

KU LEUVEN

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES
DPS16.03

APRIL 2016



Accountability and subnational tax autonomy: when do politicians lose fiscal interest?

Willem SAS

Public Economics

Faculty of Economics
and Business



*Accountability and subnational tax
autonomy:*
When do politicians lose fiscal interest?*

Willem Sas[†]

Center for Economic Studies, KU Leuven

Abstract

Devolving tax authority to lower-level jurisdictions in a federation is often argued to better align the actions of politicians with the wishes of voters. In this paper, we derive the conditions for tax autonomy to bring about local growth-enhancing policies -as the *fiscal incentives* approach of Weingast (2009) would predict- and investigate whether this is indeed beneficial to voter welfare. We add to the literature by modelling a multi-tiered, political agency setting where growth-enhancing policies produce additional public revenues. Rent-seeking incumbents can then improve their chances of re-election, by setting precisely such policies and using the additional revenues for pork-barrel targeting. Surprisingly, the resulting “discipline effect” proves stronger in a unitary setting, where all of public provision is kept at the center. However, *given* a certain degree of decentralisation and a sufficient amount of rent-seeking politicians, shoring up discipline via sharper fiscal incentives is more effectively done at lower levels of government. Expanding local tax autonomy will in this case unambiguously boost voter welfare.

JEL Classification: D72, H71.

Keywords: Fiscal decentralisation, Tax autonomy, Accountability, Fiscal incentives, Elections, Productive investment, Fiscal federalism.

*I would like to thank Susana Peralta, André Decoster, and Albert Solé-Ollé for their helpful comments and suggestions. The financial support of BELSPO (Contract TA/00/39), and the ‘Steunpunt beleidsrelevant Onderzoek Fiscaliteit en Begroting (2011-2015)’ is greatly appreciated.

[†]Center for Economic Studies (CES), KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium. Tel: +32 472 26 09 89. E-mail: Willem.Sas@kuleuven.be.

I. INTRODUCTION

Decentralising functions of government is often argued to boost the *accountability* of politicians, so that policies become more responsive to local preferences or needs.¹ Because public services are brought closer to the voters, elections held in lower-level constituencies are found to filter out the better kind of politicians at a quicker pace, or to more effectively compel rent-seeking politicians to limit rent diversion.² In other words, and using the terminology of Besley and Smart (2007), decentralisation can shore up both the latter *disciplining* effect of elections, as well as the former *selection* effect.

Essentially, such accountability gains are believed to be at their highest when taxation is to a certain degree decentralised alongside public services as well (Boadway and Tremblay, 2012).³ This way, because voters can more accurately deduce how much of their taxes go to lower-level governments, they get a better picture of the true tax cost of local public spending. The available information to gauge the quality of local politicians thus improves, especially when voters can compare their own tax cost with the tax costs in other jurisdictions. Indeed, if residents in neighbouring jurisdictions are known to pay a similar amount of taxes for better public services for example, tax autonomy provides voters with a ‘yardstick’ to size up the poorer performance of their own politicians.⁴ Now, although we can reasonably expect these information gains to make rent-seeking politicians more visible, the extent to which voters stand to gain remains moot.

To see this, consider a situation where information on political track records indeed improves because of tax autonomy, and the kind of “yardstick competition” sketched out above. Rent-seeking politicians who are trying to gain re-election -and the larger rents that come with it- by mimicking their benevolent counterparts, will then have a harder time fooling the voters.⁵ *However*, if voters can only choose from a pool of mostly incompetent, lax or corrupt politicians to replace the rotten apples, the described information gains will be of little help. As with every measure improving the selection effect of voting systems, the overall quality of politicians determines whether voters will be better off in our setting as well.⁶ Making matters worse,

¹Seabright (1996) was first to define accountability in this context as the “probability that the welfare of a given jurisdiction determines the election of the government”. Persson and Tabellini (2002a) and Hindriks and Lockwood (2009) provide further theoretical formulations of this political economy perspective on fiscal federalism.

²See Faguet (2004) or Barankay and Lockwood (2007) for empirical evidence on the positive effect of decentralisation on electoral accountability.

³For evidence on the impact of local tax autonomy on accountability, see Faguet (2004) for Bolivia and Columbia, Geys et al. (2010) for Germany, and Boetti et al. (2012) for Italy.

⁴See Besley and Case (1995), Bordignon et al. (2003), Allers and Elhorst (2005), Bosch and Solé-Ollé (2007) or Revelli and Tovmo (2007) for evidence on such “yardstick competition”.

⁵As we will discuss below, the ‘mimicking’ strategy of rent-seeking incumbents exploits imperfect voter information on the true cost of public provision. If this cost is low in a certain time period, rent-seekers can simply ‘pretend’ to be benevolent politicians charging higher costs, thus siphoning off the cost-difference as rents. These rents are then smaller than the maximum rents, which would lead to certain defeat in the next elections.

⁶Besley and Smart (2007) e.g. theoretically examine the effect of, amongst others, the increased transparency of voting mechanisms and sharpened yardstick competition on voter welfare. The proportion of able, benevolent politicians to ‘bad’ politicians drives their results.

since it became harder to fool the better informed voter, the opportunity cost for bad politicians to pretend to be benevolent -in order to extract more rents after re-election- goes up. Letting their mask fall before the elections by extracting the maximum rent becomes more attractive in other words, which will erode political discipline. Here then, and rather ironically, the information gains of tax autonomy in fact hollow out the disciplining effects of elections when they are needed the most: when most politicians are rent-seeking and selection effects no longer pay off as described above. In this latter case, decentralised tax autonomy and voter welfare indeed seem to move in opposite directions.

Importantly, as will be the focus of this chapter, there is more to decentralised taxation than improved information. Focusing on a different kind of disciplining mechanism, and abstracting from yardstick competition, Weingast (2009) makes a case for *fiscal incentives* to steer local politicians towards better performance. ‘Performance’ is understood here as any kind of policy measure enhancing regional or local growth, which in turn boosts revenues out of e.g. decentralised personal income or property taxes. Self-rewarding revenue raising is key to this mechanism in other words, allowing localities which ‘perform’ well to tap into the resulting rising tax revenues. And because office-motivated politicians never have enough funds to further their goals or win over potential voters, it will then be in their own *fiscal interest* to invest in growth-enhancing policies. Even when the pool of potential politicians consists entirely of rent-seeking politicians the argument goes, voter welfare can thus be upheld by the disciplining features of this fiscal feedback loop.⁷

Studying such a feedback mechanism from a political agency perspective in what follows, our contribution will be to micro-found the kind of political incentives identified by Weingast (2009). Following Besley and Smart (2007), our model of imperfect information spans two time-periods, allows for good as well as rent-seeking politicians, and is applicable to multi-tiered forms of government as in Hindriks and Lockwood (2009). Adding to the literature moreover, policy outcomes feed back into public revenue flows, so that growth-enhancing policies -such as productive investment or business-friendly regulatory efforts- bring about additional government income. Since we also allow for two types of voters, and because voting is probabilistic, rent-seeking incumbents may then use these additional revenues for pork-barrel targeting, winning over what we call the “priority” vote.⁸ Priority voters have specific interests over *non-valence* issues, where we think of the *valence* issue as an item on the political agenda that all voters agree on, and which consequently captures the broader notion of general welfare. In our case, this will be the kind of growth-enhancing policies Weingast (2009) is referring to.⁹ Not surprisingly then, priority voters will lean more towards the politicians targeting their specific interests,¹⁰ rather than voting for politicians

⁷See Weingast (1995, 2009, 2014) for further argumentation. Jin et al. (1999, 2005) and Kappeler et al. (2013) present some evidence.

⁸Contrary to e.g. Broilo et al. (2013), these additional revenues will never be sufficiently large to attract ever more rent-seeking candidates to the pool of politicians standing for office, or to inflate rents beyond their initial maximum proportions.

⁹As also argued by Cadot et al. (2006), infrastructure investment is not a policy issue drawing on the partisanship of voters, and thus rarely pits different parts of the electorate against each other.

¹⁰What we envisage here is a similar kind of pork-barrel targeting as in Roberson (2008), specifi-

investing in the valence issue of economic growth. For the same reason, they will also care less about the personal traits, or ideologies of their preferred politicians.

What we find is that everything hinges once more on the quality of the pool of politicians voters can choose from, but also, and crucially, on the composition of the voting population itself. If most voters do not have specific interests and only care about the valence issue of economic growth, rent-seeking politicians will be hard put to improve their chances of re-election through pork-barrel targeting. Indeed, the more priority voters there are to be won over, the more the fiscal incentives will bite, leading to less rent diversion. However, we show that this mechanism is stronger still in a unitary setting, where all of public provision is kept at the center. The reason is the “reduced pivot-probability effect”, coined by Lockwood (2006) and first introduced by Seabright (1996), making it more likely for a central government to opt for the mimicking strategy. The underlying idea is that, depending on the voting rule in place, the incumbent central government only needs to sway a pivotal subset of constituencies in its favour, whereas regional or local governments rely on a *single* constituency to ensure re-election. The mimicking strategy of bad politicians -i.e. postponing maximum rent diversion to later time periods- will thus be less costly for a central government than for a subcentral government. Indeed, and in terms of our setting here, because growth-enhancing policies only need to be introduced in a pivotal amount of constituencies for a central government to be returned to office, maximum rents can be extracted right away in the non-pivotal ones.

