Centralization and Political Accountability*

Jean Hindriks† Ben Lockwood‡

July 15, 2004

Abstract

In this paper we abstract from the usual gains and costs of decentralization (e.g. preference matching, spillovers and economies of scale). Instead we compare the political accountability of decentralized governments relative to centralized ones when there is a risk of "bad" governance. We study both the selection and incentive effects of accountability. A key aspect of centralization is to make the politician answerable to multiple constituencies subject to a common budget constraint. Our main findings are that (a) when politicians differ in competence, decentralization unambiguously dominates; and (b) when politicians differ in honesty, decentralization and centralization have conflicting accountability effects (when one provides better discipline, the other gives better selection). The analysis then identifies the circumstances under which centralization may increase voter welfare. The more general lesson that we can draw is that different institutional forms give rise to different information to the voters on which electoral accountabiltiy can be based. Therefore they differ on how effective elections can be in disciplining and selecting policymakers.

*We are grateful to the participants of the conference on "Political Economy and Institutional Performance" (Brussels, June 11-12, 2004) and seminar participants at Warwick, Nottingham, Milan and Louvain for their comments. We would also like to thank Jean Tirole, Massimo Bordignon, Micael Castanheira, and Christian Schultz for helful comments on a previous draft.

†Department of Economics, CORE, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. Email: Hindriks@core.ucl.ac.be

‡CEPR and Department of Economics, University of Warwick, Coventry CV4 7AL, England; Email: B.Lockwood@warwick.ac.uk

1. Introduction

Fiscal decentralization, the allocation of tax and spending powers to lower levels of government, is now an established policy objective, in many developed and developing countries. Moreover, it is actively promoted as a development strategy by organizations such as the World Bank (Azfar et al.(2001), World Bank(2000)). The usual advantages that are claimed for decentralization that one can find in the literature include the following (Azfar et al.(2001), Oates(1999)). First, decentralization is claimed to improve *allocative efficiency*, in the sense that the goods provided by governments in localities will be better matched to the preferences of the residents of those localities. This is sometimes known as the *preference-matching* argument. There is now a large theoretical literature evaluating the preference-matching argument.[1] There are also few more recent empirical papers.[2]

Second, decentralization is argued to increase the *accountability* of government. In the literature, this term is used in rather a broad sense, and refers to electoral rules and other institutional mechanisms that constrain the rent-seeking activities of office holders, such as taking bribes, favouring of particular interest groups, and insufficient innovation and effort. Interestingly, in this case, the lead has been taken by empirical researchers: there are now a number of cross-sectional and panel studies that show that across countries, measures of fiscal decentralization may sometimes be negatively correlated with outcomes associated with low accountability, such as corruption and poor governance (e.g. Huther and Shah(1998)), Mello and Barenstein(2001), Fissman and Gatti(2000), Triesman(2002)).

However, accountability is notoriously difficult to pin down precisely, and perhaps reflecting this, there have been rather few attempts to precisely define accountability, and analyse theoretically how the degree of accountability varies with fiscal (de)-centralization. Notable exceptions are Seabright(1996) and Bardhan and Mookherjee(2000).[3] Both of these contributions, while important steps forward, have their limitations. The first two papers have models where the government has to make an abstract policy choice (a point on the real line), and cannot capture one of the key features of fiscal centralization, namely

[1] See for example, Alesina and Spolare(1997), Besley and Coate(2003), Bolton and Roland(1997) and Cremer and Palfrey(1996), Ellingsen(1998), Gilbert and Picard(1996), Lockwood(2002), Oates(1972), Oberholzer-Gee and Strumpf(2002), Wallis and Oates(1988).

[2] Oberholzer-Gee and Strumpf(2002), Faguet(2004).

[3] We should also mention the important contributions of Besley and Smart (2003) and Belleflamme and Hindriks (2003), who focus on the impact of competition between jurisdictions (tax or yardstick) on the performance of politicans in a two-period agency framework. However, they do not compare centralisation and decentralisation: their baseline case is decentralisation without competition.

that the policy-maker faces a single budget constraint, where local public goods are funded by a national tax.[4] One of the central contributions of our paper is to argue that this institutional feature of centralization changes the incentives for a "bad" incumbent to pool with a good incumbent. These contributions are both discussed in more detail in Section 7 below.

This paper takes a somewhat different approach to modelling accountability. We consider a two-jurisdiction, two-period political agency model of the type pioneered by Rogoff(1990), and developed subsequently by many others.[5] In each period and each jurisdiction, the elected politician provides a local public good financed by taxes. At the end of the first period, the incumbent politician fights an election against a challenger. There are just two differences between decentralization and centralization. First, with decentralization, there are two politicians, one in each region, whereas with centralization, there is a single politician. Second, with centralization, the tax rate that funds public good provision is assumed the same in the two regions - *tax uniformity*.[6] Politicians - both incumbent and challenger - may be "good" or "bad". We look at two possibilities: the bad politician may be incompetent or dishonest. In this set-up, elections provide accountability in two senses. First, they allow voters to de-select bad incumbents (*selection effects*). Second, the selection effect provides an incentive for incumbents to change their behavior in order to increase the probability of re-election (*incentive or discipline effects*).

The focus of this paper is on how centralization affects these measures of accountability. Our main results are as follows. First, in the competence case (where politicians have a high or low expected cost of providing the public good), with centralization, it is unambiguously more likely that an incompetent incumbent will imitate, or "pool" with a competent one, in order to be re-elected. This is because to get re-elected the bad incumbents must produce the same record as a competent politician in *both* regions, while with decentralization a good record at home suffices to gain re-election. However, pooling is bad both in terms of selection and incentive. The bad selection effect is obvious: with pooling, it is more likely that an incompetent incumbent is re-elected. The incentive

[4]Persson and Tabellini(2000), Chapter 9.1 have a two-period version of the Seabright model with explicit modelling of fiscal policy, but again they impose the assumption of separate budget constraints. Their interpretation of this model is that it allows comparison of majoritarian and proportional electoral rules in an agency framework.

[5]Important contributions include Harrington (1993), Hess and Orphanides (1995), Besley and Case (1995), Coate and Morris(1995) and Drazen (2000b),

[6]The first feature is the basic or defining one of fiscal centraliation. The second feature, tax uniformity, is a near universal feature empirically: with very few exceptions, national tax codes do not differentiate between regions. Tax uniformity is discussed in more detail below.

effect is also bad as the incompetent incumbent distorts local public good provision (in fact, over-provides the public good) in order to imitate the good incumbent. So, the conclusion concurs with the conventional wisdom: centralization unambiguously reduces accountability, and indeed lowers voter welfare.[7]

In the honesty case, (where honest politicians are motivated only by voter welfare, and dishonest politicians only by rents diverted from tax revenue) the picture is more complex, and more interesting. First, with centralization the incumbent must pool in both regions or neither, and this constraint gives different incentives to pool in the centralized case. So depending on parameter values, centralization can make pooling more *or* less likely. Moreover, now, pooling (relative to separating) has a *good* incentive effect, although the selection effect is of course weakened by pooling. Specifically, in order to pool, the incumbent must limit the amount that is stolen from the public purse. So, centralization always has an ambiguous effect on accountability overall: when it leads to more pooling, it decreases the selection effect, but improves incentives, and when it leads to less pooling, the opposite is true. Not surprisingly, therefore, centralization may increase or decrease voter welfare.

The question then arises as to which institutional feature of centralization (a single policy-maker or tax uniformity) is driving these results. The answer depends crucially on what information voters have at the time when they vote. If voters can observe *only* the fiscal policy choices in their own jurisdictions, then tax uniformity drives the results. If, on the other hand, voters can observe the fiscal policy choices in *both* jurisdictions, then the fact of a single policy-maker drives the results. That is, it makes no difference to the equilibrium outcome whether the policy-maker can choose separate taxes in the two jurisdictions, or is constrained to a uniform tax.

The intuition is the following. A difference between centralization and decentralization only arises when regions are asymmetric ex post - i.e. when the cost of producing the public good is high in one region and low in the other. In that situation, with non-uniform taxes, and limited voter information, a bad incumbent can pool (i.e. imitate the good politician) in the low-cost region, and separate in the high-cost region, thus only losing the votes of residents of the high-cost region, and thus having a higher probability of re-election. This possibility to pool "selectively" gives the incumbent the same incentive to pool as under decentralization. With uniform taxes, and/or voter information about fiscal policy in the other region, such selective pooling is not possible: the incumbent must pool in both regions or neither, and this constraint gives different incentives to pool

[7] The measure of voter welfare is the expected present value of utility.

in the centralized case.

A final contribution of this paper is to investigate how the results are affected by the introduction of competition between jurisdictions. We focus on the case of yardstick competition, with politicians who differ in benevolence (honest or dishonest).[8] We show that if there is a pooling equilibrium with yardstick competition, decentralization always yields higher voter welfare if the correlation in the (random) cost of producing the public good is sufficiently high across the two jurisdictions. If on the other hand, there is a hybrid equilibrium with yardstick competition in which politicians play mixed strategy between pooling and separating, correlation in the cost of producing the public good makes decentralization *less* likely to dominate.

The layout of the remainder of the paper is as follows. Section 2 sets up the model. Section 3 presents the competence case and proceeds to welfare comparison between centralization and decentralization. Section 4 presents the benevolence case and describes the different equilibrium configurations before undertaking the welfare comparisons. Section 5 studies the welfare effects of yardstick comparison. Section 6 concludes with a discussion of the related literature.

