, organized by the Portuguese Central Bank in February 2006. I am grateful to the Comité Científico for useful comments. I thank Ross Hickey, Marcel Thum, Tanguy van Ypersele, and seminar participants at the “Competition Among Nation States” workshop (WZB, October 2009) the Louis-Andre Gerard Varet Conference (Marseille, June 2010), and the Lisbon Group on Institutions and Public Policy for useful comments. Financial support from Novaforum and the Fundação para a Ciência e Tecnologia is gratefully acknowledged.

†Nova School of Business and Economics, Campus de Campolide, P-1099-032 Lisboa; CEPR; Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium. Phone: +351 21 3801642. Email: peralta@novasbe.pt
1 Introduction

Fiscal decentralization is a policy objective advocated by international organizations such as the World Bank (The World Bank, 2000) and the OECD (OECD, 2001, 2002), whose application is widespread, both in developed and low-income countries (Epple and Nechyba, 2004; Bardhan, 2002). However, the degree to which it is implemented, both on its expenditure and tax collection aspects, varies a lot. Different institutional arrangements on the sharing of competencies between central (or federal) and local (or state) governments exist (Ter-Minassian, 1997; OECD, 1999). Under these various arrangements, central governments may end up having a strong impact on local policy outcomes, be it because the expenditure and revenue assignments are not equally decentralised (partial decentralisation), or because a given policy area is shared between the two government layers (shared responsibility federalism) (Jametti and Joanis, 2009).¹

The most commonly used measures of decentralisation (e.g., the ratio of local to total government expenditure or revenue) fail to capture this diversity (Blume and Voigt, 2011). Some recent attempts at more comprehensive measures include Blchliger and King (2006) and Stegarescu (2005), making use of the detailed data on local government autonomy compiled in OECD (1999), and both conclude that conventional decentralization measures overestimate the actual degree of devolution of fiscal powers to local governments.² In Germany, for instance, sub-central governments account for half of total tax revenue, but the taxes where the sub-central government may fix either the tax rate, or the tax base, or both weigh a mere 7.3%. In Austria, the figure drops from 28.7% to 3.5% (Stegarescu, 2005). Interestingly, while there has been a general trend towards greater decentralization, the degree of tax revenue autonomy of local governments is largely low and stable. While we lack expenditure data adjusted for the degree of autonomy, a comparison of unadjusted indicators suggests that the degree of decentralisation is in general lower for revenue than for expenditure. This conclusion is shared by Devarajan and Shah (2007) and Shah (2004) who survey decentralisation

¹The distinction between the two concepts is not standard in the literature. Devarajan and Shah (2007) and Shah (2004) use partial decentralisation to refer to any institutional arrangement that leads local governments to bear only part of the responsibility for policy outcomes.

²OECD (1999) classifies local taxes and tax-sharing arrangements into eight categories of decreasing autonomy, from when local governments are free to choose both the tax rate and the tax base, to centrally set tax rate and base. Some “autonomous” local government decisions countries are constrained by an allowable range for local tax rates set by the central government. Hence, if anything, these figures overestimate local governments’ revenue autonomy.
patterns in developing countries, and conclude that partial decentralisation is to blame for lack of accountability of local governments. They go on to argue that “For enhancing accountability it is desirable to match revenue means (the ability to raise revenues from own sources) as closely as possible with expenditure needs at all levels of government.”

The idea that financing expenditures with local taxes leads to greater accountability is explicitly put forward by The World Bank (2003), Shah (2004), Bahl and Martinez-Vazquez (2006), and Devarajan and Shah (2007), and documented empirically by Geys, Heinemann, and Kalb (2010). Using a panel data set of 987 German municipalities for the years 1998, 2002, and 2004, the authors document that the voters’ demand for efficient utilisation of public funds is higher in municipalities that rely less heavily on central government transfers. However, conservativeness in the decentralisation of taxing powers to local governments may result from a presumption of local governments being more prone to capture by special interests (Bardhan and Mookherjee, 2000). Tying the hands of local officials through partial decentralisation may be a means to decrease rent extraction, but it may also hamper the voters’ ability to hold local officials accountable. The study of local accountability is of particular importance in developing economies, where many decentralisation processes are accompanied by local democratisation, creating institutional setups which depart from the standard decentralisation framework of developed countries studied in the literature (Bardhan, 2002).

Our analysis shows that partial decentralisation impacts accountability even when the voters care equally and observe the central government transfer perfectly. We model provision of a local public good with uncertain provision cost, known by the politicians but not by the voters. There are two types of politicians, public-interested and self-interested, unobservable to voters. The local government provides a local public good which is funded with local taxes under full decentralisation, and with a central government transfer under partial decentralisation. The voters observe the public good, the local tax, and the transfer. The corrupt politician may misbehave in two ways, as in Besley and Smart (2007) and Hindriks and Lockwood (2009). She may either reveal her type by providing no public good and extracting maximal

---

3It may also serve to curb the negative detrimental tax competition amongst local governments (Wilson, 1999).

4As put by Bardhan (2002), “even in the relatively few democratic developing countries the institutions of local democracy and mechanisms of political accountability are often weak”, and “In many low-income countries, the decentralization issues discussed (...) are primarily about providing centrally collected tax revenue to lower levels of government, rather than seeking to empower lower levels of government to collect taxes. The focus is on public expenditure assignments, unaccompanied by any significant financial devolution.
rents (the *signaling* rent), or she may pool with the good politician by pretending that the provision cost is high when it is actually low, thus extracting the equivalent (smaller) rent (the *pooling* rent). The former behaviour reveals her type, and the voters accordingly oust her, while the second allows her to be kept in office and delay maximal rent extraction to the second period. Naturally, full decentralisation gives more latitude for local politician to extract separating rents. This explains why separation is more likely to occur under full decentralisation. Welfare-wise, full decentralisation is preferable when the proportion of good politicians is sufficiently high. We also compare the welfare levels under the two regimes with and without local elections and conclude, as expected, that local democracy delivers a higher level of welfare in both regimes. Moreover, the introduction of elections favours the move from partial to full decentralisation more often than the other way around. Finally, we characterise the optimal transfer that a benevolent central government would set under partial decentralisation. We show that the central government is willing to transfer more funds with local elections, and more so under the pooling than under the separating equilibrium.