The built-in fiscal incentives mechanism à la Weingast and the cost reductions working through reduced ‘pivot-probability’ are then shown to be mutually reinforcing in our model. As a result, the discipline effect of voting systems will always be stronger at the center. When offered the choice between more or less decentralisation of public functions subsequently, voters only prefer more decentralisation if politicians are less likely to be rent-seeking to begin with, and selection is needed more than discipline. Nevertheless, *given* a certain degree of decentralisation and a sufficient amount of rent-seeking politicians, shoring up discipline via sharper fiscal incentives is more effectively done at the lower level of government. Expanding local tax autonomy will in this case unambiguously boost voter welfare.

Of course, and as discussed above, to study voter welfare from the fiscal incentives perspective there can only be one kind of “valence” issue: economic growth. With its focus on regional economic performance -and contrary to the accountability reasoning where voters were perfectly free to fill in the notion of “good” policies themselves- the fiscal interest approach inevitably has a rather narrow take. It simply reduces welfare-enhancing policies to economically sound policies, and should therefore always be carefully weighed against other arguments in favour of (de)centralisation.

The remainder of this paper is organised as follows. In section II.1 we set out our fiscal incentives approach in further detail, and introduce the economic and political environment of the model. In section II.2 we discuss the information available to our main decision-makers, namely voters and politicians, and define the timing of their decisions. Section III then considers a fully decentralised equilibrium within a federal

cally focusing on his notion of targetable local project provision.

constellation, whilst section IV looks at a unitary setting and compares both regimes. The main effects on voter welfare under both fiscal regimes are derived in section V. Section VI offers some concluding remarks.

II. THE MODEL

Our federal economy consists of an odd number $n \geq 3$ of lower-level jurisdictions, referred to as states. We consider two time periods, and an election which is held in between both periods.

II.1. The economic and political environment

Elected state politicians provide an amount of local, growth-enhancing public goods G_i in each state i and each period, after observing unit costs θ_i of public provision. These costs can only take on two discrete levels, a high value H and a lower value L , where the probability of costs coming in at their highest level is $\Pr(\theta_i = H) = q_i$. We assume unit costs θ_i are independently and identically distributed in each period as in Besley and Smart (2007), but are fully correlated across regions.¹¹ State governments only partially cover own public expenditures with own taxes t_i , and are co-financed by use of grants t_f transferred by the federal government. As will become clear below, federal grants are fully responsive to shocks in unit costs θ_i , which coincides with many needs-based grant systems in the field. Total tax collections $T_i = t_f + t_i$ then finance public spending $\theta_i G_i$ in state i , as well as the possible rents r_i diverted by the incumbent state government, so that $T_i = \theta_i G_i + r_i$. The amount of rents r_i simply denotes the level of public spending devoted to private ends or other forms for corruption, e.g. rewarding cronies.

The spectrum of voters $[0, \bar{n}_i]$ in each state is normalised to 1, with each voter deriving the same utility from public goods adjusted for the cost of government spending. When a state government provides an amount of G_i public goods at a tax burden of T_i consequently, its voters incur a welfare level of $W_i = G_i - \mu_i C_i(T_i)$. Capturing the individual costs of taxation, $C_i(T_i)$ is a strictly convex, increasing function where the exogenous parameter μ_i denotes the marginal cost of public funds. Following Besley and Smart (2007), a rise in μ_i captures either an intensification of tax competition, the electoral passage of a (constitutional) restriction on the tax base or tax instrument, or technological and administrative complications in tax collection.

Importantly, the spectrum of voters is made up out of two types: *valence* voters occupy a share $(1 - \omega_i)$ of the field, whilst a portion ω_i of *priority* voters accounts for the rest. What distinguishes both types, is the fact that valence voters *only* derive utility from the public good G_i , whilst priority voters have other concerns as well. Put otherwise, even though all voters derive utility from the valence good G_i , priority voters may derive even more utility from other policies which they consider highly indispensable. Such policies can then range from e.g. the environment, poverty reduc-

¹¹Since G_i embodies growth-enhancing policies such as productive investment or improved regulation, unit costs can realistically be assumed more or less the same across a federation. This assumption also allows for tractable results in what follows.

tion, accessible health care or effective education, but have in common the amount of utility ϕ_i they add to priority voters' welfare. In what follows, and similar to models of political polarisation used in Besley and Burgess (2001) or Besley et al. (2010), we assume this latter gain ϕ_i to be of such a degree that only valence voters will let the provision of G_i play a part in their assessment of incumbents. In other words, priority voters will only care about their specific interests, and whether or not the incumbent has in fact heeded these.

State governments are groups of like-minded, identical politicians of type $x_i \in \{b, g\}$ which -as in Besley and Smart (2007)- can be either of the 'good' type g , or the 'bad' type b . The good kind of politician is a token benevolent leader, choosing G_i in each period to maximize voter welfare, and hence drawing no satisfaction from rents diverted from public spending r_i . Consequently, and given θ_i , the level of local public provision set by a state government consisting of benevolent politicians comes in at

$$G_i^{\theta_i}(\theta_i, \mu_i) = \arg \max G_i - \mu_i C_i(\theta_i G_i), \quad (1)$$

with $T_i^{\theta_i} = \theta_i G_i^{\theta_i}(\theta_i, \mu_i)$ the associated level of tax collections financing public spending. As mentioned above, the federal grants adjust fully to shocks in the unit cost of public provision θ_i , so that $t_f^{\theta_i} = (1 - \nu_i)T_i^{\theta_i}$ and $t_i^{\theta_i} = \nu_i T_i^{\theta_i}$, with $\nu_i \in [0, 1]$ the degree of tax autonomy in state i .¹² Under a state benevolent government lastly, voter welfare is denoted as $W_i^g(\theta_i, \mu_i)$. Unsurprisingly, both $G_i^{\theta_i}$ and W_i^g are decreasing in μ_i , since a higher marginal cost of taxation has benevolent politicians set lower taxes, resulting in lower levels of public provision.

Unlike benevolent politicians, bad politicians behave strategically by maximising rents r_i^1 in period 1 as well as discounted rents $\beta \sigma_i r_i^2$ in period 2, with β the discount rate and σ_i the probability of an incumbent government being re-elected in state i . The re-election rule, as well as the decision process of bad politicians, will be set out in section III.¹³ We also assume there to be a maximum level X_i of state tax collections -and thus also of rent diversion- that can be imposed on voters, where $T_i \in [0, X_i]$ and $X_i > T_i^L$.

Lastly, and introducing the fiscal incentives discussed earlier, we assume the incumbent government can cater to the specific needs of a significant amount of priority voters if it receives additional revenues Y_i to finance such "priority policies". Similar to Jin et al. (2005), and since the valence good G_i captures growth-enhancing measures, public provision itself is assumed to generate these additional revenue flows to the tune of $Y_i = \nu_i R_i(G_i)$.¹⁴ Here, R_i is a strictly convex, increasing function with

¹²We thus assume the federal government can observe unit costs θ_i of public provision, but does not know the type of the incumbent politicians. As a first mover, it therefore sets its federal grants based on the unit costs only. Also, the degree of tax autonomy ν_i allows for the entire spectrum between full tax autonomy and fully dependent lower-levels of government.

¹³Note also that we have, in effect, set $\beta = 0$ for benevolent politicians. As discussed in Lockwood (2005), assuming that benevolent politicians are fully myopic delivers a unique and stable equilibrium in the signalling game we will set up in the following sections.

¹⁴Of course, as in e.g. Keen and Marchand (1997) or Hindriks et al. (2008), productive public inputs could also be modelled as contributing directly to state revenues through a tax on capital earnings. Since we do not focus on inefficient over- or underproduction of public provision however, this would leave our results unchanged. Moreover, our more generic approach allows for a broader

$R_i(0) = 0$, whilst the degree of tax autonomy ν_i denotes the marginal retention rate, i.e., the share of increased revenues a state government gets to keep. We can then assign probabilities $\rho(Y_i)$ to the event of capturing a share $\eta \in [0, 1]$ of the priority vote when using the additional revenues Y_i for pork-barrel targeting, as follows:

$$\eta[Y_i(G_i)] = \begin{cases} \frac{1}{2} + \chi & \text{with } \rho[Y_i(G_i)] \\ \frac{1}{2} - \chi & \text{with } 1 - \rho[Y_i(G_i)], \end{cases} \quad (2)$$

where $0 < \chi \leq \frac{1}{2}$, $\rho(Y_i)$ is increasing in Y_i , and $\rho(0) = \frac{1}{2}$. What we find in expression (2) in other words, is nothing more than the fiscal interest mechanism in full swing. The more the incumbent state government invests in market enhancing policies or infrastructure G_i , the higher its additional revenue flows Y_i , and as a result, the easier to win over a majority share η of the priority vote ω_i .