2. The Model

2.1. Preliminaries

There are two time periods $t = 1, 2$ and two regions $i = a, b$. In each region in each time period, a politician makes a decisions about taxation and public good provision. Moreover, at the end of period 1, there is an election in which voters choose between the incumbent and a challenger, having observed only first-period fiscal policy. With decentralization, there are two politicians: one in each region. With centralization, there is one politician.

In each region, there are a continuum of measure 1 of identical voters who derive utility $u_t^i = H(g_t^i) + x_t^i$ from a regional public good g_t^i and a private good x_t^i in period t. All agents have an endowment of the private good, normalized to unity. The public good is financed by a lump-sum tax τ_t^i, so that utility of the typical voter is $H(g_t^i) + 1 - \tau_t^i$. The tax can also be interpreted as an income tax at rate τ_t^i on income of unity. It is assumed that $0 \leq \tau_t^i \leq \overline{r}$, where $\overline{r} \leq 1$ is a maximum feasible tax rate. Both voters and politicians have the same discount factor, $0 < \delta < 1$.

[8] The effect of yardstick competition in the competence model is less interesting since decentralization already dominates and observing fiscal policy in one region does not bring any useful information about the politician type in the other region (unless we assume correlated politician types across regions).

In each region in each time period, the unit cost c_t^i of producing the public good from the private good can take on one of two values: $c_t^i \in \{c_L, c_H\}$ with $c_L < c_H$. The determination of c_t^i is described in more detail below. With decentralization, there is a separate budget constraint for each region, of the form:

$$c_t^i g_t^i + r_t^i = \tau_t^i$$

where r_t^i are the rents diverted from tax revenue (if any) in region i. With centralization, the policy-maker is assumed to be able to pool tax revenues, and so faces a single budget constraint. Also, following almost all the literature on fiscal centralization, we assume that the tax rate is uniform with centralization i.e. $\tau_t^i = \tau_t$. So, the budget constraint is

$$\sum_{i=a,b} c_t^i g_t^i + r_t^i = 2\tau_t$$

where r_t^i are the rents (if any) diverted from aggregate tax revenue.[9]

Politicians may be of two types, "good" and "bad". In particular, in either region, both[10] the initial incumbent and the challenger at the election are "good" with probability π and "bad" with probability $1-\pi$. Politicians may differ in competence or benevolence, giving rise to two variants of the model.

Competence. Any politician maximises the sum or average of the utilities of citizens in his jurisdiction, and . also derives "ego-rent" from office of R. A "good" politician is more competent than the bad. The public good is provided via a technology where the probability q_t^i that the unit cost is high in region i at time t is (i) uncorrelated across time and regions, and (ii) is conditional on the competency of the incumbent. In particular, if the incumbent is good, then $q_t^i = 0$, and if the incumbent is bad, $q_t^i = q$, with $1 > q > 0$.

To avoid making technological assumptions that bias the model in favour of some institutional form and to concentrate on the accountability effect, we make the *invariance assumption* that the joint distribution of costs across regions is the same under centralization and decentralization. Note that because the cost distribution in region i at time t is assumed specific to the incumbent, the invariance assumption requires that under decentralization, incumbent types are perfectly correlated.[11]

[9] Note that as the budget constraint is national, only the aggregate rent matters.

[10] This leaves open the possibility that with decentralisation, incumbent types are either independent, or correlated. In the benevolence case, incumbent types are assumed independently drawn. In the compentence case, correlation is assumed, for reasons discussed below.

[11] This condition of homogenous incumbents with decentralization is not required in the benevolent variant of the model where costs and types are independent (either type is equally competent to produce the public good).

As in Rogoff(1990), we assume a mechanism by which a first-period bad incumbent can "hide" the true cost of public good provision from the voters.[12] So, we assume that in the first period, the incumbent can borrow freely on the international capital market at interest rate $\frac{1}{\delta} - 1$, and so in the second period, he must pay back $\$\frac{1}{\delta}$ for every \$ borrowed. The borrowing is not observable by voters.

Benevolence. A "good" politician derives utility only from the welfare of the voters in his jurisdiction: in particular, he maximises the sum or average of these utilities. A bad politician cares only about rent diversion. Either type is equally competent in producing the public good. The cost of the public good is high in either region and period with probability q.

2.2. A Benchmark

Note that in this model, there is an agency problem between voters and the incumbent: the former can only imperfectly control the behavior of the latter through electoral incentives. Note also that in setting up this model, we have abstracted from the usual features that generate a difference between centralization and decentralization in the established literature: there are no spillovers between regions, voters do not differ in tastes for the public good, either within or between regions, and the (distribution of) cost of producing the cost is the same under either institutional form. Also, by taking a fixed number of incumbents and challengers, we assume away free entry and rule out the district magnitude effect bias in favor centralization (that larger electoral districts lower barriers to entry and favor competition improving political discipline and selection).[13] So, the difference in outcome between centralization and decentralization is *entirely* due to the difference in the extent to which the voters can control, or hold accountable, the incumbent, in the two cases.

To see this, it is helpful to consider the benchmark where there is no agency problem i.e. where politicians are "good" with probability 1. In this case, it is clear that there is an equilibrium where the incumbent will always be re-elected,[14] and in either region and

[12] Otherwise, given that voters can observe both public good provision and the tax rate, the only possible equilibrium would be a separating one with no distortions in public good supply i.e. no agency "problem" in which case centralization and decentralization are again equivalent.

[13] The district magnitude effect is related to the idea suggested by Myerson (1993) that electoral rules promoting the entry of many candidates protect voters against corruption in a better way. Myerson (1999) gives overview of the performance of different electoral systems and Persson et al (2001) give evidence of the district magnitude effect.

[14] There can be other equilibria where the incumbent is no re-elected, as all voters are always indifferent

period, the incumbent will provide the public good efficiently, conditional on cost c_L or c_H. This is true whether there is centralization or decentralization. Of course, efficient provision is $g_k, k = L, H$ is implicitly defined by the Samuelson rule $H'(g_k) = c_k$. Finally, as the distribution of costs is the same under decentralization and centralization, it follows that public good provision and therefore expected voter welfare must also be the same.

2.3. Relation to the Literature

The model of competence is based on Rogoff (1990), but with some changes that make analysis of the model simpler. In particular, the assumption of international borrowing to "hide" the true resource cost of the public good is an alternative to Rogoff's infrastructure public good which is consumed with one period delay. The other key difference to Rogoff (1990) is that we assume that if voters believe the incumbent and challenger to be equally competent, they will vote for the incumbent[15]. The model of benevolence is essentially that of Besley and Smart (2003). However, as emphasized in the introduction, they did not consider centralization in this setting.

3. Competence

3.1. Decentralization

We solve the model backwards in the usual way. In the second period, conditional on first-period borrowing b, the incumbent just sets $(g_k, \tau_k + \frac{b}{\delta})$ if $c = c_k$ - quasi-linear preferences ensures that public good supply is independent of loan repayment. So, given borrowing b, second-period payoffs to voters from good incumbent is $W_L - \frac{b}{\delta}$ and expected payoff from bad incumbent is $EW - \frac{b}{\delta}$ where

$$EW = qW_H + (1-q)W_L,\ W_H = H(g_H) - c_H g_H,\ W_L = H(g_L) - c_L g_L.$$

So voters prefer to vote out bad incumbent if possible ($EW < W_L$).

In the first period, the politician observes the unit cost c_L or c_H – we drop the regional subscript for convenience - and then chooses a level of provision conditional on

between incumbent and challenger, but these will generate the same outcome as the first one when there is no agency problem.

[15]So, the voters are assumed to have a lexicographic second preference for the incumbent. This tie-breaking rule is not innocuous, as it implies that in equilibrium, the competent type behaves non-strategically,with the non-competent deciding whether to imitate him or not: this is rather different to Rogoff.

cost. Voters observe level of public good and tax collection prior to election. Voters make an inference about their incumbent's type based on observed performance and compare it to prior beliefs about the type of the challenger, and re-elect their incumbent if the probability that he is "good" is at least π. The incumbent gets rent R if he is re-elected. Given this, it is clear that good incumbents behave non-strategically, by choosing g_L, the optimal supply when cost is low. This is because if the voters observe g_L, τ_L, whatever strategy the "bad" incumbent follows, rational voters must conclude that the probability that the incumbent is "good" conditional on g_L, τ_L is at least π.

As for the bad incumbent, if cost is low, he cannot do better than set (g_L, τ_L). If cost is high, he has two options. First, he can separate and set (g_H, τ_H). Second, he can pool and imitate the public good provision and taxation of a good type (g_L, τ_L) by borrowing $\hat{b} = (c_H - c_L)g_L$. If he separates, he reveals himself to be bad and thus loses the election. If he pools with the good type, his type is not revealed and he will be re-elected.