**Related Literature**

While the literature on decentralisation is abundant, it mostly compares the two polar cases of full centralisation and full decentralisation, i.e., it bundles the expenditure and revenue assignment. The fact that this is at odds with the actual way in which devolution to local populations is tackled in most countries has fostered an emergent literature on the theory of partial decentralisation. Schwager (1999a,b) introduces the concept of *administrative federalism*, under which the central government sets quality standards for the public projects, and the local ones are choose which projects to carry out. Centralised decision favours one local government at the expense of the others, while decentralised decision fails to take account of inter-regional spill-overs. Administrative federalism may be preferred to both regimes if the spillovers are not too strong. Wilson and Janeba (2005) and Hatfield and Padró i Miquel (2008) model partial decentralisation as the devolution of a subset of a continuum of tasks to the local governments, and show that this may be desirable in order to appropriately balance vertical and fiscal externalities stemming from tax competition or to avoid excessive capital externalities stemming from inter-regional spillovers and preference heterogeneity, or on Tiebout (1956)’ analogy between local governments and profit maximising firms. For recent contributions using political economy arguments, see, inter alia, Bardhan and Mookherjee (2000), Besharov (2001), Besley and Coate (2003), Belleflamme and Hindriks (2005), Hindriks and Lockwood (2009).
taxation which hampers capital accumulation, respectively. Joanis (2008) and Jametti and Joanis (2010), in turn, model shared responsibility federalism, in the sense that the public good is jointly provided by both government levels. Shared responsibility may optimally obtain from a trade-off between complementarities in the inputs of the two government layers and diminished accountability.

Brueckner (2009) models decentralised expenditure and centralised revenue (as we do in this paper), in a Tiebout (1956) model featuring heterogeneous consumers and profit maximising developers who supply housing competitively. Since the local public good in each locality is an increasing function of costly effort by the local officials, there may be heterogeneity in public good provision even with a common budget. The local government maximises the developers’ profits, net of public good provision and effort cost. Brueckner (2009)’s main result is that welfare improves when moving from full centralisation (i.e., uniform provision) to partial decentralisation, and from this to full decentralisation, since in each case one is allowing for a greater variety of local public goods, hence a better preference matching. When Brueckner (2009) introduces Leviathan local officials who favour high local budgets, full decentralisation leads to over-provision of the local public good, and partial decentralisation may be optimal, given its ability to tame the Leviathan. As we shall see, taking the more balanced view that rent-motivated politicians co-exist with purely benevolent ones mitigates this conclusion.6

A few papers analyse the impact of local democratisation on accountability, defined as the ability of local governments to target the poorest households in the community. Foster and Rosenzweig (2004) compare a two-party political competition equilibrium with the outcome that would arise under “aristocratic government”, i.e. a system in which the local government is a perfect agent of the landed households. Using Indian data, where considerable cross-state variation about the degree of democratisation of local governments exist, the authors show that local democratization serves the interests of the poor.7

The remaining of the paper is organized as follows. The next section presents the model and equilibrium under each regime. Section 3 compares

---

6This is reminiscent of Besley and Smart (2007)’s conclusion that tax competition is not necessarily welfare improving in the presence of both types of politicians, contrary to Edwards and Keen (1996)’s original insight.

voters’ expected welfare under the two regimes. Section 4 analyses the relative effectiveness of elections in boosting welfare under the two regimes. The optimal central government transfer is analysed in Section 5, and Section 6 concludes. All proofs are relegated to the Appendix.

2 The model

Local governments provide a local public good $G$, whose provision cost is uncertain. The cost $\theta$ can be high ($\theta = H$) or low ($\theta = L$), with $H > L$. We follow Hindriks and Lockwood (2009) in assuming that the probability of a high cost, $q$, is greater than $1/2$.\footnote{This assumption rules out the hybrid equilibrium derived by Besley and Smart (2007), which Lockwood (2005) has shown not to survive the Cho-Kreps stability criterion.} Voters derive utility from the consumption of the public good and dislike high local taxes. The local revenue is denoted $x \in [0, X]$, i.e., there is a no-debt constraint and a maximum local tax. The utility function is

$$W(G, x) = G - C(x),$$

where $C(x)$ is a strictly increasing and strictly convex function (an analogous utility function is used by Besley and Smart (2007)). We define $G^H$ and $G^L$ as follows

$$G^\theta = \arg \max G - C(\theta G)$$

and let $x^\theta = \theta G^\theta$. Convexity of the $C$ function ensures that $x^L > x^H$, i.e., the public sector is optimally larger when the provision cost is low, and $G^L > G^H$, i.e., the public good decreases with the provision cost. Naturally, the maximum local revenue covers the provision cost in all circumstances, i.e., $X \geq x^L$.

Local politicians may be of one of two types: good ($\eta = g$) or bad ($\eta = b$). Good politicians always pursue the interests of the electorate, while bad ones care about the rents $r$ they manage to extract. The proportion of good politicians is $\pi$, which may be interpreted as the quality of the polity.

The timing of the model is as follows. There are two periods. In the first period, the incumbent chooses the policy to implement. At the end of the first period, an election takes place, and the incumbent may be re-elected or ousted, in which case nature randomly selects a politician to be in office in the second period. The game ends at the end of the second period, with no further elections. The future is discounted by $\beta \in (0, 1)$. For future reference, let $G_\eta$ (resp., $x_\eta$) denote the public good level (resp., local tax) decided by a politician of type $\eta = g, b$ in periods $t = 1, 2$.\footnote{This assumption rules out the hybrid equilibrium derived by Besley and Smart (2007), which Lockwood (2005) has shown not to survive the Cho-Kreps stability criterion.}
The local revenue $x$ stems from local taxes set by the local government under full decentralisation, and from central government transfers under partial decentralisation. The central government does not know the true provision cost and sets a transfer $\tilde{x}$. This transfer need not, but may be funded by local taxes. The transfer is at most equal to the maximum optimal local tax, i.e.,

$$\tilde{x} \leq x^L \quad (2)$$

We analyse in Section 5 the consequences of allowing the central government to optimise over this transfer and show that none of the results derived below changes.