II.2. Information and timing

At the end of period 1, an election is held in each state where one group of politicians challenges the group in office. The group winning the majority of votes wins the election. Whether the incumbent politicians at the beginning of period 1 -as well as the challengers- are of the good type g or the bad type b is defined by independent draws from an identical distribution. With a probability $\Pr(x_i = g) = \pi_i$, a group of politicians -incumbent or challenger- in a given state i will be benevolent. The ensuing game between incumbent state politicians and voters is then defined as follows.

At the beginning of period 1, the type $x_i \in \{b, g\}$ of the group of incumbent politicians is drawn for each state i . These incumbents then observe the unit costs of public provision θ_i and their federal grant $t_f^{\theta_i}$, after which they decide on state taxation t_i , rents r_i^1 , and public goods G_i . Ahead of the elections the voters observe the amount of public goods G_i provided in their state, as well as the collected taxes t_i and t_f to finance public spending. The unit costs θ_i of public provision however, together with the type of both the incumbent and challenging state politicians, remain unobserved. What voters do observe is the probability q_i that unit costs θ_i are high, the probability π_i that politicians are benevolent, and the degree of state tax autonomy ν_i . After the elections, the elected group of politicians again sets G_i and r_i^2 . Since there are no elections after period 2, even newly-elected challengers can be considered “lame ducks” whose actions will not be influenced by electoral pressure.

Clearly, since the actual type of politicians as well as the rents essentially remain hidden to the voter’s eye, the game described above has a distinct structure of imperfect information. To figure out whether the incumbent is benevolent or not, the only option open to valence voters is to scrutinize incumbent performance during period 1, and weigh their -as such- updated beliefs about the incumbents’ type against their prior beliefs about the challengers. We elaborate on the resulting Bayesian equilibria in the following sections.

Notice lastly how -contrary to e.g. the career-concerns models developed by Persson and Tabellini (2002b)- politicians can be good or bad, and are *equally* competent perspective on all possible sources of government income.

to produce the desired amount of public goods at either unit cost $\theta_i \in \{H, L\}$. Moreover, politicians are fully aware of this competence ex ante -in stead of ex post as in Persson and Tabellini (2002b)- and will thus be able to hide their true type from the voters if needed. Why they would want to do so, will become clear below.

III. DECENTRALISED EQUILIBRIUM

We solve the game of incomplete information described above by applying a type of backward induction, obtaining a unique Bayes-Nash equilibrium in each state. We therefore start with period 2, and turn first to the interaction between politicians and valence voters.

As there are no elections following period 2, the group of politicians in office in that period will no longer be constrained by electoral discipline. Good behaviour will never lead to re-election and future rents, which has bad politicians divert the maximum amount of rents $r_i^2 = X_i$. Setting state taxes t_i so that $T_i = X_i$ is diverted away from public provision, G_i will be equal to zero as a result. Inversely, good politicians never divert rents, set $r_i^2 = 0$, and consequently decide on G_i following (1).

Since second-period strategies are the same for bad incumbents or challengers alike, i.e. extracting full rents, the best strategy for valence voters is to weed out as many bad politicians they can during the elections. Their sequential voting rule will as a result be to re-elect the incumbent group of period 1 if they think this group is more likely to be benevolent than the challengers. In other words, if the *posterior* probability they ascribe to the incumbents being benevolent surpasses the *prior* probability π_i of the challengers, they re-elect the incumbents. The voter's posterior beliefs will thus inevitably be based on incumbent performance during period 1 only, and follow from the equilibrium strategies of first-period incumbents.

Zooming in on period 1, a benevolent state government then simply maximises voter welfare following (1), and chooses (G_i^H, T_i^H) with probability q_i , or (G_i^L, T_i^L) with probability $(1 - q_i)$. Logically then, since the voter has this information, his posterior beliefs will assign probability zero to the incumbent being of the good type at any *other* information set (G_i, T_i) . Consequently, in any perfect Bayesian equilibrium, $\Pr(g|T_i) = 0$ if $(G_i, T_i) \neq (G_i^{\theta_i}, T_i^{\theta_i})$. At any such information set, the valence voter elects the challenger.

As a result, a group of bad incumbents has only three possible strategies left in terms of first-period tax collection t_i , so that, residually, $T_i \in (T_i^L, T_i^H, X_i)$. In the latter case the bad incumbents claim the full rent $r_i^1 = X_i$ as in period 2, revealing their true type b and as such “separating” from the good politicians. In the first two cases on the other hand, incumbents undertake at least some measure of public provision to hide their true type, thus mixing in or “pooling” with the good politicians. The reason for this masquerade is the re-election motive, in full effect when the sum total of expected rents over both periods outweighs rents X_i in period 1, so that

$$r_i^1 + \beta\sigma_i X_i > X_i, \tag{3}$$

where we have replaced rents r_i^2 in period 2 with the maximum value of X_i , which was the incumbent strategy of bad politicians in period 2. Now, filling in the blanks in expression (3) are the unit costs of public provision, $\theta_i \in (L, H)$. Suppose the bad incumbents face low unit costs L in period 1. By setting tax collections t_i so that $T_i^{\theta_i} = T_i^H$, and providing the corresponding amount of public goods G_i^H , they are able to siphon off rents to the extent of $\hat{r}_i^1 = (H - L)G_i^H$. Inversely, when $\theta_i = H$, the incumbents will not be able to claim any rents without revealing their type. In such a situation, where the pooling strategy does not pay any rents in period 1 so that $\hat{r}_i^1 = 0$, the separating strategy always dominates. Indeed, $r_i^1 = X_i$ exceeds expected second-period rents $\beta\sigma_i X_i$ to be gained after re-election. For exactly the same reasons, valence voters always re-elect the incumbent group after observing (G_i^L, T_i^L) in period 1, so that in any equilibrium we get that

$$\Pr(g|T_i^L) = 1. \quad (4)$$

Arriving at the proper posterior beliefs based on the observation (G_i^H, T_i^H) subsequently, is more intricate. Sure enough, valence voters know of the risk that a group of bad politicians might pretend to be benevolent in order to improve its re-election chances, yet it remains an uncertainty. They therefore assign probability λ_i to this pooling strategy where

$$\lambda_i = \Pr(T_i = T_i^H | \theta_i = L, x_i = b). \quad (5)$$

Based on all available information, and using Bayes rule, valence voters then infer the posterior probability that first-period tax collections T_i^H were levied by benevolent incumbent politicians as

$$\Pr(g|T_i^H) \equiv \Pi_i = \frac{\pi_i q_i}{\pi_i q_i + (1 - \pi_i)(1 - q_i)\lambda_i}, \quad (6)$$

which allows us to derive lemma 1 below.¹⁵

Lemma 1 *Given the posterior probability $\Pr(g|T_i^H) = \Pi_i$ defined in (6), and assuming that $q_i > \frac{1}{2}$, the valence voter will always re-elect the incumbent when observing first period public provision of G_i^H .*

Suppose now an incumbent government of bad politicians would only have to worry about winning over valence voters. Its first-period strategies would then be straightforward at this point. If first-period unit costs θ_i are low, and given lemma 1, incumbent politicians will face a re-election probability of $\sigma_i = 1$ if they provide G_i^H at a total tax take of T_i^H . From (3), we then deduce that the pooling strategy $\hat{r}_i^1 = (H - L)G_i^H$ will always be more beneficial than full rent extraction if and only if

$$\hat{r}_i^1 + \beta\sigma_i X_i > X_i. \quad (7)$$

¹⁵A simple proof is provided in appendix V.A. Following Hindriks and Lockwood (2009), we assume that $q_i > 1/2$ in all states. This rules out the hybrid equilibrium derived by Besley and Smart (2007), which was proven unstable in the Cho-Kreps sense by Lockwood (2005).

If condition (7) does not hold however, or in the case that unit costs come out on the high side H , bad incumbents will always separate and reveal their type. Their probability of re-election σ_i is reduced to zero because of this.

Of course, and crucially, the voting population does not simply consist of valence voters. Priority voters also influence the probability of re-election σ_i in (7) which, in turn, alters first-period incumbent strategies as well. This is where the fiscal interest mechanism kicks in, and where outcomes become rather less clear-cut as a result. Indeed, by providing a certain level of market-enhancing public goods G_i incumbent politicians will also generate additional revenues Y_i , which can be used to finance priority policies. Pulling in a larger share of the total vote, bad incumbents can thus win the day in two ways: by pretending to be benevolent as before, and by winning the hearts of priority voters. Since a group of bad incumbents will never set (G_i^L, T_i^L) since this would violate (7), we focus on the probability σ_i of re-election when the incumbent sets (G_i^H, T_i^H) .