His payoff to separating is therefore

$$W_D^{sep} = W_H + \delta\left[\pi W_L + (1-\pi)EW\right]$$

that is, in the second period, the incumbent rationally anticipates that he will be replaced by a challenger who is competent with probability π. His payoff to pooling is

$$W_D^{pool} = W_L + \delta\left[R + EW - \frac{\hat{b}}{\delta}\right]$$

where, in the second period, the incumbent rationally anticipates that he will win the election, and that the debt incurred in order to pool must be repaid. So, bad incumbent will pool if $W_D^{pool} \geq W_D^{sep}$ which reduces to

$$R \geq \frac{\Delta}{\delta} + \pi q(W_L - W_H) \equiv R_D, \tag{3.1}$$

where $\Delta = W_H - (H(g_L) - c_H g_L)) > 0$. On the RHS are the *incentive* and *selection* costs of pooling to all voters - and thus to the "bad" incumbent himself. The incentive cost, $\Delta > 0$ is the welfare loss of the distortion in public good supply. The selection cost, $\pi q(W_L - W_H)$ is the loss in second-period expected utility from not replacing a bad incumbent. So, there is a pooling equilibrium if (3.1) holds, and a separating equilibrium otherwise.

3.2. Centralization

Now there is a single decision maker for both regions, and the tax rate that funds public good provision in both regions must be the same. The incumbent gets total rent $2R$ if he

is re-elected. The analysis is the same as in the decentralization case, except for the first-period pooling/separating decision of the bad incumbent. In particular, in the second-period , an incumbent of either type chooses optimal public good provision, conditional on cost, and the good incumbent chooses the same in the first-period.

Now consider the bad incumbent in the first-period. There are different cost states to consider. In "symmetric" cost state[16] LL, the bad incumbent cannot do better than set (g_L, τ_L) in both regions and get re-elected. In state HH, the pooling constraint is (3.1) again, as the costs and benefits of pooling are as in the decentralized case, scaled up by a factor of 2. In "asymmetric" cost states HL or LH, however, the choice to pool is different. Here, if the bad incumbent separates, it is optimal to do so by providing the public good at the optimal level in each region, and setting a tax $\tau = \frac{1}{2}(c_H g_H + c_L g_L)$. His payoff to separating is therefore

$$W_C^{sep} = W_H + W_L + 2\delta\left[\pi W_L + (1-\pi)EW\right]$$

If he pools, he only need distort public good supply in the high-cost region which requires to borrow $\hat{b}$. His payoff to pooling is therefore

$$W_C^{pool} = 2W_L + \delta\left[2R + 2EW - \frac{\hat{b}}{\delta}\right]$$

Let us call this strategy *partial pooling* as it only prescribes pooling in the asymmetric cost state. The condition for pooling in the asymmetric cost state HL is $W_C^{pool} \geq W_C^{sep}$, which gives

$$R \geq \frac{\Delta}{2\delta} + \pi q(W_L - W_H) \equiv R_C, \tag{3.2}$$

Note that in an asymmetric cost state, pooling requires distortion in public good supply in *only* the high-cost region, so the incentive cost of pooling is $\frac{\Delta}{\delta}$, the same as in the decentralized case, whereas the selection cost and the gain in rent to the incumbent are double their values in the decentralized case. Thus, comparing (3.1) and (3.2), we see that $R_C < R_D$. So, if $R_C \leq R < R_D$, we will say that there is *partial pooling* with centralization i.e. pooling only in asymmetric cost states, HL, LH, not in HH. The following Lemma summarizes the equilibrium configurations.

Lemma 1: *In the competence model, the equilibrium outcome has the following features: (a) with decentralization, there exists a separating equilibrium if $R < R_D$ and a pooling equilibrium if $R \geq R_D$;(b) with centralization, there exists a separating equilibrium if $R < R_C$; a partial pooling equilibrium if $R_C \leq R < R_D$; and a pooling equilibrium if $R \geq R_D$.*

[16] In what follows, state XY refers to regional costs distribution c_X , c_Y, with $X, Y = H, L$.

3.3. Welfare Comparisons

This is very simple. First, we note that because of the cost invariance assumption, the probability of any of the cost states HH, HL, LH, LL is the same under centralisation and decentralization. So, the expected payoff to a voter under both decentralization and centralization is the same, except in asymmetric costs states where also $R_C \leq R < R_D$. Then, decentralization gives voters a higher payoff, as they are unambiguously worse off under pooling (both incentive and separation effects of pooling are negative). So, we can summarise as follows.

Proposition 1. *The outcome, and voter welfare with centralization and decentralization is the same, except if (i) the cost is high in one region and low in another, and (ii) $R_C \leq R < R_D$. In this event, the bad incumbent will separate under decentralization but will pool under centralization, making all voters better off under decentralization.*

The intuition underlying this result is the following. With centralization, politicians are held accountable to a larger extent, because to get re-elected they must produce a good record (the same record as a competent politician) in *both* regions. With decentralization only a good record at home suffices to gain re-election. This gives extra incentive to pool with centralization. However pooling is bad in two respects. First, it induces inefficient provision of public good (worse incentive). Second it reduces the chance of detecting and replacing bad politicians (worse selection). It then follows that centralization is unambiguously worse than decentralization.

Note that it is reelection that is the source of inefficiency in this model. If politicians were not subject to reelection, they would not distort public good provision in the first place. It follows that accountability is less desirable the greater the distortion in current incentive relative to the selection gain from reelection.[17]

3.4. What is the Key Feature of Centralization?

As remarked in the introduction, there are two key changes in the move to centralization: one policy-maker replaces two, and one common budget constraint (with a single tax) replaces two. Which of these generates the difference in outcome described in Proposition 1?

[17] For an interesting analysis of the tradeoff between accountable and non-accountable officials, see Maskin and Tirole (2001).

Assume for the moment that fiscal policy in one region is not observable by voters in the other region, and suppose that with centralization, the politician can now levy separate taxes on the two regions (non-uniform taxation). Given the additional degrees of freedom in setting tax, the incumbent has the new option to separate selectively in only one region. If the state is HL, he can provide public good levels (g_H, g_L) and set separate taxes $(\tau_H,\ \tau_L)$, exactly as under decentralization. This choice implies that all voters in the low-cost region will vote for the incumbent (he pools in this region with the good type), whereas in the high-cost region, he reveals himself to be incompetent, and so all will vote for the challenger. As both regions have equal population, with neutral electoral rule, the incumbent and challenger are elected with equal probability. So, in this case, his payoff is

$$\widetilde{W}_C^{sep} = W_H + W_L + \frac{1}{2}\left[2\delta(EW + R) + 2\delta(\pi W_L + (1-\pi)EW)\right]$$

where the last two term in the square brackets reflects the fact that the incumbent is re-elected or not with equal probability, enjoying payoffs $2(EW+R)$ or $2(\pi W_L+(1-\pi)EW)$ in the following period respectively.

If R is large enough that the incumbent wishes to regain office,[18] as assumed above, then $\widetilde{W}_C^{sep} \geq W_C^{sep}$. That is the incumbent prefers to separate selectively in one region preserving his chance of re-election in the other region, and separate taxes enables him to do so. Thus, with non-uniform taxes, the incumbent will pool iff $W_C^{pool} \geq \widetilde{W}_C^{sep}$. After some rearrangement, we see that the condition for pooling is simply iff (3.1) holds, which is the same as in the case of decentralization. So, with separate taxes, decentralization and centralization are equivalent. The intuition is that uniform taxation makes it more costly to separate in one region, because then the incumbent's type is revealed to voters in *both* regions.

Of course, all this depends on the assumption that fiscal policy in one region is not observable by voters in the other region. If it is observable, then the incumbent cannot separate selectively in only one region, and then the welfare difference between centralization and decentralization will arise again with exactly the same outcome as for uniform taxes.

4. Benevolence

A key feature of the competence model is that it is the re-election motive that is the source of inefficient public good provision. As a result, both the incentive and selection

[18] Namely, if $EW + R \geq \pi W_L + (1-\pi)EW$, or $R \geq q\pi(W_L - W_H)$.

effects are associated with pooling lower voter welfare. As we now see, when politicians differ in benevolence, the incentive effect associated with pooling raises voter welfare - in order to get re-elected, the bad incumbent must control his predilection for diverting rent This conflict makes for richer predictions - in particular, centralization can dominate decentralization from the point of view of the voter.

4.1. Decentralization

Again, we solve backward. In the second period, the honest policy-maker will provide optimal public good level g_k given the cost realization c_k, and set tax $\tau_k = c_k g_k$. The dishonest policy-maker will just thus take maximum rent and set a tax of $\tau = \overline{r}$. So, all voters prefer the honest policy-maker.

In the first-period, a good type incumbent in either region behaves non-strategically and so will make exactly the same policy choices as in the second period.[19] Moreover, when cost is high, the bad type always prefers to take maximum rent in the first period, rather than imitates the good type in exchange for re-election: this is because discounting the future, it is better to take maximum rent now, and nothing later, rather than the opposite. This leaves the bad type in either region with three options when cost is low. First, he can set (g_H, τ_H) and take $\hat{r} = g_H(c_H - c_L)$ in the form of rents: call this the pooling strategy. Second, he can set $g = 0$, and take maximal rents, by setting $\tau = \overline{r}$:call this the separating strategy. Finally, he can randomise between these two alternatives, choosing the first with probability λ: call this the hybrid strategy.

His choice among these three strategies will depend on how each affects his probability of re-election. The key issue is how voters will vote if they observe (g_H, τ_H). A voter's posterior belief that the incumbent is benevolent, conditional on this tax rate is

$$\Pr(b \,|\tau_H, g_H) = \frac{\pi q}{\pi q + (1-\pi)(1-q)\lambda} \tag{4.1}$$

Let σ be the probability of re-election when (g_H, τ_H). Then if $\Pr(b \,|\tau_H, g_H) > \pi$, $\sigma = 1$, if $\Pr(b \,|\tau_H, g_H) < \pi$, $\sigma = 0$, and if $\Pr(b \,|\tau_H, g_H) = \pi$, voters are prepared to randomise, so $\sigma \in [0, 1]$.