In what follows, we make the natural assumption that the maximum local tax under full decentralisation is bounded, and for technical convenience we set the upper bound such that

Assumption 1

$$C(X) - C(\tilde{x}) \leq \frac{1 - q}{\beta + q} \left( G^H - C(x^H) - (\tilde{G}^H - C(\tilde{x})) \right)$$

Where $\tilde{G}^g$ is the public good provided by a good politician under partial decentralisation when the provision cost is high and is defined in (4) below.

One may consider further restricting the local tax raising autonomy. The voters know that no local tax above $x^L$ is optimal, so it is conceivable that they would not pay a higher amount. On the other hand, it may happen that the central government adopts a precautionary principle of always allowing the highest possible optimum public good provision. We shall sometimes restrict our attention to this case, i.e. when the maximum local tax is set to its minimum, and the central government transfer to its maximum, and they are both equal to the maximum optimum total provision cost.

Assumption 2

$$X = \tilde{x} = x^L$$

Assumption 2 is a special case of Assumption 1.

2.1 Preliminary results

Voters observe the policy implemented by the incumbent, and the central government transfer, if any, and use Baye’s rule to compute the posterior probability $\Pi(G)$ that she is good, given the observed record. If $\Pi(G)$ is
greater than the probability that the randomly selected official is good, \( \pi \), the incumbent is reelected.\(^9\) Otherwise, she is voted out of office.

We now describe the politicians’ behaviour in each decentralisation regime.

**Full decentralisation** A good politician implements

\[
G_{gt} = G^{\theta}, \quad x_{gt} = x^{\theta}, \ t = 1, 2
\]

in both periods. Hence, any other policy vector perfectly signals a bad politician. Proceeding by backwards induction, it is straightforward that a bad politician extracts maximal rents in the second period since she is no longer concerned by re-election, i.e.

\[
x_{b_2} = X, \ G_{b_2} = 0
\]

In the first period, she either pools with the good type by choosing \( G = G^{\theta}, x = x^{\theta} \), for \( \theta = L \) or \( \theta = H \), or she behaves as in the second period.\(^{10}\) If the provision cost is \( \theta = H \), the bad incumbent gets a negative rent of \((L - H)G^L\) by providing \( G^L \), no rent by choosing \( G^H \) and the maximal rent when she provides no public good. The two former strategies are dominated by the latter.\(^{11}\) A similar dominance argument eliminates the provision of \( G^L \) when the cost is \( \theta = L \). The politician either provides no public good

\[
G_{b_1} = 0, \ x_{b_1} = X,
\]

extracting a rent of \( r_{b_1} = X \), or

\[
G_{b_1} = G^H, \ x_{b_1} = x^H,
\]

with a rent of \( r_{b_1} = (H - L)G^H \). Let the two possible rent values under full decentralisation be denoted

\[
R^f = X, \ r^f = (H - L)G^H,
\]

where \( R^f \) refers to maximal or signaling rents, while \( r^f \) refers to pooling rents.

We may summarize the behavior of the bad politician as follows. She never provides the optimal amount of the public good. If the cost is high,

\(^9\)It will become clear in the subsequent that it suffices to write \( \Pi \) with the single argument \( G \).

\(^{10}\)Any other policy vector is a perfect signal of her type and is dominated by \( (G = 0, x = X) \).

\(^{11}\)To see why, note that \( G = 0 \) yields a payoff of \( X \), whereas the other two actions give at most \((L - H)G_L + \beta X\) and \( \beta X \), respectively.
she extracts maximal rents and is voted out of office. When the cost is low, she may use one of two possible strategies. She either separates from the good type by providing no public good and is voted out, or she pools with the good type under a high cost, thus keeping her re-election chances. In so doing, she foregoes current rents in return for a probability of reelection (hence, future rents).

From the above, it is immediate that $G^L$ can only be implemented by the good politician, hence $\Pi(G^L) = 1 > \pi$, and the voters reelect the incumbent. Also, $\Pi(0) = 0 < \pi$, hence the incumbent is voted out of office. Finally, we have that

$$\Pi(G_H) = \frac{\pi q}{\pi q + (1 - \pi)(1 - q)}$$

when the bad politician pools with the good one.

**Partial decentralisation** We now turn to partial decentralisation. The good politician chooses $G$ as to maximise (1), under the budget constraint $\theta G \leq \tilde{x}$. Hence, she provides

$$G_{g_t} = \tilde{G}^{\theta} = \frac{\tilde{x}}{\theta}, \ t = 1, 2$$

in both periods. The bad politician provides

$$G_{b_2} = 0$$

and collects the central government transfer as a rent, $r_{b_2} = \tilde{x}$, in the second period. As regards the first period, her behaviour is analogous to the case of full decentralisation. When the provision cost is high she chooses

$$G_{b_1} = 0,$$

thus extracting $r_{b_1} = \tilde{x}$, while when the provision cost is low, she may opt for

$$G_{b_1} = 0,$$

with a rent of $r_{b_1} = \tilde{x}$ or

$$G_{b_1} = \tilde{G}_H,$$

with a rent of $r_{b_1} = \tilde{x} \frac{H - L}{H}$. Let signaling and pooling rents under partial decentralisation be denoted

$$R^p = \tilde{x}, \ r^p = \tilde{x} \frac{H - L}{H}$$
Voters do not reelect when \( G = 0 \) and reelect with probability one if \( G = \hat{G}^L \). Conditional on observing \( G = \hat{G}^H \), and provided that the bad type chooses the pooling strategy, the probability that the incumbent is good is given by \( \Pi(\hat{G}^H) = \Pi(G^H) \), as defined in (3).

Using the facts that \( \hat{x} \leq x^L \), and \( X \geq x^L \), it is immediate that signaling rents are not lower under full decentralisation,

\[
R^f = X \geq \hat{x} = R^p
\]

while pooling rents are higher (resp., lower) if \( x^H > \hat{x} \). Indeed, we may write

\[
r^f = (H - L)G^H = x^H \frac{H - L}{H},
\]

so that \( r^f / r^p = x^H / \hat{x} \).