Adding to the realism and applicability of the model, we furthermore introduce a probabilistic framework at this point, of the kind extensively used in Persson and Tabellini (2002b). In this light, we assume that valence voters also care about a second policy dimension, orthogonal to public provision G_i . We refer to this second dimension as the “ideology” of a group of politicians, yet it could apply to any personal characteristics of the politicians themselves. When casting their vote, valence voters thus base their voting decisions not only on incumbent performance as captured by lemma 1, but also on the ideologies of both competing groups of politicians. Specifically, a given valence voter j will now re-elect the incumbent group of politicians if

$$\Pi_i > \pi_i + \gamma_i^j + \delta_i. \quad (8)$$

As in Persson and Tabellini (2002b), the ideological policy dimension comes in through both terms on the right side of (8). Here, γ_i^j is an individual-specific parameter which captures voter j 's *individual* ideological bias towards the groups of politicians, which can take on negative as well as positive values. Voters for whom $\gamma_i^j = 0$ are ideologically neutral, whilst voters where $\gamma_i^j < 0$ are ideologically biased in favor of the incumbent group, and vice versa. We assume γ_i^j is uniformly distributed on the interval $[-\frac{1}{2}, \frac{1}{2}]$. Second, the parameter δ_i reflects the *aggregate* popularity of both political groupings across the population as a whole, which can also be positive or negative and is again uniformly distributed, but now on the interval $[-\frac{1}{2\xi}, \frac{1}{2\xi}]$.¹⁶ Building on the totality of our framework, we can then derive the overall probability of re-election σ_i in lemma 2.¹⁷

Lemma 2 *When the incumbent group of politicians provides a level of public goods G_i^H , and with $\alpha_i = \frac{\omega_i(2\rho[Y_i(G_i)]-1)\chi}{(1-\omega_i)} \geq 0$, its re-election probability σ_i of winning over a majority of both priority as well as valence voters μ_i , is given by*

¹⁶Both distributional assumptions facilitate closed form solutions. For a discussion of their generalisation, we refer to Persson and Tabellini (2002b).

¹⁷See appendix V.B for the derivations.

$$\Pr \left[\mu_i \geq \frac{1}{2} \right] = \sigma_i(\alpha_i + (\Pi_i - \pi_i)) = \begin{cases} 1 & \text{if } \alpha_i + (\Pi_i - \pi_i) > \frac{1}{2} \\ \frac{1}{2} + \xi[\alpha_i(Y_i, \chi, \omega_i) + (\Pi_i - \pi_i)] & \text{Otherwise} \\ 0 & \text{if } \alpha_i + (\Pi_i - \pi_i) < -\frac{1}{2}. \end{cases} \quad (9)$$

For a good understanding, suppose the probability of winning expressed by (9) lies between zero and one. Plugging (9) into (7), we then arrive at the necessary condition for a group of bad incumbents to provide G_i^H -in other words, opt for the pooling strategy- which is

$$\hat{r}_i^1 + \beta \sigma_i(\alpha_i(Y_i, \chi, \omega_i) + (\Pi_i - \pi_i)) X_i > X_i, \quad (10)$$

where, using (5) and (6), Π_i is defined by setting $\lambda_i = 1$ in the latter expression. In other words, if the sum total of expected rents characterised by the left hand side of (10) exceeds the rents to be captured in period 1, the incumbents will always mimic the benevolent politicians to be re-elected. In any other case they separate, and are voted out. We summarise in lemma 3.

Lemma 3 *As long as $\hat{r}_i^1 + \beta \sigma_i(\alpha_i + (\Pi_i - \pi_i)) X_i > X_i$, a group of bad incumbents will always choose the pooling strategy, i.e. set (G_i^H, T_i^H) when $\theta_i = L$, and will be re-elected as a result. They separate otherwise, and extract the full rent $r_i^2 = X_i$.*

Now, this equilibrium clearly hinges on condition (10) and the probability of re-election σ_i which, compared to a setting without priority voters, in turn depends entirely on α_i as defined by lemma 2. We investigate in proposition 1.

Proposition 1 *When growth-enhancing policies lead to additional state revenues, rent-seeking incumbent politicians can improve their chances of re-election by introducing exactly such policies, using the additional revenues to win over priority voters. The extent to which they will do so, depends on*

1. *The marginal retention rate ν_i : the more revenues flow back into state coffers, the more priority votes can be won over;*
2. *The share of priority voters ω_i : as the share of priority voters grows, ideology and popularity shocks grow less important;*
3. *The marginal cost of public funds μ_i : a lower marginal cost of taxation implies higher levels of G_i^H , and higher revenues.*

Proof From section II we know that $Y_i = \nu_i R_i(G_i)$, where $R_i(0) = 0$ and $Y_i'(G_i) > 0$. Since $\rho(Y_i = 0) = \frac{1}{2}$, and zooming in on the expression for α_i given in lemma 2,

$$\alpha_i = \frac{\omega_i (2\rho[Y_i(G_i)] - 1) \chi}{(1 - \omega_i)}, \quad (11)$$

we know that (11) will be equal to zero when $G_i = 0$. Moreover, since $\rho(Y_i)$ is increasing in Y_i , we have that $\frac{d\alpha_i(G_i, \nu_i, \omega_i, \chi)}{dG_i} > 0$ and $\alpha_i(G_i, \nu_i, \omega_i, \chi) > 0$ for all other possible values of $G_i, \nu_i, \omega_i, \chi$, given that states enjoy some degree of tax autonomy

$\nu_i \in]0, X_i]$. From (11) we also learn that $\frac{d\alpha_i(G_i, \nu_i, \omega_i, \chi)}{d\nu_i} > 0$ and $\frac{d\alpha_i(G_i, \nu_i, \omega_i, \chi)}{d\omega_i} > 0$, which, together with the fact that $\alpha_i(G_i, \nu_i, \omega_i, \chi) > 0$ when $G_i = G_i^H$ as shown above, proves points 1) and 2) of proposition 1 as higher values of α_i increase the probability that condition (10) holds. Likewise, since lower marginal costs of public funds μ_i translate into higher public provision G_i^H following (1) and $\hat{r}_i^1 = (H - L)G_i^H$, we know that $\frac{d\alpha_i(G_i, \nu_i, \omega_i, \chi)}{d\mu_i} > 0$ and $\frac{d\hat{r}_i^1(G_i)}{d\mu_i} > 0$, which proves point 3) of proposition 1. \square

What we learn from lemma 3 and proposition 1, is that the mere presence of priority voters provides bad incumbents with a second incentive to work for re-election, aside from pure reputation building. Indeed, without priority voters (9) would reduce to the usual trade-off between reputational gains $(\Pi_i - \pi_i)$ -achieved by the incumbents after setting G_i^H - and popularity shocks δ_i . With priority voters on the other hand, the fiscal incentive “feedback loop” results in more politicians choosing for the pooling strategy, rather than simply separating. The stronger the feedback loop, and the more priority voters, the higher this kind of electoral discipline. Lastly, lower marginal costs of public funds μ_i not only fatten potential rents r_i^1 as in Besley and Smart (2007), but also bring about larger additional revenues Y_i which strengthens the fiscal incentive mechanism. Following lemma 3, both effects improve political discipline.

IV. CENTRALISATION AND COMPARISON OF FISCAL REGIMES

Suppose now that instead of having n different state governments deciding on public provision in their own state, a central government decides on the full set of regional policies so that -since all public functions are now centralised- tax autonomy ν_i also equals zero in each state. Second-period strategies remain unchanged for both types of politicians in this scenario: a benevolent central government would still optimise (1) for each state i , and a group of bad politicians would still extract maximum rents $r_i^2 = X_i$. As a result, valence voters again only look at first-period policies when casting their vote, and since a benevolent central government also optimises (1) in period 1, he or she again ascribes probability $\Pr(g|T_i) = 0$ to any situation where $(G_i, T_i) \neq (G_i^{\theta_i}, T_i^{\theta_i})$. At any such information set consequently, and in any perfect Bayesian equilibrium, the valence voter elects the challenger.

The optimal strategy for a group of bad incumbents however, is different from what we had before under decentralisation. This for the simple reason that a central government usually does not require a majority of the votes in all of its constituencies to be re-elected. As discussed in Seabright (1996) or Hindriks and Lockwood (2009), the probability that a certain constituency is pivotal in the electoral outcome is diminished once we move from a set of decentralised political entities to a unitary constellation made up out of many constituencies. Following Hindriks and Lockwood (2009) we apply a simple electoral rule here, where the central government only has to gain a majority in $m = (n + 1)/2$ states to be re-elected. Consequently, the pooling strategy of mimicking the benevolent politicians when $\theta_i = H$ becomes more attractive, as (G_i^H, T_i^H) only needs to be set in m pivotal states, whilst the full rent can

now be extracted in $(n - m)$ states in period 1 as well. We derive the equilibrium consequences of this kind of “selective pooling”¹⁸ in lemma 4.¹⁹

Lemma 4 *A central government of bad incumbents chooses the pooling strategy in $m = (n + 1)/2$ states, i.e. sets (G_i^H, T_i^H) when $\theta_i = H$, if and only if*

$$\frac{m}{n\sigma_i(\alpha_{ic} + (\Pi_i - \pi_i))} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) < \beta, \quad (12)$$

and is re-elected as a result. It separates otherwise, and extracts full rents $nr_i^2 = nX_i$.