Based on these strategies, there are then three possible types of equilibrium[20]: First, a *pooling* equilibrium, where the bad politician imitates the good one when the cost of

[19] Here, as in the remainder of the paper, we must assume a non-strategic behavior of the "good" incumbent (in the first period) or else,as Christian Schultz pointed out, there would be reasonable beliefs out-of-equilibrium inducing the good incumbent to distort public good supply in equilibrium to gain reelection in place of bad challenger. This assumption is implicit in Besley-Smart (2003).

[20] We ignore non-generic cases where behaviour can be different in the two different regions.

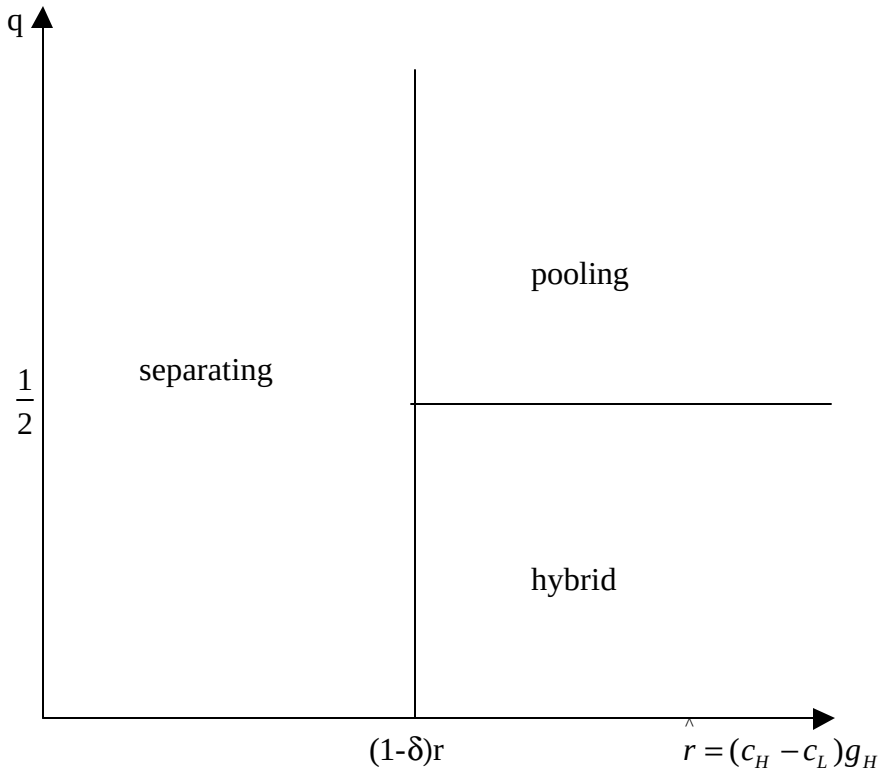

Figure 4.1:

public good provision is low, and is re-elected with probability 1 in that event. Second, a *separating* equilibrium where bad politician does not imitate the good one even when the cost of public good provision is low, and is re-elected with probability 0 in that event. Finally, a *hybrid* equilibrium, where he pools with probability $0 < \lambda < 1$ and is re-elected with probability $0 < \sigma < 1$. We then have the following result, which is simply a restatement of Lemma 1 of Besley and Smart(2003):

Proposition 2. *With decentralization, a pooling equilibrium exists iff* $q \geq 0.5$ *and* $\hat{r} \geq (1-\delta)\overline{r}$, *a separating equilibrium exists iff* $\hat{r} \leq (1-\delta)\overline{r}$, *and a hybrid equilibrium, with* $\lambda = q/(1-q) = \lambda_D$ *and* $\sigma = (\overline{r} - \hat{r})/\delta\overline{r}$ *exists iff* $q < 0.5$ *and* $\hat{r} \geq (1-\delta)\overline{r}$.

This result is illustrated in Figure 1 below:

4.2. Centralization

Solving backward, in the second-period, the benevolent policy-maker will provide optimal public good level in each region given local costs and charge uniform tax equal to average

cost. The non-benevolent policy-maker will provide no public and take maximum rent, regardless of the cost configurations. So, all voters prefer the benevolent policy-maker.

In the first-period, the benevolent incumbent behaves non-strategically and so will make exactly the same policy choices as in the second period. As for the non-benevolent strategy, when cost is high in both regions HH the same argument as in the decentralized case implies that the non-benevolent type always prefers to take maximum rent in the first period, rather than imitate the benevolent type in exchange for re-election. This leaves the following cost configurations with possible strategies (taking it as given that randomization between these strategies is also an option).

- *In asymmetric cost state* HL or LH: the bad incumbent can (i) pool by setting (τ_H, g_H) thus taking rent $\hat{r}$; (ii) separate by choosing maximal rents and no public good provision.
- *In symmetric cost state* LL : the bad incumbent could either:(i) pool by setting (τ_H, g_H) thus taking rent $2\hat{r}$; (ii) pool by setting g_H and g_L and $\tau = (\tau_H + \tau_L)/2$, thus taking rent $\hat{r}$, or (iii) separate by choosing maximal rents and no public good provision.

However, in event LL, option (ii) with partial rent diversion, can be shown never to be optimal for the incumbent, and so we do not consider it further.[21] Let the probability of pooling in asymmetric state HL (or LH) be ξ, and the probability of pooling in symmetric state LL be λ. Then, given ξ, λ, voters' posterior belief that the incumbent is benevolent, conditional on the tax rate and public good supply must be:

$$\begin{aligned}
\Pr(b\,|\tau_L, g_L) &= \Pr(b\,\Big|\frac{\tau_L + \tau_H}{2}) = 1 \\
\Pr(b\,|\tau_H, g_H) &= \frac{q^2\pi}{q^2\pi + (1-q)^2\lambda(1-\pi) + 2q(1-q)\xi(1-\pi)} \\
\Pr(b\,|\tau, g) &= 0 \text{ otherwise}
\end{aligned}$$

The best response by voters is to re-elect only when this probability exceeds π.

[21] Indeed if the incumbent pools when HL or LH we must have $P(b|\tau_H, g_H) \geq \pi$ but then deviating when LL to take rent $2\hat{r}$ preserves re-election and because $\hat{r} \geq (1-\delta)\overline{r}$ the bad type would prefer deviating. Alternatively, if the incumbent does not pool when HL or LH we must have $P(b\,|\tau_H, g_H) = 1$ which again implies that bad type would prefer deviating when LL to take rent $2\hat{r}$.

First, we look for an equilibrium with the maximum degree of pooling i.e. $\lambda = \xi = 1$. We call this equilibrium a *full pooling equilibrium.* Then

$$\Pr(b\,|\tau_H, g_H) = \frac{q^2\pi}{q^2\pi + (1-q^2)(1-\pi)}$$

This can be an equilibrium only if $\Pr(b\,|\tau_H, g_H) \geq \pi$ (otherwise, the politician would prefer to take maximum rents, anticipating not being re-elected), which requires $q \geq 1/\sqrt{2}$. Note that the condition for full pooling here is stronger than the condition for pooling in the decentralized case (i.e. $q \geq 1/2$).[22] Also, it requires that whatever the cost realizations, the politician prefers to take the rent from pooling rather than go for maximum rent extraction. The rent from pooling is either $\hat{r}$ or $2\hat{r}$, so $\hat{r} + \delta 2\overline{r} \geq 2\overline{r}$, or $\hat{r} \geq 2(1-\delta)\overline{r}$ is required. Comparing to Proposition 2, we see that a full pooling equilibrium is less likely than pooling with decentralization (the conditions on q and $\hat{r}$ are tighter).

The alternative for the incumbent is to pool only when the possible rents from doing so are large i.e. $2\hat{r}$. We call this a *partial pooling equilibrium.* This requires $\lambda = 1,\ \xi = 0$, in which case

$$\Pr(b\,|\tau_H, g_H) = \frac{q^2\pi}{q^2\pi + (1-q)^2(1-\pi)}$$

This can be an equilibrium only if $P(b\,|\tau_H, g_H) \geq \pi$, which requires $q \geq 1/2$. Note that in this case, the condition on q is the same as in the decentralized case. The equilibrium additionally requires that the rent from pooling $2\hat{r}$, satisfies $2\hat{r} + \delta 2\overline{r} \geq 2\overline{r}$, or $\hat{r} \geq (1-\delta)\overline{r}$.

Next, for a *separating equilibrium* to exist, it is necessary and sufficient that be the case that whatever the cost realizations, the bad politician prefers to take maximum rents $2\overline{r}$ now rather than pool. This occurs when $2\overline{r} > 2\hat{r} + \delta 2\overline{r}$, or $\hat{r} < (1-\delta)\overline{r}$.