Before we proceed, it is useful to introduce the following notation. Let

\[
\gamma^d = \frac{R^d - r^d}{R^d} \in (0,1), \; d = f,p
\]

denote the cost from pooling with the good type, or the discipline cost. The discipline cost is equal to

\[
\gamma^p = 1 - \frac{H - L}{H} = \frac{L}{H}
\]

under partial, and

\[
\gamma^f = 1 - \frac{(H - L) x^H}{X}
\]

under full decentralisation. Hence, the discipline cost is higher under full decentralisation. This is a straightforward consequence of the fact that the total budget is exogenous for the politician under partial decentralisation, and when deciding to signal her type, she provides no public goods and diverts the fixed transfer of the central state. When she can set local taxes, in turn, she signals her type by both providing no public good and setting the maximum possible local tax, which she captures as a rent. It is important to notice that this is unrelated to the amount of information about the central government transfer that the voters have.

2.2 A benchmark: no elections

In this setup, the relationship between the voters and the local politician is a principal-agent one in which the principle has one single instrument – elections – to provide incentives to the agent. In order to grasp the usefulness of this instrument under the two decentralisation scenarios, it is useful to characterise the politicians’ behaviour in the absence of elections. The existence of elections does not change the behaviour of the good politicians, whose
objectives are perfectly aligned with those of the principal. Bad politicians, in turn, always provide no public good and extract maximal rents, that is, they behave in the first period in office exactly as in the second.

In the context of a low-income country Foster and Rosenzweig (2004), compare local democracy with the benchmark of an “aristocratic government”, that is, a local government which serves the interest of the local elite. Contrary to Foster and Rosenzweig (2004), in our set-up all the voters agree on the welfare-maximising policy. The natural benchmark is local decisions taken by a randomly selected public official, who will behave according to the voters’ interest with probability $\pi$, and care for rents otherwise.

### 2.3 Equilibrium

We are now in a position to present the equilibrium of the game. Given that the good politician does not behave strategically in this model, as in Besley and Smart (2007), the equilibrium depends on the strategic interaction between voters and the bad incumbent. There are two possible equilibria. In the pooling equilibrium, the bad incumbent facing a low cost mimics the good one facing a high provision cost, and is re-elected by the voters, while in the separating equilibrium, the bad incumbent provides no public good and is ousted by the voters.

Not surprisingly, the equilibrium has a similar structure under both full and partial decentralisation.\(^{12}\)

**Lemma 1** A pooling equilibrium arises if and only if the discipline cost is lower than the discount factor, i.e.,

$$\gamma^d \leq \beta, \ d = f, p$$

Otherwise, the equilibrium is separating.

The equilibrium behaviour is ruled by the trade-off between current and expected (discounted) future rents. When pooling rents are too low, separation occurs. When they are high, pooling occurs at equilibrium. Given that the probability of the high cost is sufficiently high, voters are willing to believe that $G^H$ (or $\tilde{G}^H$) was implemented by a good politician and they reelect the incumbent with probability one.\(^{13}\)

\(^{12}\)The equilibrium is similar to the one described in Besley and Smart (2007)’s Lemma 1.

\(^{13}\)In the opposite scenario of $q < 1/2$, Besley and Smart (2007) construct a hybrid equilibrium where the voters randomise between re-electing the incumbent or not, and the bad incumbent randomises between the pooling and the separating strategies. Lockwood

11
Lemma 1 creates three relevant intervals for the discount factor. To fix ideas, let us say that \( \beta \) is low when \( \beta < \gamma^p \), high when \( \beta > \gamma^f \), and intermediate otherwise, i.e., \( \gamma^p < \beta < \gamma^f \). We may state the following Corollary.

**Corollary 1** When \( \beta \) is low (resp., high), there is a separating (resp., pooling) equilibrium in both regimes. Otherwise, partial decentralisation leads to a pooling equilibrium, while full decentralisation leads to a separating equilibrium.

While both regimes are equivalent in the equilibrium structure, they do not yield the same payoffs. Given that the discipline cost is higher under full decentralisation, the separating equilibrium is more likely under this regime. Put otherwise, intermediatively myopic politicians are disciplined under partial decentralisation, and signal their type under full decentralisation.

One implication of this results is that, all else equal, the turnover of local politicians should increase with the degree of decentralisation.

We now turn to the welfare comparison of the two regimes. The welfare difference comes from the different incentives and the different values of rent extraction under the two regimes.

### 3 Welfare comparison

In this section, we compute the expected welfare at the beginning of the first period for both regimes. We simplify matters by supposing that the central government transfer is funded by local taxes. Alternatively, one may interpret the utility function as reflecting the trade off between public good provision and cost which a benevolent planner would take into account for the purpose of comparing the two regimes.

As usual in this type of models, expected voter welfare can be decomposed into into three parts (Besley and Smart, 2007).

The first is the *no-election utility*, that is, the one we would obtain in the absence of elections. It is equal to

\[
(1 + \beta) \left( \pi \left( q\tilde{G}^H + (1 - q)\tilde{G}^L \right) - C(\tilde{x}) \right)
\]

(5)

(2005) has shown that this equilibrium does not survive the Cho-Kreps stability criterion, since the good politician has an incentive to distort public good provision when the cost is high, in order to signal her type and avoid the likely outcome of being replaced by a bad politician.
under partial, and
\[
(1 + \beta) \left[ \pi \left( q \left( G^H - C(x^H) \right) + (1 - q) \left( G^L - C(x^L) \right) \right) - (1 - \pi)C(X) \right]
\] (6)
under full decentralisation. The difference between full and partial decentralisation is ambiguous. On the one hand, partial decentralisation leads to lower rent diversion on behalf of the bad incumbent and, on the other hand, it entails a misallocation cost, due to the good politician not implementing the first-best public good level. Naturally, the negative effect dominates when the proportion of good politicians is sufficiently high, in which case full decentralisation is preferable. This is stated in the next Proposition.

**Proposition 1** Suppose that local officials are selected randomly. Then, full decentralisation outperforms partial decentralisation provided that the proportion of good politicians is sufficiently high.

The impact of elections on voter welfare comes about in two distinct effects. On the one hand, the bad incumbent may refrain from extracting maximal rents in the first period (*discipline effect*). On the other hand, the bad incumbent may signal her type by extracting maximal rents, be voted out and, with some probability, be replaced by a good politician in the second period (*selection effect*).