As we show in appendix, the trade-off captured by (12) is nothing more than condition (10) solved for β , adjusted for selective pooling and the fact that now $\nu_i = 0$, which alters α_i .²⁰ Indeed, setting $m = n$ and $\alpha_{ic} = \alpha_i$ has lemma 4 reduce to lemma 3. In both cases the cost of fiscal restraint $\left(1 - \frac{\hat{r}_i^1}{X_i}\right)$ -i.e. of not extracting the full rent X_i in period 1- is weighed against the value of re-election, captured by the discount factor β . The higher β relative to $\left(1 - \frac{\hat{r}_i^1}{X_i}\right)$, the more attractive the pooling strategy for incumbent politicians, since fiscal restraint in period 1 becomes less costly, and future rents more valuable. Furthermore, a higher probability σ_i of re-election pushes the left hand side of (12) downwards, hence lowering the ‘trigger value’ $\beta_i^C = \frac{m}{n\sigma_i(\cdot)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right)$ at which point bad incumbents start choosing for the pooling strategy, as also implicitly expressed by proposition 1. What distinguishes lemma 4 from the previous decentralisation case however, is the selective pooling effect expressed by $\frac{m}{n}$ in (12). We elaborate in proposition 2.

Proposition 2 *The disciplining effect of elections will be weaker under decentralisation, since the pooling strategy is more attractive when public provision is centralised. However, given decentralisation, shoring up discipline via sharper fiscal incentives is more effectively done at the state level.*

Proof The first part of proposition 2 is trivial, as it is clear that

$$\frac{m}{n\sigma_i(\alpha_{ic} + (\Pi_i - \pi_i))} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) = \beta_i^C < \beta_i^D = \frac{1}{\sigma_i(\alpha_i + (\Pi_i - \pi_i))} \left(1 - \frac{\hat{r}_i^1}{X_i}\right), \quad (13)$$

since $\frac{m}{n} < 1$, $\hat{r}_i^1 \leq X_i$, σ_i is a probability and $\alpha_i < \alpha_{ic}$. The latter inequality holds because $\nu_i = 0$, which yields a higher value of α_i as expressed by (11) since $Y_i = (1 - \nu_i)R_i(G_i)$ with centralisation, and $\rho(Y_i)$ is increasing in Y_i . In any case, what we learn from (13) is that the triggering value for bad incumbents to pool is as a result lower in the centralised case, since $\beta_i^C < \beta_i^D$. Suppose now that $\sigma_i(\alpha_i + (\Pi_i - \pi_i))$ increases under any of the possible scenarios given in proposition 1. The amount by which β_i^C drops as a result, will then always be only $\frac{m}{n}$ of the amount by which β_i^D would drop. Increasing the probability of re-election thus boosts discipline more in the decentralised case. \square

¹⁸Hindriks and Lockwood (2009) coined the term, and were first to introduce Seabright’s “reduced pivot probability” mechanism to Besley and Smart’s (2007) political agency framework, which allows for selection as well as discipline effects.

¹⁹See appendix V.C for the derivations.

²⁰The proof of proposition 2 below provides further insight in the adjusted α_{ic} under centralisation.

As in Hindriks and Lockwood (2009), the option of selective pooling under centralisation leads to more discipline and less political turnover, and this compared to a decentralised setting. The fact that the probability of re-election σ_i is endogenous to pork-barrel targeting through additional spending Y_i , only reinforces this effect as we can see in (13). However, once a decentralised constellation is *given*, endogenous election probabilities can considerably improve local discipline on the margin. Indeed, in any of the scenarios described in proposition 1 - an increase in the retention rate ν_i , a larger share of priority voters ω_i , or smaller marginal costs of taxation μ_i - the resulting rise in σ_i will be more pronounced compared to a similar upshot under centralisation. As a result, the triggering value for bad incumbents to choose the pooling strategy will drop more sharply at the state level. Given the decentralisation of certain public functions consequently, extending the degree of tax autonomy ν_i will more effectively improve discipline compared to similar endeavours in a centralised setting. When deciding on which tax bases to decentralise moreover, choosing a tax base which is liable to induce less tax competition -and thus incurs a smaller marginal cost μ_i - has the same effect. Of course, the question remains what proposition 1 and proposition 2 mean in terms of voter welfare. We tackle this question in the following section.

V. WELFARE IN BOTH REGIMES

To size up the welfare effects of decentralising both public functions and tax autonomy as specified above, we first need to describe voter welfare levels in several relevant, or counterfactual, situations. In this light, we first of all write expected per-period voter welfare when a benevolent group of type g is in office as

$$EW_i^g(\mu_i) = q_i W_i^g(H, \mu_i) + (1 - q_i) W_i^g(L, \mu_i), \quad (14)$$

with q_i the probability that the unit costs of public provision are high, or $\theta_i = H$, as defined above. Next, when the incumbent group of politicians is of type b and furthermore extracts the full rent (X_i), welfare can be spelled out as

$$W_i^b(\mu_i) = -\mu_i C_i(X_i). \quad (15)$$

Using (14) and (15), expected voter welfare when an unknown type of politician is in office -and when, if bad, incumbents always separate- then becomes

$$W_i^0 = \pi_i EW_i^g(\mu_i) + (1 - \pi_i) W_i^b(\mu_i), \quad (16)$$

with π_i the probability that a politician is good, as specified earlier. Now, in order to compare different political constellations in terms of welfare, we use a baseline welfare level similar to the benchmark used in Hindriks and Lockwood (2009), given by

$$\bar{W}_i = W_i^0 + \beta (\pi_i EW_i^g(\mu_i) + (1 - \pi_i) W_i^b(\mu_i)). \quad (17)$$

Our two-period benchmark \bar{W}_i thus captures expected voter welfare in a baseline scenario where (10) holds -and all politicians are re-elected as a result- but where

bad politicians nevertheless separate. They are replaced by good politicians setting welfare according to (14) with probability π_i , or by bad politicians providing only (15) with probability $(1 - \pi_i)$, as expressed by $(1 - \pi_i)W_i^0$ in (17). Given this definition of \bar{W}_i , and allowing for pooling, we can then write welfare in the decentralised setting discussed above, as

$$EW_i^D(\mu_i) = \bar{W}_i + \lambda_i(1 - \pi_i)(1 - q) (\Delta_i^d - \beta\pi_i\Delta_i^s). \quad (18)$$

What we see in (18) is that, if -with probability $\lambda_i(1 - \pi_i)(1 - q)$ - a group of bad incumbents pools rather than separates, expected voter welfare $EW_i^D(\mu_i)$ will diverge from the baseline in two important ways. First, voters face a “selection loss” in period 2. They miss out on the welfare they otherwise would have gained if the same group of incumbents were to have separated, to be replaced by benevolent politicians with probability π_i . The present value of this second-period welfare loss returns as $\beta\pi_i\Delta_i^s$ in (18), with $\Delta_i^s = EW_i^g(\mu_i) - W_i^b(\mu_i)$ the gap between counterfactual welfare $EW_i^g(\mu_i)$ and actual welfare $W_i^b(\mu_i)$ in period 2. Second, voters also enjoy a “discipline benefit” in period 1. Because the group of bad incumbents exerts fiscal restraint rather than diverting the maximum rent, voters attain a welfare level of $W_i^g(H, \mu_i)$ in stead of $W_i^b(\mu_i)$, as expressed by $\Delta_i^d = W_i^g(H, \mu_i) - W_i^b(\mu_i)$ in (18). Now, using the same reasoning, we can write voter welfare in a centralised constellation as

$$EW_i^C(\mu_i) = \bar{W}_i + \lambda_i(1 - \pi_i)(1 - q) \left(\frac{m}{n} \Delta_i^d - \beta\pi_i\Delta_i^s \right). \quad (19)$$

Comparing (19) and (18), a first difference lies in the benefits from pooling. Whereas in a decentralised setting the benefits would be reaped in all n states, selective pooling of the central government limits the benefits to m states. Second, and following proposition 2, the pooling probabilities λ_i will also differ in both political regimes. We formalise this comparison in proposition 3.

Proposition 3 *Depending on the triggering value for incumbents to pool, introduced in lemma 4, we can distinguish between the following welfare scenarios:*

1. If $\beta \leq \frac{m}{n\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right)$, $EW_i^D = EW_i^C$,
2. If $\frac{m}{n\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) < \beta \leq \frac{1}{\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right)$, $EW_i^D > EW_i^C \Leftrightarrow \pi_i > \frac{m}{n} \frac{\Delta_i^d}{\beta\Delta_i^s}$,
3. If $\frac{1}{\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) < \beta$, $EW_i^D > EW_i^C$.

As a result, a decentralised system can only potentially Pareto-dominate a centralised framework if a sufficiently large fraction of politicians is benevolent, so that $\pi_i > \frac{m}{n} \frac{\Delta_i^d}{\beta\Delta_i^s}$.