Finally, we consider *the hybrid equilibrium.* This occurs iff $q < 1/2$ and $\hat{r} \geq (1-\delta)\overline{r}$. From the second condition, the bad politician always wants to pool: the problem is that if he does so fully ($\lambda = \xi = 1$) or even partially ($\lambda = 1, \xi = 0$), the voter's posterior belief that he is good will be too low for him to be re-elected. So, we need λ, ξ such that

$$\Pr(b\,|\tau_H, g_H) = \frac{q^2\pi}{q^2\pi + (1-q)^2\lambda(1-\pi) + 2q(1-q)\xi(1-\pi)} = \pi$$

which requires

$$q^2 = (1-q)^2\lambda + 2q(1-q)\xi$$

[22] The reason for this is that with centralization, tax τ_H will only be set by the good incumbent if the cost is high in *both* regions, and the probability that this occurs is $q^2 < q$: consequently, q must be higher to compensate for the lower probability of this event.

Also, the bad politician gets a higher payoff from pooling in state LL than in state HL or LH : so, will set $\lambda = 1$ if possible. But setting $\lambda = 1$ in the above expression, we get $2q - 1 = 2q(1-q)\xi$ which is impossible given $q < 1/2$. So, we have $\xi = 0$ (no pooling in state HL or LH) and

$$\lambda = \frac{q^2}{(1-q)^2} = \lambda_C$$

Then, the bad politician must be indifferent between rent extraction and pooling in state LL;this requires $2\hat{r} + \sigma\delta 2\overline{r} = 2\overline{r}$ which requires

$$\sigma = \frac{\overline{r} - \hat{r}}{\delta\overline{r}}$$

So, in the hybrid equilibrium, the probability of re-election is as in BS with decentralization: but, the probability of pooling is lower with centralization: $\left(\frac{q}{1-q}\right)^2 < \frac{q}{1-q}$ as $q < 0.5$. We can summarise the discussion as follows:

Proposition 3. *With centralization, a full pooling equilibrium exists iff* $q \geq 1/\sqrt{2}$ *and* $\hat{r} \geq 2(1-\delta)\overline{r}$, *a partial pooling equilibrium exists iff* $q \geq 1/2$ *and* $\hat{r} \geq (1-\delta)\overline{r}$, *a separating equilibrium exists iff* $\hat{r} \leq (1-\delta)\overline{r}$, *and a hybrid equilibrium, with* $\lambda = q^2/(1-q)^2 = \lambda_C$ *and* $\sigma = (\overline{r} - \hat{r})/\delta\overline{r}$ *exists iff* $q < 0.5$ *and* $\hat{r} \geq (1-\delta)\overline{r}$.

Comparing equilibria with centralization and decentralization, we can make the following observations. The conditions for separating and hybrid equilibria are the same. The condition for pooling equilibrium is also the same but centralization involves two different degrees of pooling: full pooling when both the rent from pooling and the probability of high cost are high enough; otherwise pooling is partial. The equilibrium configuration under centralization is illustrated below.

4.3. Welfare Comparisons

Comparing Propositions 2 and 3, we can derive quite complete results on the effect of centralization on voter welfare. We measure voter welfare as the sum of voter utilities in the two regions (no regional welfare distribution concern). We proceed by making the comparison separately for different regions of the parameter space, in particular, the set of feasible values of $q, \hat{r}$, as suggested by Figures 1-2.

Case 1. $\hat{r} < (1-\delta)\overline{r}$. Here, only a separating equilibrium exists with either centralization and decentralization. Then the outcome is identical with decentralization and centralization, so voter welfare is identical.

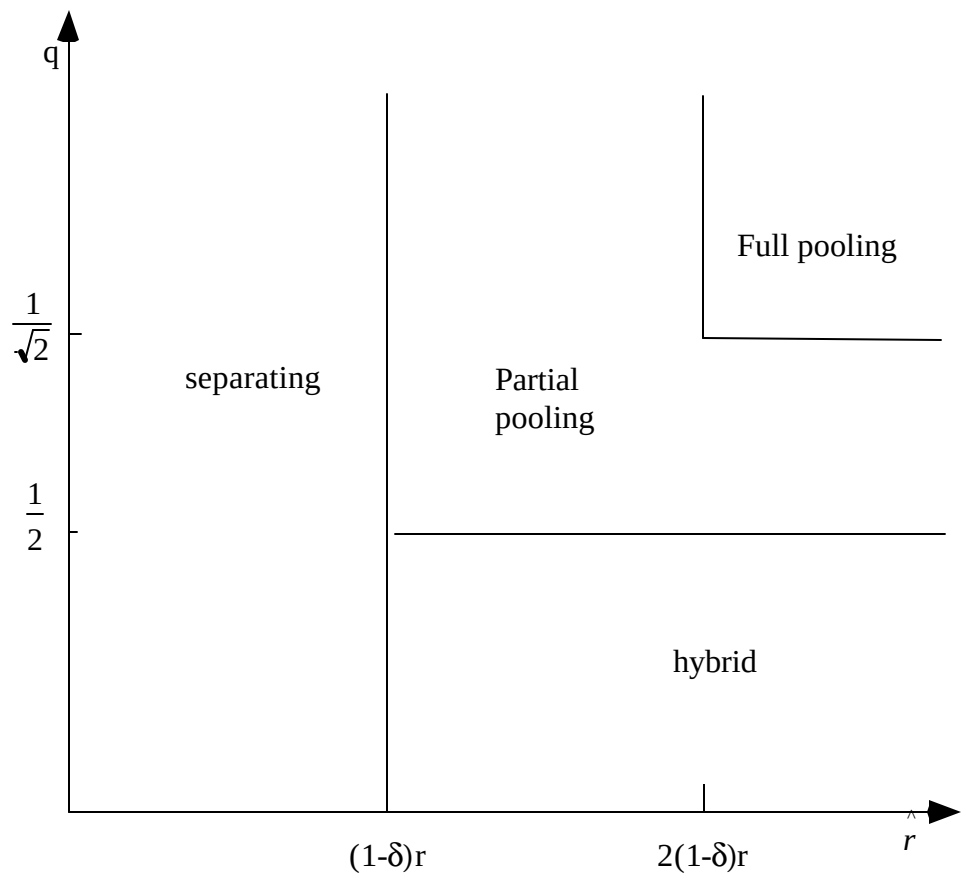
q
Full pooling
1/√2
separating
Partial pooling
1/2
hybrid
(1-δ)r
2(1-δ)r
r̂

Figure 4.2:

Case 2. $\hat{r} \geq 2(1-\delta)\overline{r}$, $q \geq 1/\surd 2$. In this case, when the cost is the same in both regions -states (H, H) or (L, L), the two types of incumbent act in the same way under centralization and decentralization and so the outcomes are the same. On the other hand, if the state is HL or LH, with decentralization the bad incumbent in *only* the low-cost region pools, whereas with centralization, the (single) incumbent *always* pools. So, the incentive effect is *stronger* with centralization, but the selection effect is *weaker*.

Case 3. $(1-\delta)\overline{r} \leq \hat{r} \leq 2(1-\delta)\overline{r}$, or $0.5 \leq q \leq 1/\surd 2$, or both. In this case, when the cost is the same in both regions -states HH or $LL-$ the two types of incumbent act in the same way under centralization and decentralization and so the outcomes are the same. On the other hand, if the state is HL or LH, with decentralization the bad incumbent in *only* the low-cost region pools, whereas with centralization, the (single) incumbent *never* pools. So, the incentive effect is *weaker* with centralization, but the selection effect is *stronger*.

Case 4: $\hat{r} \geq (1-\delta)\overline{r}$, $q < 0.5$. In this case, when the cost is high in both regions -state $HH-$ the two types of incumbent act in the same way under centralization and decentralization i.e. bad incumbents always take maximum rents and so the outcomes are the same. On the other hand, if the state is HL or LH, with decentralization the bad incumbent in *only* the low-cost region pools with probability λ_D, whereas with centralization, the (single) incumbent *never* pools (since with hybrid equilibrium we have $\xi = 0$). So, in these two asymmetric states, the incentive effect is *weaker* with centralization, but the selection effect is *stronger*. Finally, if the state is LL, with decentralization, both bad incumbents pool with probability λ_D^2, both separate with probability $(1-\lambda_D)^2$, and one separates with probability $2\lambda_D(1-\lambda_D)$. With centralization, the single incumbent pools with probability λ_C which is less than λ_D. In this case, the change in the incentive and selection effects with centralization goes again in the same direction: worse incentive (due to less expected pooling) but better selection.

So, we have already established that incentive and selection effects may be either stronger *or* weaker with centralization. Of course, merely identifying how the incentive and selection effects differ following a switch from decentralization to centralization (or vice versa) is just the first step. The welfare gains associated with increased incentive and selection effects have to be evaluated and compared. This is straightforward but involves some algebra. The results are summarised in the following Proposition, proved in the Appendix:

Proposition 4. *In case of separating equilibrium, decentralization is equivalent to centralization. Otherwise decentralization dominates centralization if and only if* $\delta\pi \geq \frac{W_H+\overline{r}}{EW+\overline{r}}$

in case of full pooling equilibrium; and if and only if $\delta\pi \leq \frac{W_H+\overline{r}}{EW+\overline{r}}$ *in case of partial pooling or hybrid equilibrium.*

The intuition is simple. In the full pooling case, for example, the selection effect is *stronger* with decentralization (i.e. it is more likely that the bad incumbent is voted out). The expected discounted benefit of replacing the bad incumbent by the challenger is $\delta[\pi EW + (1-\pi)(-\overline{r}) - (-\overline{r})] = \delta\pi(EW + \overline{r})$. However, this gain is made at the cost of removing the incentive to pool of the incumbent in the first period, which costs $W_H+\overline{r}$. So, the overall gain to decentralization is $\delta\pi(EW+\overline{r})-(W_H+\overline{r})$ which is positive when $\delta\pi \geq \frac{W_H+\overline{r}}{EW+\overline{r}}$. The case with partial pooling is just the other way around with more pooling in decentralization and thus less selection. Interestingly the case with hybrid equilibrium turns out to be the same as the partial pooling case. To see why this is, note that in spite of the apparent difference in the incumbent behavior, decentralization involves in expected terms more pooling and less selection. In particular, in both hybrid and partial pooling there is no pooling with centralization when cost state is asymmetric wherever pooling arises in the low cost region with decentralization.