The *discipline effect* occurs with a pooling equilibrium, and is given by
\[
(1 - \pi)(1 - q)\tilde{G}^H
\] (7)
under partial, and
\[
(1 - \pi)(1 - q) \left( G^H - C(x^H) + C(X) \right)
\] (8)
under full decentralisation. When both regimes lead to a pooling equilibrium, the gain from discipline is higher under full decentralisation. To see why, use \( X \geq x^L \geq \tilde{x} \) to obtain
\[
G^H - C(x^H) + C(X) - \tilde{G}^H \geq C(x^L) - \frac{x^L}{H} - \left( C(x^H) - \frac{x^H}{H} \right) > 0,
\]
where the last inequality follows from the fact that \( C' - 1/H > 0 \), together with \( x^L > x^H \). This is a consequence of the difference between signaling and pooling rents, which is amplified under full decentralisation. Naturally, when pooling only occurs under partial decentralisation, this regime dominates full decentralisation.

The *selection effect* arises when the equilibrium is separating, and is equal to

\[13\]
\[ \beta(1 - \pi)\pi(1 - q)\bar{G}^L \] 
under partial, and

\[ \beta(1 - \pi)\pi(1 - q)\left( G^L - C(x^L) + C(X) \right) \] 
under full decentralisation.

The gain from selecting bad incumbents is always higher under full decentralisation. When separation only occurs under this regime, this is obviously the case. When it occurs under both regimes, the fact that \( x^L \geq x \geq \bar{x} \) implies that

\[ G^L - C(x^L) + C(X) - \bar{G}^L > 0 \]

The fact that the discipline and selection gains from having elections are higher under full decentralisation (with the exception of when partial decentralisation provides discipline and full decentralisation does not) is a consequence of the very high rents that a bad incumbent may extract when she is empowered to raise taxes.

As one would expect, incentive and selection only matter to the extent that some bad politicians exist. If \( \pi = 1 \), the existence of elections has no impact on welfare. Obviously, when the agent shares the principal’s objective with probability one, the incentive instrument used by the principal has no impact on welfare.

We wrap up the welfare comparison in the following Proposition.

**Proposition 2** Suppose that local officials are elected. When \( \beta \) is high, if Assumption 1 holds, full decentralisation outperforms partial decentralisation. When \( \beta \) is low or intermediate, full decentralisation outperforms partial decentralisation provided that the proportion of good politicians is sufficiently high.

Propositions 1 and 2 together show that full decentralisation is desirable when the proportion of good politicians is high, irrespective of the existence of local elections. Indeed, without elections, partial decentralisation is desirable because it leads to lower rent diversion, and it is undesirable because it entails a misallocation cost; the former stems from the presence of bad politicians, and the latter from the good ones. The discipline and selection effects of elections naturally have a higher impact to the extent that there is a likelihood that the incumbent is a bad politician, so adding them up keeps the result unchanged. As in Brueckner (2009), full decentralisation is the best regime in the absence of rent-motivated politicians. However, when
all the politicians are rent motivated, then partial decentralisation always dominates. This is not the case in Brueckner (2009), where this only happens if local heterogeneity is not too high and the Leviathan motives of the politicians are strong.

When $\beta$ is high, both regimes lead to a pooling equilibrium. In this case, while we are sure that full decentralisation is the preferred regime when the proportion of good politicians is sufficiently high, we cannot rank the welfare of the two regimes, in general, when there are many bad politicians. Assumption 1, by imposing an upper bound on the maximum rent extraction, ensures that full decentralisation is better than partial. Dropping Assumption 1 and letting $X$ become sufficiently high, creates a range of $\pi$ close to 0 where partial decentralisation outperforms full decentralisation.

Importantly, neither Proposition 1, nor Proposition 2 exclude the possibility that the cut-off $\pi$ above which full decentralisation is preferable is equal to 0. We now provide conditions for this to arise.

3.1 A stronger case for full decentralisation

From the above, it is apparent that the main cost of full decentralisation stems from the excessive rent diversion that may arise when the politician has the necessary autonomy to launch local taxes. As it turns out, restraining this ability by imposing Assumption 2 greatly improves the case for full decentralisation, as stated in the next Proposition.

**Proposition 3** Suppose that Assumption 2 holds. Then, full centralisation always outperforms partial decentralisation, except when local officials are elected, $\beta$ is intermediate, and the proportion of good politicians is low.

Under Assumption 2, maximum rent diversion is equalised in the two decentralisation regimes, which eliminates the main cost of full decentralisation. The only scenario in which it may pay to implement partial decentralisation is when the discipline is improved under this regime, i.e., when $\beta$ is intermediate, since in this case a bad politician extracts maximal rents under full, but not under partial decentralisation.

4 Local elections and full decentralisation: substitutes or complements?

Local elections trivially increase voter welfare under both regimes, for any $\pi > 0$. However, they do not do so equally. We now analyse the extent
to which local elections change the trade-off between the two regimes. This is an important policy question, since the debate about decentralisation in developing and transition economies usually goes hand in hand with the one on democratisation (Bardhan, 2002) and it is generally acknowledged that one of the caveats of decentralisation in developing economies is capture by self-interested politicians or local elites (Bardhan and Mookherjee, 2006).

Propositions 1 and 2 show that, with or without local elections, there is a threshold $\pi \geq 0$ above which full decentralisation is preferable, welfare-wise, to partial decentralisation. We shall say that full decentralisation and local democratisation are complements when introducing elections decreases this threshold, thus strengthening the case for full decentralisation. When $\pi$ is between the two thresholds, introducing elections makes a switch from partial to full decentralisation become desirable, welfare-wise. Conversely, when introducing elections increases the threshold, we say that full decentralisation and local democratisation are substitutes. In this case, the introduction of local elections may necessitate a move from full to partial decentralisation, in order to increase voter welfare.

The answer depends on the incentives put in place by the two regimes. When both full and partial decentralisation lead to the same (pooling or separating) equilibrium, democratisation and decentralisation are complements.

Proposition 4 When the discount factor is either low or high, full decentralisation and local elections are complements.