Proof Subtracting (19) from (18), we can write the potential welfare gains of decentralisation as

$$\frac{EW_i^D(\mu_i) - EW_i^C(\mu_i)}{(1 - \pi_i)(1 - q)} = (\lambda_i^D - \lambda_i^C) (\Delta_i^d - \beta\pi_i\Delta_i^s) + \lambda_i^C \left(1 - \frac{m}{n}\right) \Delta_i^d, \quad (20)$$

with λ_i^D and λ_i^C the pooling probabilities under decentralisation and centralisation respectively. Following lemma 4 and proposition 2, we distinguish the following three scenarios using the triggering values of pooling. First, when $\frac{m}{n\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) \geq \beta$, incumbents will separate in the decentralised as well as the centralised setting, so that $\lambda_i^D = \lambda_i^C = 0$ and (20) will be equal to zero. Inversely, when $\frac{1}{\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) < \beta$, pooling strategies are aligned so that $\lambda_i^D = \lambda_i^C = 1$, yet welfare will be higher under decentralisation as (20) collapses to $\frac{EW_i^D(\mu_i) - EW_i^C(\mu_i)}{(1 - \pi_i)} = \lambda_i^C \left(1 - \frac{m}{n}\right) \Delta_i^d > 0$. Lastly, when $\frac{m}{n\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) < \beta \leq \frac{1}{\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right)$, pooling strategies differ depending on the fiscal regime. Whereas incumbents will no longer pool in the decentralised setting, a central government still would -because of selective pooling- so that $\lambda_i^D = 0$ and $\lambda_i^C = 1$. Welfare gains under decentralisation then reduce to

$$\frac{EW_i^D(\mu_i) - EW_i^C(\mu_i)}{(1 - \pi_i)(1 - q)} = (\beta\pi_i\Delta_i^s - \Delta_i^d) + \left(1 - \frac{m}{n}\right) \Delta_i^d, \quad (21)$$

which will only be positive when $\pi_i > \frac{m}{n} \frac{\Delta_i^d}{\beta\Delta_i^s} = \bar{\pi}_i$. For β sufficiently large and since $\Delta_i^s > \Delta_i^d$, we have that $0 < \bar{\pi}_i < 1$. \square

Zooming in first on scenario 1 and 3 in proposition 3, where equilibrium strategies are aligned across fiscal regimes, we see that welfare nevertheless diverges between centralisation and decentralisation in scenario 3. When incumbents are certain to choose the pooling strategy in both regimes in other words, voters are better off when decision making is decentralised. The selective pooling reflex of centralised government is at play here, undermining the full potential of an outcome where the pooling strategy would be chosen in each state. In scenario 2 of the proposition electoral strategies do differ between fiscal regimes, as incumbents pool in a centralised system but separate under decentralisation. Voter welfare now depends on the “quality” of the pool of politicians voters can choose from. Characterised by a threshold value of $\bar{\pi}_i = \frac{m}{n} \frac{\Delta_i^d}{\beta\Delta_i^s} > 0$, decentralisation only improves voter welfare when the quality of politicians is sufficiently high, so that $\pi_i > \bar{\pi}_i$.

What emerges in this second scenario in other words, and similar to Besley and Smart (2007), is the relative importance of the selection effects vis-à-vis the disciplining effects of an election. If the pool of politicians mostly consists of benevolent politicians, strengthening the selection effect -here through increased separation with decentralisation- serves voter welfare more than improving discipline, and vice versa. Indeed, since pooling incumbents will eventually divert maximum rents in a future term, replacing bad incumbents as soon as possible is welfare-improving if a sufficient amount of benevolent alternatives is at hand. Unsurprisingly then, a similar trade-off between selection and discipline effects presents itself when a change in the re-election probabilities δ_i of incumbent politicians leads to a shift in equilibrium strategies. We analyse the welfare effects of such a shift in proposition 4, focusing on the decentralised setting where -following proposition 2- the impact of changes in δ_i will be highest.²¹

²¹See appendix V.D for a proof.

Proposition 4 *Rent-seeking politicians facing a higher probability of re-election, either through an increase in tax autonomy ν_i or a larger share of priority voters ω_i , will be quicker to pool. If the quality of politicians is sufficiently low -such that $\pi_i < \frac{\Delta_i^d}{\beta\Delta_i^s}$ - these gains in discipline always improve voter welfare, and vice versa.*

Naturally, proposition 4 also implies that when a large share of politicians turns out to be benevolent, so that $\pi_i \geq \frac{\Delta_i^d}{\beta\Delta_i^s}$, improved discipline will in fact undermine voter welfare. In this case, the degree to which voters value the selection effect denoted by $\beta\pi_i\Delta_i^s$ in (18) rises, as a larger weight is ascribed to the selection loss Δ_i^s as opposed to the discipline benefit Δ_i^d . Also, and importantly, what is omitted in proposition 4 is the effect of a change in the marginal cost of taxation μ_i on voter welfare. The reason is that, although such a shift will also alter the probabilities of re-election for incumbent politicians, a general welfare effect also comes into play via (1). A separate analysis in corollary 1 is therefore in order.²²

Corollary 1 *If the quality of politicians is low, so that $\pi_i \leq \frac{\Delta_i^d}{\Delta_i^s\beta}$, rising marginal costs of taxation unambiguously curtail voter welfare as more bad incumbents decide to separate. Inversely, when $\frac{\Delta_i^d}{\Delta_i^s\beta} < \pi_i$, voter welfare may increase as discipline subsides. These welfare shifts are more pronounced in a setting with priority voters.*

Contrary to common knowledge, and also pointed out by Besley and Smart (2007), increasing the inefficiency of a tax system through the marginal cost of taxation does not necessarily pay off in terms of voter welfare. Tax competition for example, which is thought to improve discipline and reign in rent-seeking, can in fact lead to the opposite outcome here. Driving up μ_i leads to lower rents r_i^1 and lower probabilities of re-election for incumbents, as specified above, and thus to more separation. If most politicians standing for office are also rent-seeking, voter welfare is hollowed out.

VI. CONCLUDING REMARKS

Decentralising tax authority to lower-level jurisdictions in a federation is often argued to improve the accountability of local politicians. In this paper, we derived the necessary conditions for tax autonomy to bring about local growth-enhancing policies -as the fiscal incentives approach of Weingast (2009) would predict- and investigate whether this mechanism is indeed beneficial to voter welfare. In this sense, we are first to model a multi-tiered, political agency setting where policy outcomes feed back into revenue flows, which indeed keeps rent-seeking politicians in line.

What we find is that everything in effect hinges on the quality of the pool of politicians voters can choose from, as well as on the composition of the voting population itself. If most voters do not have specific concerns and only care about economic growth, rent-seeking politicians will be hard put to improve their chances of re-election through pork-barrel targeting. Indeed, the more priority voters, the more the fiscal incentives will bite, and the less rents are diverted. However, we show that this “discipline” effect is stronger still in a unitary setting where all of public provision is kept at

²²A proof is given in appendix V.E.

the center. The reason is the “reduced pivot-probability effect”, coined by Lockwood (2006) and first introduced by Seabright (1996), where single jurisdictions become less pivotal in ensuring the re-election of a central government. Investing in a bare minimum of constituencies then suffices for a central government to be re-elected, making it more attractive for bad politicians to opt for this strategy of postponing maximum rent extraction.

Discipline will be more effective at the center in other words, *despite* the built-in fiscal incentives which are in fact mutually reinforcing in this case. When offered the choice between fiscal regimes consequently, voters would only prefer decentralisation if politicians are less likely to be rent-seeking to begin with, so that selection is needed more than discipline. Nevertheless, *given* a certain degree of decentralisation and a sufficient amount of rent-seeking politicians, shoring up discipline via sharper fiscal incentives is more effectively done at the lower level of government. Expanding local tax autonomy will in this case unambiguously boost voter welfare.

APPENDIX A. PROOF OF LEMMA V.1

The valence voter will always re-elect the incumbent after observing first period public provision of G_i^H when his posterior beliefs Π_i outweigh his prior beliefs π_i :

$$\Pr(g|T_i^H) = \Pi_i = \frac{\pi_i q_i}{\pi_i q_i + (1 - \pi_i)(1 - q_i)\lambda_i} > \pi_i. \quad (22)$$

Solving (22) for λ_i we get that

$$\Pi_i > \pi_i \Leftrightarrow \frac{q_i}{(1 - q_i)} > \lambda_i, \quad (23)$$

which, since $\lambda_i \in [0, 1]$, will always be the case as long as $q_i > \frac{1}{2}$. \square

APPENDIX B. DERIVATION OF LEMMA V.2

Let us first look at a ‘swing’ valence voter s whose ideological bias makes him indifferent between the two parties so that, after observing (G_i^H, T_i^H) in period 1, we get

$$\gamma_i^s = \Pi_i - \pi_i - \delta_i. \quad (24)$$

All valence voters j with $\gamma_i^j \leq \gamma_i^s$ thus prefer the incumbent grouping of politicians. Consequently, given our distributional assumptions, and using (2), the incumbent group can expect to win the following *overall* vote share μ_i after setting (G_i^H, T_i^H) in period 1:

$$\mu_i = \omega_i E(\eta[Y_i(G_i)]) + (1 - \omega_i) \left(\gamma_i^s + \frac{1}{2} \right). \quad (25)$$

Plugging in (24), we obtain

$$\mu_i = \omega_i E(\eta[Y_i(G_i)]) + (1 - \omega_i) \left(\Pi_i - \pi_i - \delta_i + \frac{1}{2} \right), \quad (26)$$

from which we get that

$$\Pr \left[\mu_i \geq \frac{1}{2} \right] = \Pr \left[\omega_i E(\eta[Y_i(G_i)]) + (1 - \omega_i) \left(\Pi_i - \pi_i - \delta_i + \frac{1}{2} \right) \geq \frac{1}{2} \right]. \quad (27)$$