The basic lesson of proposition 5 is that decentralization need not improve accountability. Actually it suggests rather that centralization is equally likely to dominate.

4.4. What is the Key Feature of Centralization?

As the case of varying competence, when centralization occurs, there are two changes: (i) there is a single, rather than two, policy-makers: (ii) there is a single, rather than two, budget constraints. Which of these changes generates the difference between the two cases?

Again, consider a situation where the single policy-maker sets a different tax in each of the two regions. Further assume as we have done so far that voters only observe tax and public good provision in their own region. Then in an asymmetric cost state HL or LH, the non-benevolent incumbent will always separate in the high-cost region, and he has two options in the low cost region. The first is to separate by setting $g = 0, \tau = 1$.The second is to pool by setting $g = g_H,\ \tau = \tau_H$. Pooling implies that all voters in the low-cost region will vote for the incumbent, whereas in the high-cost region, all will vote for the challenger. As both regions are assumed to have equal population, the outcome is that by separating selectively in only one region, the incumbent is re-elected with probability 0.5, whereas with complete separation, the incumbent is re-elected with probability 0.

A pooling equilibrium requires that $\Pr(b\,|\tau_H, g_H) \geq \pi$, which now requires that $q \geq 0.5$ as in the decentralized case. Moreover, the incumbent has payoff from selective separation

of $\hat{r} + \overline{r} + 0.5\delta(2\overline{r})$, whereas from complete separating, he has payoff $2\overline{r} + 0$. So, he will pool if $\hat{r} \geq (1-\delta)\overline{r}$, as in the decentralized case. This argument establishes that without fungibility, and with voters in one region unable to observe the fiscal policy in the other region, centralization and decentralization are equivalent.

5. Correlated Costs and Yardstick Competition

We now introduce the possibility of correlated cost shocks. We first consider how this affects the outcome with centralization in Section 5.1 below. Whether correlation affects the outcome with decentralization turns on whether we suppose that voters in region i can observe the fiscal policy choices of region j. If they can, then yardstick competition is possible, and as emphasized by Besley and Smart(2003). This may work in favour of decentralization (but does not necessarily do so). So, we then analyse the case of decentralization with yardstick competition in Section 5.2 below: this analysis extends the results of Besley and Smart(2003) on yardstick competition[23]. Finally, Section 5.3 compares centralization and decentralization with correlated costs.

5.1. Centralization with Correlated Cost Shocks

We model correlation as follows. The joint probability distribution over possible states of the world is:

$$\begin{aligned} \Pr(HH) &= q^2 + \rho q(1-q),\ \Pr(LL) = (1-q)^2 + \rho q(1-q), \\ \Pr(HL) &= \Pr(LH) = q(1-q) - \rho q(1-q) \end{aligned} \tag{5.1}$$

where $\rho \in [0,1]$ is the correlation coefficient. So, with perfectly correlated cost shocks, $\Pr(HH) = q$, $\Pr(LL) = 1-q$ and $\Pr(HL) = \Pr(LH) = 0$.

With decentralization, if voters cannot observe policy choice in the other region, then the outcome is the same as with independent shocks i.e. Proposition 2 still applies. With centralization, however, the outcome is affected by correlation, even when voters have limited information. The possible types of equilibrium are the same (full pooling, partial pooling, separating, hybrid), but the condition determining whether a full or partial pooling equilibrium occurs does depend on ρ, as does the probability of pooling in the

[23] Specifically, their analysis is restricted to the case where $q = 0.5$, whereas we characterise equilibrium with yardstick competition for all q.

hybrid equilibrium. Formally, our characterization of equilibrium with correlation is now as follows (see Appendix for the proof):

Proposition 5. *With centralization and correlated shocks, a full pooling equilibrium exists iff (i)* $\hat{r} \geq 2(1-\delta)\overline{r}$ *and (ii)* $q \geq q(\rho)$, *where* $q(\rho)$ *is strictly decreasing in* ρ *with* $q(0) = 1/\sqrt{2}$, $q(1) = 1/2$. *The conditions for partial pooling, separating and hybrid equilibria are the same as without correlation. However, the probability of pooling in the hybrid equilibrium is higher with correlation* $\lambda(\rho) = \frac{q^2+\rho q(1-q)}{(1-q)^2+\rho q(1-q)} > q^2/(1-q)^2$, *but the probability of re-election keeps the same* ρ.

So, with correlation, full pooling is easier to achieve in equilibrium than with independent cost shocks.

The intuition is simple. When $q > 1/2$, the higher the correlation ρ, the more likely is the state HH than is any one of the other three states. But in state HH the record (τ_H, g_H) will only be produced by a benevolent type, whereas in any one of the other three states, it will be set by a dishonest type (with full pooling). So, other things equal, the higher ρ, the higher $\Pr(b\,|\tau_H, g_H)$, and consequently q can be lower while satisfying the re-election condition $\Pr(b\,|\tau_H, g_H) \geq \pi$.

Note finally that the welfare result, Proposition 4, applies as stated to the case of correlated costs. Indeed, the proof in the Appendix holds for any correlation rate $\rho \in (0,1)$. The only difference is that in the statement of cases 2 and 3, the critical value of q, $1/\sqrt{2}$, is replaced by $q(\rho)$,and in the statement of case 4 the probability of pooling in hybrid equilibrium λ_C is replaced by $\lambda(\rho)$.

5.2. Decentralization with Yardstick Competition

Note that if relative performance is not observable by voters in the case of decentralization, Proposition 5 still applies to the comparison of voter welfares under decentralization and centralization. However, of course with correlation, the probability of asymmetric cost states, and thus the difference in expected voter welfare between the two regimes, is smaller than without correlation.

However, it is equally plausible that performance comparison is possible with decentralization, and indeed, it is often stated in the policy literature that the possibility of performance comparison is one of the key advantages of decentralization.

However we rule out strategic interaction between regions by assuming that each incumbent does not observe, or cannot condition policy on, the shock in the other region and the type of the other incumbent. We also assume that $\widehat{r} \geq (1-\delta)\overline{r}$ (i.e. no separating

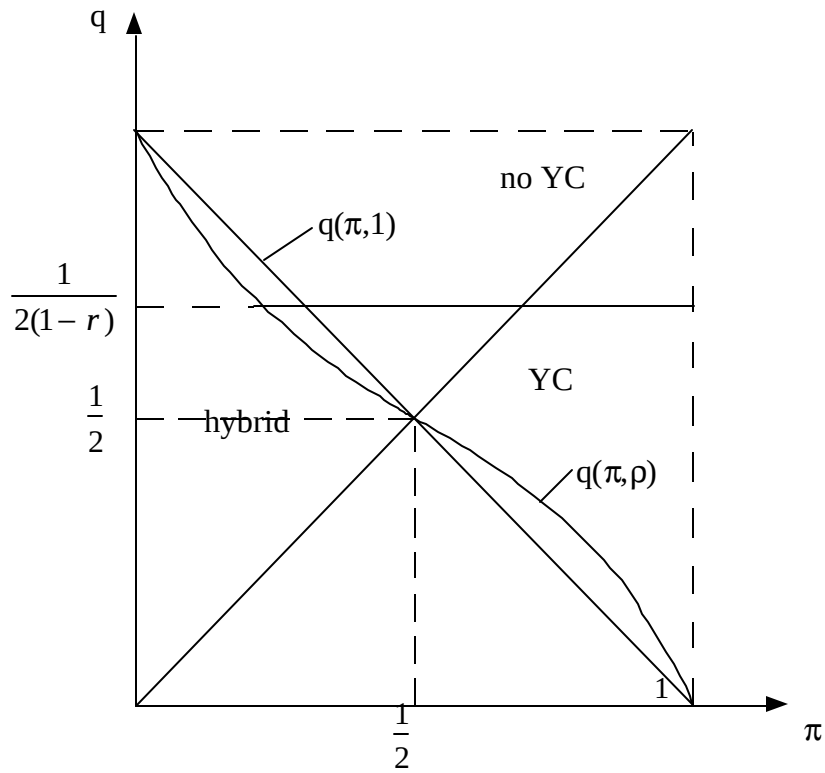

Figure 5.1:

equilibrium). We then have the following result. Let

$$\widehat{q}(\rho) = \frac{1}{2(1-\rho)}, \;\; \pi(q,\rho) = \frac{(0.5-q)}{(1-2q)^2 + 4\rho q(1-q)}$$

and let $\tilde{q}(\pi,\rho) = \pi^{-1}(q,\rho)$. Then we have the following Proposition, proved in the Appendix:

Proposition 6. *Assume* $\hat{r} \geq (1-\delta)\overline{r}$. *Then, a pooling equilibrium with no yardstick competition exists iff* $q \geq \tilde{q}(\pi,\rho)$ *and* $q \geq \widehat{q}(\rho)$, *a pooling equilibrium with yardstick competition exists iff* $\widehat{q}(\rho) \geq q \geq \tilde{q}(\pi,\rho)$ *and a hybrid equilibrium with yardstick competition exists iff* $q \leq \widehat{q}(\rho)$ *and* $q \leq \tilde{q}(\pi,\rho)$.