The message is clear. Even if partial decentralisation may be the best regime without local elections, due to an excessive prevalence of rent-motivated local officials, it pays to switch to full decentralisation once the democratisation process reaches the local government tier. This is because elections put in place the right incentives to limit the harmfulness of the excessive number of bad politicians.

This need not be the case, however, when the discount factor is intermediate and full decentralisation leads to a separating equilibrium, while partial decentralisation leads to a pooling one. Then, depending on the parameters, it may happen that partial decentralisation dominates full decentralisation more often (i.e., for a wider range of $\pi$) with than without elections. This possibility is the object of the following Proposition. Proposition 3 states that, that under Assumption 2, full decentralisation is the preferred regime in the absence of local elections, irrespective of the proportion of good politicians. It turns out that introducing elections in this case may change the welfare ranking of the two regimes and make partial decentralisation preferable to full decentralisation, as shown in our next result.
Proposition 5 Under Assumption 2, when the discount factor is intermediate, full decentralisation and local elections are substitutes.

Recall that when the discount factor is between the discipline costs of the two regimes, partial decentralisation leads to a pooling equilibrium whereas full decentralisation leads to a separating one. The intuition for this result is better grasped if one supposes that there are almost no good politicians, i.e., $\pi \approx 0$. Then, there are no gains from selection because of the very low probability that nature selects a good politician to substitute the bad one who is ousted at the end of the first period, hence introducing elections under full decentralisation has a negligible impact on welfare. However, the gains from discipline are important, since almost all politicians are rent-motivated, so it pays to introduce elections in the partial decentralisation regime. In this case, democratisation and decentralisation are substitutes when it comes to voter welfare, and it is conceivable that welfare improves by the simultaneous move towards local democracy and reduced local autonomy.

It is important to highlight that Proposition 5 shows that when $\beta$ is intermediate, it may happen that local elections changes the welfare ranking of the two regimes in favour of partial decentralisation, not that it always does so. By contrast, Proposition 4 states that for low and high $\beta$ the introduction of elections always changes the welfare ranking in favour of full decentralisation. One may interpret our results as suggesting that local democratisation is more likely to favour the move from partial to full decentralisation than the other way around.

5 The optimal central government transfer under partial decentralisation

Even with the limited information at its disposal, the central government may arguably do better than just transferring an arbitrary amount to the local government. The natural question is then whether an optimally set central government transfer may improve the case for partial decentralisation. We take up this exercise in this Section. Following Brueckner (2009), we optimise the transfer to the local government that a benevolent central government would choose. Straightforward algebra allows us to establish that the optimal central government transfer, $\tilde{x}^o$, respects

$$C''(\tilde{x}^o) = \pi \left( \frac{q}{H} + \frac{1-q}{L} \right)$$

in the absence of elections,
in the pooling equilibrium, that is, when $\beta < \gamma^p$, and

$$C'(\tilde{x}_{\text{pool}}) = \frac{\pi q + (1 - \pi)(1 - q)}{H} + \frac{\pi(1 - q)}{L}$$

(12)

in the separating equilibrium, that is, for $\beta > \gamma^p$.

The optimal transfer equalises the marginal cost of public funds to its expected marginal benefit. Without elections, only the good politicians use the transfer for public good provision, hence the expected benefit is equal to the expected value of the public good, weighed by the probability that a good politician is in office. Adding discipline and selection increases the likelihood that the transfer be used to fund the public good when the provision cost is low (probability of $(1 - q)$). In the pooling equilibrium, the benefit is immediate, whereas in the separating one, it is delayed until the subsequent period.

Comparing the above conditions allows us to state the following proposition.

**Proposition 6** In the partial decentralisation regime, the optimal central government transfer is higher under the pooling equilibrium than under the separating one, and higher in both than without local elections.

The existence of local elections ensures that the funds received from the central government are more likely to be used in the interest of the voters, hence it is optimal to increase them. Recall that the optimal transfer results from the trade-off between the marginal cost of public funds and its expected marginal benefit. Given that the likelihood that the transfer funds the public good instead of rents is higher in the pooling than in the separating equilibrium, the transfer is higher in the former. This is because the expected benefit is delayed until the next period under the separating equilibrium, and this type of equilibrium exists precisely when the future is heavily discounted.

Finally, it is noteworthy the optimal central government transfer is in all cases smaller than the maximum local tax set by a benevolent politician ($x^L$), but it need not be greater than the smallest local tax ($x^H$). To see why, note that even in the best possible scenario in which there are only good politicians, they provide at best $\tilde{x}/L$, with probability $(1 - q)$. Hence, the expected marginal benefit of the transfer is in all cases lower than $1/L$.

The fact that the optimal transfer is lower than $x^L$ is consistent with our assumption on $\tilde{x}$, as stated in (2). Hence, all the results above carry through
when the central government transfer is set to its optimal level, implying that the case for full decentralisation is not weakened.

6 Concluding remarks

This paper sheds light on the merits of partial decentralisation from the viewpoint of political accountability. We use a political agency model to compare full decentralisation, where the local public good is funded with local taxes set by the local government, and partial decentralisation, where the local public good is funded with a transfer from the central government.

As usual in this type of models, there is a conflict between discipline (i.e., the bad incumbent refraining from rent extraction) and selection (i.e., the bad incumbent signaling his type via rent extraction). We show that full decentralisation improves selection, at the expense of discipline, when compared to partial decentralisation. This implies that one should observe more turnover in a fully decentralised regime.

In general, full decentralisation delivers a higher welfare level when the proportion of public-interested local politicians is sufficiently high, irrespective of whether local elections allow the voters to oust a seemingly rent-motivated incumbent at the end of the first period in office. We also treat the related question of whether a move towards local democracy should be accompanied by more or less local autonomy, from the viewpoint of voter welfare. We show that democratisation may favour either of the two regimes, depending on how myopic are the politicians, although it is more likely to favour full decentralisation.

Finally, we compute the optimal transfer of the central government under partial decentralisation and show that this possibility does not change any of the results obtained with an exogenous transfer. In other words, allowing the central government to optimise over the transfer does not weaken the case for full decentralisation.