Since we know that

$$E(\eta[Y_i(G_i)]) = \rho[Y_i(G_i)] \left(\frac{1}{2} + \chi \right) + (1 - \rho[Y_i(G_i)]) \left(\frac{1}{2} - \chi \right) = (2\rho[Y_i(G_i)] - 1)\chi + \frac{1}{2} > 0, \quad (28)$$

we can write (27) as

$$\Pr \left[\mu_i \geq \frac{1}{2} \right] = \Pr_{\delta_i} \left[\frac{\omega_i (2\rho[Y_i(G_i)] - 1)\chi}{(1 - \omega_i)} + (\Pi_i - \pi_i) \geq \delta_i \right], \quad (29)$$

or, setting $\frac{\omega_i (2\rho[Y_i(G_i)] - 1)\chi}{(1 - \omega_i)} = \alpha_i$, as

$$\Pr \left[\mu_i \geq \frac{1}{2} \right] = \Pr_{\delta_i} [\alpha_i + (\Pi_i - \pi_i) \geq \delta_i]. \quad (30)$$

Using (30), and given our distributional assumptions on δ_i , the probability for the group of incumbents of winning the elections then becomes

$$\Pr \left[\mu_i \geq \frac{1}{2} \right] = \sigma_i (\alpha_i + (\Pi_i - \pi_i)) = \begin{cases} 1 & \text{if } \alpha_i + (\Pi_i - \pi_i) > \frac{1}{2} \\ \frac{1}{2} + \xi (\alpha_i + (\Pi_i - \pi_i)) & \text{Otherwise} \\ 0 & \text{if } \alpha_i + (\Pi_i - \pi_i) < -\frac{1}{2}. \end{cases} \quad (31)$$

Of course, the implicit assumption behind the previous argumentation is that the bad incumbents will always use the additional revenue gains Y_i to cater to the priority vote, and thus improve their re-election chances σ_i . We assume this will be more beneficial than simply diverting away these additional revenues as rents, so that

$$\widehat{r}_i^1 + Y_i + \beta_i \sigma_i (\Pi_i - \pi_i) X_i < \widehat{r}_i^1 + \beta_i \sigma_i (\alpha_i + (\Pi_i - \pi_i)) X_i, \quad (32)$$

keeping in mind that $\alpha_i(G_i^H, \chi, \omega_i) = 0$ when no additional revenue Y_i is invested in the priority vote, because $\rho(Y_i = 0) = \frac{1}{2}$. Since Y_i concerns additional revenues, and X_i constitutes the bulk of public spending furthermore, (32) will hold in most realistic situations.

APPENDIX C. DERIVATION OF LEMMA V.3

The necessary condition for a group of bad incumbents to provide G_i^H when $\theta_i = H$ in m states -in other words, opt for the selective pooling strategy- will be

$$m\widehat{r}_i^1 + (n - m) X_i + n\beta\sigma_i (\alpha_i + (\Pi_i - \pi_i)) X_i > nX_i. \quad (33)$$

Note that, because the unit costs of public provision θ_i are assumed fully correlated across states, the central government can decide to pool in all of the $m = (n + 1)/2$ necessary states. Reworking (33) then gives us

$$m(\widehat{r}_i^1 - X_i) > -n\beta\sigma_i (\alpha_i + (\Pi_i - \pi_i)) X_i, \quad (34)$$

or, solving for β ,

$$\frac{m}{n} \frac{(X_i - \hat{r}_i^1)}{\sigma_i(\alpha_i + (\Pi_i - \pi_i)) X_i} < \beta, \quad (35)$$

so that, collecting terms, we obtain

$$\frac{m}{n\sigma_i(\alpha_i + (\Pi_i - \pi_i))} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) < \beta. \quad (36)$$

APPENDIX D. PROOF OF PROPOSITION V.4

Focusing on an increase of ν_i or ω_i , and thus limiting our attention to a rise in $\sigma_i(\alpha_i(G_i, \nu_i, \omega_i, \chi) + (\Pi_i - \pi_i))$ only, we have that

$$\frac{1}{\sigma_i(\alpha'_i + (\Pi_i - \pi_i))} \left(1 - \frac{\hat{r}_i^{1'}}{X_i}\right) = \beta_i^{D'} < \beta_i^D = \frac{1}{\sigma_i(\alpha_i + (\Pi_i - \pi_i))} \left(1 - \frac{\hat{r}_i^1}{X_i}\right), \quad (37)$$

where β_D and β'_D again denote the triggering values for the incumbents to pool, but now before and after a shift in ν_i or ω_i respectively. This upwards shift leads to higher levels of $\alpha_i(G_i, \nu_i, \omega_i, \chi)$ as proven in proposition V.1, which we mark out in (37) as $\alpha'_i > \alpha_i$. Since σ_i denotes a probability furthermore and $\frac{d\sigma_i}{d\alpha_i} > 0$ following lemma V.2, the direction of the inequality sign in (37) follows. Turning now to welfare effects, we write post-increase welfare as

$$EW_i^{D'}(\mu_i) = \bar{W}_i + \lambda'_i(1 - \pi_i)(1 - q) (\Delta_i^d - \beta\pi_i\Delta_i^s), \quad (38)$$

with $\lambda_i^{D'}$ the altered pooling probabilities after the increase. Subtracting (18) from (38), we then derive the potential welfare gains of an increase in ν_i or ω_i as

$$\frac{EW_i^{D'}(\mu_i) - EW_i^D(\mu_i)}{(1 - \pi_i)(1 - q)} = (\lambda_i^{D'} - \lambda_i^D) (\Delta_i^d - \beta\pi_i\Delta_i^s). \quad (39)$$

Using lemma V.3 and (37), we can once more distinguish three scenarios using the triggering value of pooling. First, when $\frac{1}{\sigma_i(\alpha'_i)} \left(1 - \frac{\hat{r}_i^{1'}}{X_i}\right) \geq \beta$, incumbents will separate in the decentralised as well as the centralised setting, so that $\lambda_i^D = \lambda_i^{D'} = 0$ and (39) will be equal to zero. Inversely, when $\frac{1}{\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right) < \beta$, pooling strategies are aligned so that $\lambda_i^D = \lambda_i^{D'} = 1$ and (39) is again equal to zero. Lastly, when $\frac{1}{\sigma_i(\alpha'_i)} \left(1 - \frac{\hat{r}_i^{1'}}{X_i}\right) < \beta \leq \frac{1}{\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i}\right)$, pooling strategies will be different before and after the shift in ν_i or ω_i . Where incumbents would have separated before the rise in re-election probabilities, they will now keep on pooling so that $\lambda_i^D = 0$ and $\lambda_i^{D'} = 1$. Welfare gains after the increase then become

$$\frac{EW_i^{D'}(\mu_i) - EW_i^D(\mu_i)}{(1 - \pi_i)(1 - q)} = (\Delta_i^d - \beta\pi_i\Delta_i^s), \quad (40)$$

which will only be positive when $\pi_i < \frac{\Delta_i^d}{\beta\Delta_i^s}$. Again, for β sufficiently large and since $\Delta_i^s > \Delta_i^d$, we have that $0 < \bar{\pi}_i < 1$. \square

APPENDIX E. PROOF OF COROLLARY V.1

Taking the total derivative of (18) with respect to μ_i , we obtain

$$\frac{\partial EW_i^D(\mu_i)}{\partial \mu_i} = \frac{\partial \bar{W}_i(\mu_i)}{\partial \mu_i} + \frac{\partial \lambda_i}{\partial \mu_i} (1-\pi_i)(1-q) \left(\Delta_i^d - \beta \pi_i \Delta_i^s \right) + \lambda_i (1-\pi_i)(1-q) \left(\frac{\partial W_i^g(H, \mu_i)}{\partial \mu_i} + C_i(X_i) - \beta \pi_i \left(\frac{\partial EW_i^g(\mu_i)}{\partial \mu_i} + C_i(X_i) \right) \right), \quad (41)$$

with,

$$\frac{\partial \bar{W}_i(\mu_i)}{\partial \mu_i} = \frac{\partial W_i^0(\mu_i)}{\partial \mu_i} + \beta \left(\pi_i \frac{\partial EW_i^g(\mu_i)}{\partial \mu_i} + (1-\pi_i) \frac{\partial W_i^0(\mu_i)}{\partial \mu_i} \right) < 0, \quad (42)$$

$$\frac{\partial W_i^0(\mu_i)}{\partial \mu_i} = \pi_i \frac{\partial EW_i^g(\mu_i)}{\partial \mu_i} - (1-\pi_i) C_i(X_i) < 0, \quad (43)$$

and,

$$\frac{\partial EW_i^g(\mu_i)}{\partial \mu_i} = q \frac{\partial W_i^g(H, \mu_i)}{\partial \mu_i} + (1-q) \frac{\partial W_i^g(L, \mu_i)}{\partial \mu_i} < 0, \quad (44)$$

$$\frac{\partial W_i^b(\mu_i)}{\partial \mu_i} = -C_i(X_i) < 0. \quad (45)$$

Suppose now that for $\mu_i = \bar{\mu}_i$, we have that (10) becomes

$$\hat{r}_i^1(\mu_i) + \beta \sigma_i (\alpha_i (G_i, \nu_i, \chi, \omega_i) + (\Pi_i - \pi_i)) X_i = X_i. \quad (46)$$