This equilibrium configuration is illustrated in the figure below

Note that Proposition 6 generalizes Lemma 2 of Besley and Smart(2003). To see this, note that $\widehat{q}(\rho) \geq 0.5$, so if $q = 0.5$, yardstick competition is always used in equilibrium, as in their Lemma 2, and a pooling equilibrium exists if $\pi \geq 1/2$,and a hybrid equilibrium exists if $\pi < 1/2$.

The welfare comparison is the same as before when yardstick competition is not binding. So we shall compare welfare when the yardstick effect is present. We first restrict attention to pooling with yardstick competition.[24]

5.3. Welfare Comparisons

We perform welfare comparison for equilibrium preserving yardstick competition. If yardstick comparison does not affect the equilibrium behavior of incumbents in period 1, it will however modify his probability of reelection in cost state LL. This is because the bad incumbent pooling on τ_H is no longer re-elected if the other incumbent is good, $\sigma(\tau_H, \tau_L) = 0$. In this event, the selection gain from voting out bad type is $W^0 + \overline{r}$ which occurs with probability $2\pi(1-\pi)$. For other cost states HH and HL (or LH) probabilities of reelection are the same: with cost H the bad incumbent takes maximum rent and is voted out irrespective of the outcome in other region; with cost L the bad incumbent is re-elected when pooling on τ_H and is reelected independently of the outcome in the other region.[25] So as before, when comparing centralization and decentralization, there is no welfare difference in cost state HH, and the welfare difference in either cost state HL or LH keeps the same (say,$EW_D - EW_C$). Using these facts, the expected welfare difference between decentralization and centralization is $2\pi(1-\pi)(W^0 + \overline{r})$ in cost state LL with probability $(1-q)^2 + \rho q(1-q)$;and is $EW_D - EW_C$ in cost states HL (or LH) with probability $2(q(1-q) - \rho q(1-q))$. Let EW_{YC} be the expected welfare with yardstick competition, then welfare difference with centralization is

$$\begin{aligned} EW_{YC} - EW_C &= [2q(1-q)(1-\rho)]\,(EW_D - EW_C) + \\ &\quad [(1-q)(1-q(1-\rho))]\,2\delta\pi(1-\pi)(W^0 + \overline{r}) \end{aligned}$$

where the first term (possibly negative) is the original welfare difference (without yardstick competition) and the second term (which is always positive) is the selection gain

[24]Note that in this case, the bad incumbent is less likely to be re-elected when pooling, which changes the threshold for no separating equilibrium to $\widehat{r} \geq (1 - \delta(1 - \pi + \pi(1-\rho)q))\overline{r}$ (which is strictly greater than $(1-\delta)\overline{r}$ and increasing with ρ). In the following we suppose this condition holds.

[25]This is true also if the outcome is $\overline{r}$ in the other region because then voters know that cost in that region is H and with positive correlation $\Pr(g\,|\tau_H, \overline{r}\,) > \Pr(g\,|\tau_H, \tau_H\,) > \pi$.

from yardstick competition. As correlation ρ increases, more weight is put on the second term relative to the first, making decentralization with yardstick competition more likely to dominate centralization. Thus we have:

Proposition 7. *In Case of pooling equilibrium with yardstick competition, the expected welfare with decentralization increases with correlation and there exists correlation rate $\rho^* < 1$ such that decentralization dominates if $\rho > \rho^*$ and centralization dominates if $\rho < \rho^*$.*

It remains to see the effect of yardstick comparison on hybrid equilibrium. When $q \leq \widehat{q}(\rho)$, the incumbent faces the re-election rule with yardstick competition, $\sigma(\tau_H, \tau_L) = 0$. To preserve chances of re-election, the incumbent can use mixed strategy. This requires $\Pr(b\,|\tau_H, \tau_L) = \pi$ which implies

$$q(1-q)(1-\rho) = (1-q)(1-(1-\rho)q)\lambda$$

which reduces to

$$\lambda = \frac{q(1-\rho)}{1-(1-\rho)q} \equiv \lambda_{YC}$$

Thus the probability of pooling decreases with correlation. Our next proposition establishes that in case of hybrid equilibrium, yardstick competition makes decentralization less likely to dominate:

Proposition 8. *In Case of hybrid equilibrium with yardstick competition, decentralization dominates if and only if $\delta\pi \leq \gamma(\rho)\left(\frac{W_H+\bar{\tau}}{EW+\bar{\tau}}\right)$ with $\gamma(\rho) < 1$, $\gamma(\rho)$ decreasing and $\gamma(\rho) < 0$ for $\rho > \rho^*$.*

Therefore the previous conclusion (of proposition 7) that yardstick competition offers decentralization the best chance of dominating centralization is no longer true with hybrid equilibrium. The reason is that to preserve chance of re-election bad incumbents optimally respond to yardstick comparison by pooling less and thereby reduces the relative advantage of decentralization over centralization (viz. more pooling with less selection).

6. Related Literature and Conclusions

6.1. Related Literature

There are a number of literatures related to the topic of this paper. First, as mentioned, there are two papers, namely Seabright (1996) and Bardhan and Mookherjee(2000) which

directly address the relationship between fiscal decentralization and accountability. However, both papers have their limitations in the modelling of the key concept of accountability. Seabright's paper posits a number of regions, and that the welfare of all the residents of any region is determined by policy choice and a region-specific taste shock, which is uncorrelated across regions. The electoral rule is majoritarian: with centralization, the national policy-maker must capture the votes of at least a majority of regions,[26] and with decentralization, the regional policy-maker must capture at least a majority of districts within regions.[27] In this setting, a national policy-maker is less accountable than a set of regional policy-makers, as to be re-elected, the former only needs to choose policy in such a way that a *majority* of regions prefer the incumbent, whereas the latter needs to choose policy in such a way that *all* the residents within a region prefer the incumbent. Therefore his analysis relies on heterogenous tastes across regions and homogenous taste within region, and the dominance of decentralization emerges naturally as a preference matching device.

Bardhan and Mookherjee (2000) focus on lobbying by an interest group in a setting of Downsian political competition between two parties (that is, parties can precommit to policy positions ex ante) The focus is on the amount of lobbying and the consequent distortion in the policy outcome away from the preference of the median voter under centralization and decentralization. One limitation of the model is that it is not completely micro-founded, because campaign contributions are simply assumed to affect the voting decisions of uninformed voters. As the work of Prat (2002a,b) shows, the informational role of campaign contributions can be explicitly modelled. Another problem is that in their framework,it is difficult to make explicit welfare comparisons between centralization and decentralization, as voters are heterogenous.

Second, this paper builds on the large literature on electoral accountability that can be traced back to the work of Barro (1973) and Ferejohn (1986). The closest to our work is the important recent paper by Besley and Smart (2003): indeed, we use an extension of their model as a vehicle for our analysis. We should stress that in their paper, they do not consider centralized decision-making: their benchmark is decentralization without "competition" between regions, and then the impact on selection and incentive effects. of introducing either tax or yardstick competition is studied[28]. A further related contribution

[26] In fact, Seabright allows for a more general electoral rule where some arbirary fraction of electoral districts must be won in order to win the election overall.

[27] Given that the shock to welfare is identical for all residents within a region, either all voters will vote for the incumbent, or none will. So, the precise form of the electoral rule does not matter here.

[28] See also Carillo. and Mariotti (2001).

is Belleflamme and Hindriks (2003) who assess the effect of yardstick competition in a different context where bad incumbents distort decision to get re-elected (they use inefficient sneaky methods of redistribution instead of more transparent efficient methods). One consequence of this approach is that yardstick competition can never hurt the voters, because bad incumbents are more easily detected (selection gain) and so they are less likely to distort inefficiently their policy choice for re-election purpose (incentive gain).

There is also a set of political models of corruption assessing the relative performance of different electoral rules (majoritarian rule versus proportional representation). In Myerson (1993) voting behavior is endogenous to the electoral rule but corruption is assumed to be an exogenous feature of each politician. The ability of voters to hold corrupt incumbents accountable is worse with the majoritarian rule as voters are less willing to switch their vote on a "good" challenger who is perceived to have little chance of winning. Persson and Tabellini (1999) compare political bevior under proportional representation in a nation-wide district and the majority rule in a number of local districts. Competition is stronger with local majority rule as politicians are interested to win a majority, not national-wide, but in the marginal districts containing more swing voters. As these voters are more responsive to policy outcome, politicians are more disciplined and divert less rents.