This paper relates the different degrees of decentralisation of revenue and expenditure assignments to local government with local accountability. Differently from previous contributions, which mostly take the viewpoint of the ability of decentralisation to target the poorest in the community, we define accountability as decreased rent diversion, in the absence of ideological conflict. While this paper analyses one important aspect of decentralisation design, there are numerous features which need to be taken into account by the literature in order to better understand the impact of the actual design of decentralisation on political outcomes at the local level.
Appendix

Proof of Lemma 1
We look at each type of equilibrium in turn. In the pooling equilibrium, we obtain from (3) that $\Pi(G_H) > \pi$, since $q > 1/2$. Hence, the voters are willing to re-elect the incumbent who provides $G_H$. The bad incumbent is willing to provide $G_H$ if and only if

$$r^d + \beta R^d \geq R^d \iff \beta \geq \gamma^d, \ d = f, p$$

In the separating equilibrium, the voters are willing to re-elect the incumbent who provides $G_g$, since it is for sure the good type. Hence, the bad type gets $R^d$ if she does not deviate, or $r^d + \beta X$ if she deviates by providing $G_g$. Hence, she is willing to separate if and only if $\beta < \gamma^d, \ d = f, p$. □

Proof of Proposition 1
Using (5) and (6), one may write the expected welfare difference as

$$\Delta W_o(\pi) = (1 + \beta) \left[ \pi \left( q \left( G^H - C(x^H) - \tilde{G}^H + C(\tilde{x}) \right) + (1 - q) \left( G^L - C(x^L) - \tilde{G}^L + C(\tilde{x}) \right) \right) + (1 + \beta) \left[ (1 - \pi) \left( C(\tilde{x}) - C(X) \right) \right] \right]$$

Notice that $\left( G^\theta - C(x^\theta) - \tilde{G}^\theta + C(\tilde{x}) \right) > 0$, given that $G^\theta$ is the optimal provision level when the cost is $\theta$. Hence, $\Delta W_o$ is increasing in $\pi$, since

$$\kappa_1 = \frac{d\Delta W_o}{d\pi} = (1 + \beta) \left( q \left( G^H - C(x^H) - \tilde{G}^H + C(\tilde{x}) \right) + (1 - q) \left( G^L - C(x^L) - \tilde{G}^L + C(\tilde{x}) \right) \right) + (1 + \beta) \left( (1 - \pi) \left( C(\tilde{x}) - C(X) \right) \right)$$

is positive.

One may thus write $\Delta W_o(\pi) = \kappa_1 \pi - \kappa_2$, with $\kappa_1 > 0$, and $\kappa_2 = (1 + \beta) \left( C(X) - C(\tilde{x}) \right) \geq 0$.

Then,

$$\Delta W_o(0) = -\kappa_2 \leq 0$$

$$\Delta W_o(1) = \kappa_1 - \kappa_2 = \left( q \left( G^H - C(x^H) - \tilde{G}^H + C(\tilde{x}) \right) + (1 - q) \left( G^L - C(x^L) - \tilde{G}^L + C(\tilde{x}) \right) \right) > 0$$

Hence, there exists a $\hat{\pi}_o = \frac{\kappa_2}{\kappa_1} \in [0, 1)$ such that $\Delta W_o(\pi) \leq 0$ when $\pi \leq \hat{\pi}_o$, and $\Delta W_o(\pi) > 0$ when $\pi > \hat{\pi}_o$. □
Proof of Proposition 2
We look at each range of $\beta$ separately.

(i) When $\beta$ is low, the equilibrium is separating under both regimes. Using (5), (6), (9), and (10), the expected welfare difference is equal to

\[
\Delta W_{ss}(\pi) = \Delta W_o(\pi) + \beta(1 - \pi)\pi(1 - q) \left(G^L - C(x^L) + C(X) - \tilde{G}^L\right)
\]

\[= \kappa_1 \pi - \kappa_2 + \kappa_3 \left(\pi - \pi^2\right)
\]

with $\kappa_3 = \beta(1 - q) \left(G^L - C(x^L) + C(X) - \tilde{G}^L\right) > 0$, where the last inequality follows from the fact that $\tilde{x} \leq x^L \leq X$.

$\Delta W_{ss}(\pi)$ is a quadratic concave function of $\pi$, with $\Delta W_{ss}(0) = \Delta W_o(0) \leq 0$, and $\Delta W_{ss}(1) = \Delta W_o(1) > 0$. Moreover,

\[
\frac{d\Delta W_{ss}(\pi)}{d\pi}\bigg|_{\pi=0} = \kappa_1 + \kappa_3 > 0
\]

Hence, there exists a $\hat{\pi}_{ss} \in (0, 1)$ such that $\Delta W_{ss}(\pi) \leq 0$ when $\pi \leq \hat{\pi}_{ss}$, and $\Delta W_{ss}(\pi) > 0$ when $\pi > \hat{\pi}_{ss}$.

(ii) When $\beta$ is intermediate, the equilibrium is separating under full, and pooling under partial decentralisation. Using (5), (6), (7), and (10), the expected welfare difference is equal to

\[
\Delta W_{sp}(\pi) = \Delta W_o(\pi) + \beta(1 - \pi)\pi(1 - q) \left(G^L - C(x^L) + C(X) - \tilde{G}^H\right)
\]

\[= \kappa_1 \pi - \kappa_2 + \kappa_4 \left(\pi - 1\right) + \kappa_5 \left(\pi - \pi^2\right)
\]

where $\kappa_4 = (1 - q)\tilde{G}^H > 0$, and $\kappa_5 = \beta(1 - q) \left(G^L - C(x^L) + C(X)\right) > 0$. $\Delta W_{sp}(\pi)$ is a quadratic concave function of $\pi$, with $\Delta W_{sp}(0) = \Delta W_o(0) - (1 - q)\tilde{G}^H < 0$, and $\Delta W_{sp}(1) = \Delta W_o(1) > 0$. Moreover,

\[
\frac{d\Delta W_{sp}(\pi)}{d\pi}\bigg|_{\pi=0} = \kappa_1 + \kappa_4 + \kappa_5 > 0
\]

Hence, there exists a $\hat{\pi}_{sp} \in (0, 1)$ such that $\Delta W_{sp}(\pi) \leq 0$ when $\pi \leq \hat{\pi}_{sp}$, and $\Delta W_{sp}(\pi) > 0$ when $\pi > \hat{\pi}_{sp}$. □
When $\beta$ is high, the equilibrium is pooling under both regimes. Using (5), (6), (7), and (8), the expected welfare difference is equal to

$$\Delta W_{pp}(\pi) = \Delta W_o(\pi) + (1-\pi)(1-q) \left( G^H - C(x^H) + C(X) - \tilde{G}^H \right) = \kappa_1 \pi - \kappa_2 + \kappa_6 (1-\pi)$$

where $\kappa_6 = (1-q) \left( G^H - C(x^H) + C(X) - \tilde{G}^H \right) > 0$.