Since $\frac{d\hat{r}_i^1(\mu_i)}{d\mu_i} < 0$ and $\frac{d\sigma_i}{d\mu_i} < 0$ as specified earlier, we obtain for all values $\mu_i < \bar{\mu}_i$ that

$$\hat{r}_i^1(\mu_i) + \beta \sigma_i (\alpha_i (G_i, \nu_i, \chi, \omega_i) + (\Pi_i - \pi_i)) X_i > X_i, \quad (47)$$

or that $\frac{1}{\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i} \right) < \beta$. From lemma 3 we then find that in the resulting pooling equilibrium $\frac{\partial \lambda_i}{\partial \mu_i} = 0$, and $\lambda_i = 1$. This allows us to write (41) as

$$\frac{\partial EW_i^D(\mu_i)}{\partial \mu_i} = \frac{\partial \bar{W}_i(\mu_i)}{\partial \mu_i} + (1-\pi_i)(1-q) \left(\frac{\partial W_i^g(H, \mu_i)}{\partial \mu_i} + C_i(X_i) - \beta \pi_i \left(\frac{\partial EW_i^g(\mu_i)}{\partial \mu_i} + C_i(X_i) \right) \right), \quad (48)$$

which, plugging in (42) and (43), and collecting terms, yields

$$\frac{\partial EW_i^D(\mu_i)}{\partial \mu_i} = \pi_i (1 + \beta + \beta q - \beta q \pi_i) \frac{\partial EW_i^g(\mu_i)}{\partial \mu_i} - (1-\pi_i) \left((q + \beta(1-\pi_i) + \beta(1-q)\pi_i) C_i(X_i) - (1-q) \frac{\partial W_i^g(H, \mu_i)}{\partial \mu_i} \right) = \Psi_i, \quad (49)$$

where, using (44) and (45), we know that $\Phi_i < 0$. Inversely, when $\mu_i \geq \bar{\mu}_i$, we get that

$$\hat{r}_i^1(\mu_i) + \beta \sigma_i (\alpha_i (G_i, \nu_i, \chi, \omega_i) + (\Pi_i - \pi_i)) X_i \leq X_i. \quad (50)$$

From lemma V.3 we know that when $\beta \leq \frac{1}{\sigma_i(\alpha_i)} \left(1 - \frac{\hat{r}_i^1}{X_i} \right)$ results in a separating equilibrium where $\frac{\partial \lambda_i}{\partial \mu_i} = 0$ and $\lambda_i = 0$. We can then write (41) simply as

$$\frac{\partial EW_i^D(\mu_i)}{\partial \mu_i} = \frac{\partial \bar{W}_i(\mu_i)}{\partial \mu_i} < 0. \quad (51)$$

In the neighbourhood of $\bar{\mu}_i$ lastly, we know a value $\mu_i \lesssim \bar{\mu}_i$ exists for which a marginal increase implies a shift from the pooling to the separating equilibrium according to lemma 3, so that $\frac{\partial \lambda_i}{\partial \mu_i} = -1$ and $\lambda_i = 1$. From (41), and using (49), we now obtain

$$\frac{\partial EW_i^D(\mu_i)}{\partial \mu_i} = \Psi_i - (1 - \pi_i)(1 - q) (\Delta_i^d - \beta \pi_i \Delta_i^s), \quad (52)$$

which will only be positive if and only if

$$(1 - \pi_i)(1 - q) (\Delta_i^d - \beta \pi_i \Delta_i^s) < \Psi_i. \quad (53)$$

Since we know from (49) that $\Psi_i < 0$, this implies that $\frac{\Delta_i^d}{\Delta_i^s \beta} < \pi_i$ is a necessary condition for (53) to hold. The last part of the corollary is proven by considering a discrete jump off μ_i rather than thinking on the margin. As proven in proposition V.2, any discrete jump in μ_i -as well as the resulting change in β_i defining the interval where politicians switch equilibrium strategies- will be larger the larger α_i since $\frac{d\beta_i}{d\alpha_i} > 0$, and thus the more priority voters in the population as $\frac{d\alpha_i(G_i, \nu_i, \omega_i, \chi)}{d\omega_i} > 0$. This proves the last part of the proposition. \square

REFERENCES

- Allers, M. A. and Elhorst, J. P. (2005). Tax mimicking and yardstick competition among local governments in the netherlands. *International tax and public finance*, 12(4):493–513.
- Barankay, I. and Lockwood, B. (2007). Decentralization and the productive efficiency of government: Evidence from swiss cantons. *Journal of public economics*, 91(5):1197–1218.
- Besley, T. and Burgess, R. (2001). Political agency, government responsiveness and the role of the media. *European Economic Review*, 45(4):629–640.
- Besley, T. and Case, A. (1995). Incumbent behavior: Vote seeking, tax setting and yardstick competition. *American Economic Review*, 85(1):25–45.
- Besley, T., Persson, T., and Sturm, D. M. (2010). Political competition, policy and growth: theory and evidence from the us. *The Review of Economic Studies*, 77(4):1329–1352.
- Besley, T. and Smart, M. (2007). Fiscal restraints and voter welfare. *Journal of Public Economics*, 91(3):755–773.
- Boadway, R. and Tremblay, J.-F. (2012). Reassessment of the tiebout model. *Journal of public economics*, 96(11):1063–1078.
- Boetti, L., Piacenza, M., and Turati, G. (2012). Decentralization and local governments' performance: how does fiscal autonomy affect spending efficiency? *FinanzArchiv: Public Finance Analysis*, 68(3):269–302.

- Bordignon, M., Cerniglia, F., and Revelli, F. (2003). In search of yardstick competition: a spatial analysis of italian municipality property tax setting. *Journal of Urban Economics*, 54(2):199–217.
- Bosch, N. and Solé-Ollé, A. (2007). Yardstick competition and the political costs of raising taxes: An empirical analysis of spanish municipalities. *International Tax and Public Finance*, 14(1):71–92.
- Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2013). The political resource curse. *American Economic Review*, 103(5):1759–96.
- Cadot, O., Röller, L.-H., and Stephan, A. (2006). Contribution to productivity or pork barrel? the two faces of infrastructure investment. *Journal of public Economics*, 90(6):1133–1153.
- Faguet, J.-P. (2004). Does decentralization increase government responsiveness to local needs?: Evidence from bolivia. *Journal of public economics*, 88(3):867–893.
- Geys, B., Heinemann, F., and Kalb, A. (2010). Voter involvement, fiscal autonomy and public sector efficiency: evidence from german municipalities. *European Journal of Political Economy*, 26(2):265–278.
- Hindriks, J. and Lockwood, B. (2009). Decentralization and electoral accountability: incentives, separation and voter welfare. *European Journal of Political Economy*, 25(3):385–397.
- Hindriks, J., Peralta, S., and Weber, S. (2008). Competing in taxes and investment under fiscal equalization. *Journal of public economics*, 92(12):2392–2402.
- Jin, H., Qian, Y., and Weingast, B. R. (1999). Regional decentralization and fiscal incentives: federalism.
- Jin, H., Qian, Y., and Weingast, B. R. (2005). Regional decentralization and fiscal incentives: Federalism, chinese style. *Journal of public economics*, 89(9):1719–1742.
- Kappeler, A., Solé-Ollé, A., Stephan, A., and Väililä, T. (2013). Does fiscal decentralization foster regional investment in productive infrastructure? *European Journal of Political Economy*, 31:15–25.
- Keen, M. and Marchand, M. (1997). Fiscal competition and the pattern of public spending. *Journal of Public Economics*, 66(1):33–53.
- Lockwood, B. (2005). A note on the hybrid equilibrium in the besley-smart model.
- Lockwood, B. (2006). The political economy of decentralization. In *Handbook of multi-level finance*, chapter 2. Ahmad, Ehhisham and Brosio, Giorgio.
- Persson, T. and Tabellini, G. (2002a). Political economics and public finance. *Handbook of public economics*, 3:1549–1659.

- Persson, T. and Tabellini, G. E. (2002b). *Political economics: explaining economic policy*. MIT press.
- Revelli, F. and Tovmo, P. (2007). Revealed yardstick competition: Local government efficiency patterns in norway. *Journal of Urban Economics*, 62(1):121–134.
- Roberson, B. (2008). Pork-barrel politics, targetable policies, and fiscal federalism. *Journal of the European Economic Association*, 6(4):819–844.
- Seabright, P. (1996). Accountability and decentralisation in government: An incomplete contracts model. *European Economic Review*, 40(1):61–89.
- Weingast, B. R. (1995). The economic role of political institutions: market-preserving federalism and economic development. *Journal of Law, Economics, & Organization*, pages 1–31.
- Weingast, B. R. (2009). Second generation fiscal federalism: The implications of fiscal incentives. *Journal of Urban Economics*, 65(3):279–293.
- Weingast, B. R. (2014). Second generation fiscal federalism: political aspects of decentralization and economic development. *World Development*, 53:14–25.