Firstly note that, $\Delta W_{pp}(\pi)$ is an increasing function of $\pi$, since

$$\frac{d\Delta W_{pp}(\pi)}{d\pi} = \kappa_1 - \kappa_6 = [(1+\beta)q - (1-q)] \left( G^H - C(x^H) + C(X) - \tilde{G}^H \right)$$

$$+ (1+\beta) \left( q \left( G^H - C(x^H) - \tilde{G}^H + C(\tilde{x}) \right) + (1+\beta) \left( C(X) - C(\tilde{x}) \right) \right),$$

where the first term is positive, because $1+\beta > 1$, and $q > 1-q$. Moreover, $\Delta W_{pp}(0) = -\kappa_2 + \kappa_6 > 0$ under Assumption 1. $\Box$

**Proof of Proposition 3**

Under Assumption 2, we have $\Delta W_o(0) = 0$, hence $\Delta W_o(\pi) \geq 0$, $\forall \pi \geq 0$. The result then follows from the facts that $\Delta W_{ss}(0) = \Delta W_o(0) = 0$, $\Delta W_{sp}(0) = \Delta W_o(0) - (1-q)\tilde{G}^H < 0$, and $\Delta W_{pp}(0) = \Delta W_o(0) + \kappa_6 > 0$. $\Box$

**Proof of Proposition 4**

We look at each range of $\beta$ separately.

(i) When $\beta$ is low, the expected welfare difference is given by (14). Recall that $\tilde{\pi}_o = \kappa_2/\kappa_1$. Then, we may write

$$\Delta W_{ss}(\tilde{\pi}_o) = \frac{\kappa_3 \kappa_2}{\kappa_1} (\kappa_1 - \kappa_2) > 0$$

In addition,

$$\Delta W_{ss}(0) < 0 < \Delta W_{ss}(\tilde{\pi}_o),$$

and

$$\Delta W_{ss}(1) = \kappa_1 - \kappa_2 > \Delta W_{ss}(\tilde{\pi}_o),$$

where the last inequality stems from the fact that $\kappa_1 = \kappa_2 + \kappa_3 + (1+\beta)q \left( G^H - C(x^H) - \tilde{G}^H + C(\tilde{x}) \right) + (1-q) \left( G^L - C(x^L) - \tilde{G}^L + C(\tilde{x}) \right)$, from which
\[
\frac{\Delta W_{ss}(1)}{\Delta W_{ss}(\hat{\pi}_o)} = \frac{\kappa_3 \kappa_2}{\kappa_1^2} < 1
\]

Hence, \(\Delta W_{ss}(\pi)\) is increasing at \(\pi = \hat{\pi}_o\), which, together with the fact that \(\Delta W_{ss}(\hat{\pi}_o)\), implies that \(\hat{\pi}_{ss} < \hat{\pi}_o\).

(ii) For high \(\beta\), straightforward algebra allows us to write

\[
\hat{\pi}_{pp} - \hat{\pi}_o = \frac{\kappa_6 (\kappa_2 - \kappa_1)}{\kappa_1 (\kappa_1 - \kappa_6)} < 0.
\]

\(\Box\)

Proof of Proposition 5

Under Assumption 2, \(\hat{\pi}_o = 0\), and \(\Delta W_{ss}(0) = -\kappa_2 - \kappa_4 < 0\). In addition, we know from (16) that \(\Delta W_{ss}(\pi)\) is increasing at \(\pi = 0\). Hence, \(\hat{\pi}_{sp} > 0 = \hat{\pi}_o\).

\(\Box\)

Proof of Proposition 6

Using (11), (12), (13), it is straightforward that \(C'(\tilde{x}^o_{pool}) > C'(\tilde{x}^o)\), and \(C'(\tilde{x}^o_{sep}) > C'(\tilde{x}^o)\), which, together the convexity of \(C(x)\), implies that \(\tilde{x}_{pool} > \tilde{x}^o\), and \(\tilde{x}_{sep} > \tilde{x}^o\).

We now use (12) and (13), together with the fact that the separating equilibrium exists for \(\beta < \gamma_p = L/H\), to obtain

\[
C'(\tilde{x}^o_{sep}) < \pi \left( \frac{q}{H} + \frac{(1-q)(1 + (1-\pi))}{H} \right) = C'(\tilde{x}^o_{pool}) - \frac{(1-\pi)^2(1-q)}{H}
\]

Again, by the convexity of \(C(x)\), \(\tilde{x}^o_{pool} > \tilde{x}^o_{sep}\).

\(\Box\)

Proof that \(\tilde{x}^o\), \(\tilde{x}^o_{pool}\) and \(\tilde{x}^o_{sep}\) are smaller than \(x^L\)

Recall that \(C'(x^L) = 1/L\). From (11), (12), and (13),

\[
C'(\tilde{x}^o) - C'(x^L) = \pi q \left( \frac{1}{H} - \frac{1}{L} \right) - \frac{1-\pi}{L} < 0
\]

\[
C'(\tilde{x}^o_{pool}) - C'(x^L) = \pi q \left( \frac{1}{H} - \frac{1}{L} \right) + (1-\pi) \left( \frac{1-q}{H} - \frac{1}{L} \right) < 0
\]

\[
C'(\tilde{x}^o_{sep}) - C'(x^L) = \pi q \left( \frac{1}{H} - \frac{1}{L} \right) - \frac{(1-\pi)(1-\pi\beta(1-q))}{L} < 0
\]

The result follows from the convexity of \(C(x)\).

\(\Box\)
References


———. 2003. Reforming intergovernmental fiscal relations in the Czech Republic.