Finally, there is a large body of literature on decentralization within organizations. The basic trade-off is between delegation of decision-making to better informed agents and the associated loss of control and agency costs (since informed agents may not have all the incentives to act in the best interests of the delegating authority). The optimal organizational form emerges from this simple trade off. For a long time it has been a challenge to explain decentralization with this framework. The reason is that from the "revelation principle" any decentralized organization can be replicated by a centralized one in which all agents with different sets of information report their information to a center who then makes all the relevant decision (Myerson(1982). This revelation principle has for a long time been the cornerstone of the weak superiority of centralization over decentralization. However recent theoretical developments have shown that relaxing some of the underlying assumptions can create a preference for decentralization. For example if the center cannot commit not to renegotiate the initial contract after agents have reported their information, then decentralization is an effective commitment device by preventing ex-post opportunistic renegotiations from the center (Beaudry and Poitevin(1995)). When agents can collude when reporting their information to the center, then decentralization by creating a conflict of interest between agents can prevent any collusion and again decentralization to informed agents may be optimal (Laffont and Martimort(1998)). When communication is costly, decentralization may also be optimal because it reduces commu-

nication costs (see Melumad, Mookherjee and Reichelstein(1997)).

However, the model of this paper is rather different to the contractual approach. First, models of electoral incentives (as is ours) take an incomplete contracting approach. First delegation of decision making to politicians. The "right" policy choice depends on non-verifiable information about the state of the environment, that cannot be contracted upon ex-ante. Consequently it must be delegated to politicians, thereby creating an incentive problem since giving any residual right to an agent to make collective choices opens up the possibility of opportunistic behavior and abuse of power. Second, the decision maker is selected by election: the politician is not controlled by a contract but by the public opinion. If he wants to remain in office, he must win elections.

6.2. Conclusions

In this paper, we have considered the impact of different assignments of fiscal powers (centralized and decentralized) on accountability of politicians. There are just two differences between centralization and decentralization. With decentralization there is a distinct decision maker in each jurisdiction and separate taxation. With centralization it is the same decision maker for both jurisdictions with uniform taxation and common budget constraint. Politicians are either "good" or "bad". Given homogenous electorate there is no ambiguity about what is a good incumbent. As for a bad incumbent we consider two possibilities: either the bad incumbent is incompetent or dishonest. For each case we study which organizational form performs better in terms of its ability to screen and discipline politicians. Our main results are that (i) in the competence case, decentralization unambiguously dominates, and (ii) in the honesty case, centralization and decentralization have conflicting effects on accountability, in the sense that one institutional form is better to discipline when the other is better to screen politicians. A distinctive feature of centralization in our analysis is to extend accountability to multiple constituencies with whom the politician may or may not be congruent subject to an overall budget constraint. In contrast, decentralization makes politicians only answerable to their local constituency.

References

[1] Alesina, A. and E. Spolare (1997), "On the Number and Size of Nations," *Quarterly Journal of Economics*, 112, 1027-56.

[2] Azfar, O., S. Kahkonen, and P. Meagher (2001) "Conditions for Effective Decentralised Governance: A Synthesis of Research Findings", IRIS Center, University of Maryland.

[3] Bardhan, P. and D. Mookerjeee (2000), "Capture and Governance at Local and National Levels," *American Economic Review*, 90(2), 135-139.

[4] Barro, R., (1973), "The control of politician: an economic model", *Public Choice* 14, 19-42.

[5] Belleflamme, P. and J. Hindriks (2003) "Yardstick competition and political agency problems", *Social Choice and Welfare* (forthcoming).

[6] Besley, T. and A. Case (1995), "Incumbent Behavior: Vote-seeking, Tax-Setting, and Yardstick Competition", *American Economic Review*, 85, 25-45.

[7] Besley, T. and S. Coate (2003), "Centralized versus decentralized provision of local public goods: a political economy approach", *Journal of Public Economics*, 87, 2611-37.

[8] Besley, T. and M. Smart (2003), "Does Tax Competition Raise Voter Welfare?".mimeo, London School of Economics.

[9] Besley, T. and R. Burgess (2000), "The Political Economy of Government Responsiveness: Theory and Evidence from India," *Quarterly Journal of Economics,* 117 (4), 1415-51.

[10] Besley, T. and A. Prat (2001), "Handcuffs for the Grabbing Hand? The Role of the Media in Political Accountability", mimeo, London School of Economics.

[11] Beaudry, P. and M. Poitevin (1995), "Contract Renegotiation: a Simple Framework and Implications for Organization Theory", *Canadian Journal of Economics*, 28, 302-35.

[12] Bolton, P. and G. Roland (1997), "The Break-Up of Nations: A Political Economy Analysis" *Quarterly Journal of Economics*, **112**, 1057-90.

[13] Carillo, J. and T. Mariotti (2001) "Electoral Competition and Politicians Turnover" *European Economic Review*, 45, 1-26

[14] Coate, S. and S. Morris (1995), "On the Form of Transfers to Special Interests", *Journal of Political Economy*, 103, 1210-35.

[15] Cremer. J. and T.Palfrey, (1996), "In or Out? Centralization by Majority Vote", *European Economic Review* **40**, 43-60.

[16] Drazen, A. (2000a), *Political Economy in Macroeconomics*, Princeton University Press, Princeton, NJ.

[17] Ellingsen, T.(1998), "Externalities vs. Internalities: A Model of Political Integration", *Journal of Public Economics*, **68**, 251-268.

[18] Ferejohn, J. (1986), Incumbent performance and electoral control, *Public Choice* 50, 5-25.

[19] Fissman, R., and R.Gatti (2001) "Decentralization and Corruption: Evidence Across Countries," *Journal of Public Economics,* 83(3), 325-45.

[20] Gilbert, G. and P.Picard (1996), "Incentives and the Optimal Size of Local Jurisdictions", *European Economic Review* **40**, 19-43.

[21] Harrington, J.E., Jr. (1993), "Economic Policy, Economic Performance, and Elections", *American Economic Review*, 83, 27-42.

[22] Hess, G. and A. Orphanides (1995), "War Politics: An Economic, Rational Voter Framework", *American Economic Review*, 85, 828-46.

[23] Huther, J. and A.Shah (1998) "Applying a Simple Measure of Good Governance to the Debate on Fiscal decentralization", World Bank Policy Research Working Paper 1894.

[24] Laffont, J.-J., and D. Martimort (1998), "Collusion and Delegation", *Rand Journal of Economics*, 29, 280-305.

[25] Laffont, J.-J. (2000), "*Incentives and Political Economy*", Oxford University Press: New York

[26] Lockwood, B., (2002), "Distributive Politics and the Costs of Centralization", *Review of Economic Studies*, 69(2), 313-337.

[27] Maskin E. and J. Tirole (2001), "The Politician and the Judge: Accountability in Government", mimeo Princeton University.

[28] Mello,L. and M.Barenstein (2001), "Fiscal Decentralization and Governance: A Cross-Country Approach", IMF Working Paper 01/71.

[29] Melumad, N., D. Mookherjee, and S. Reichelstein (1997), "Contract Complexity, Incentives, and the Value of Delegation", *Journal of Economics and Management Strategy*, 6, 257-89.

[30] Myerson, R. (1982), "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems" *Journal of Mathematical Economics* 10, 67-81.

[31] Myerson, R. (1993), "Effectiveness of Electoral Systems for Reducing Government Corruption: a Game Theoretic Analysis", *Games and Economic Behavior* 5, 118-32.

[32] Myerson, R. (1999), "Theoretical Comparisons of Electoral Systems" *European Economic Review* 43, 671-97.

[33] Oates, W. (1972), *Fiscal Federalism,* Harcourt-Brace, New York.

[34] Oates, W (1999), "An Essay on Fiscal Federalism," *Journal of Economic Literature,* 37(3), 1120-49.

[35] Oberholzer-Gee, Felix and Koleman S. Strumpf (2002): "Endogeneous Policy Decentralization: Testing the Central Tenet of Economic Federalism," *Journal of Political Economy, 110, 1-36.*

[36] Persson, T. and G. Tabellini (1999), The Size and the Scope of Government: Comparative Politics with Rational Politicians, *European Economic Review* 43, 699-735.

[37] Persson, T. and G. Tabellini (2000), *Political Economics: Explaining Economic Policy*, MIT Press, Cambridge, MA..

[38] Persson, T., G. Tabellini, and F. Trebbi (2000), Electoral Rules and Corruption" NBER Working Paper No. 8154.

[39] Prat, A. (2002a), "Campaign Spending With Office-Seeking Politicians, Rational Voters, and Multiple Lobbies", *Journal of Economic Theory*, 103, 162-89.

[40] Prat, A. (2002b), "Campaign Advertising and Voter Welfare", *Review of Economic Studies*, 69, 997-1017.

[41] Rogoff, K. (1990), Equilibrium political budget cycles, *American Economic Review* 80:21-37.

[42] Seabright, P. (1996), "Accountability and Decentralization in Government:an Incomplete Contracts Model", *European Economic Review* **40**, 61-91.

[43] Treisman, Daniel (2002), "Decentralization and the Quality of Government", unpublished paper, Department of Political Science, UCLA.

[44] Wallis, J. and W.E.Oates (1988), "Decentralization in the Public Sector: an Empirical Study of State and Local Government", in Rosen, H.J. (ed.), in *Fiscal Federalism: Quantitative Studies*, University of Chicago Press.

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

A. Appendix

Proof of Proposition 4. (i) *Full pooling case.* The expected voter payoff under decentralization is as follows. W.l.o.g., suppose that the state is HL. With probability π^2, there are two good incumbents, in which case voter welfare is $W_H + W_L$ in period 1. These incumbents will both be re-elected, and provide voter welfare $EW = qW_H + (1-q)W_L$ in the second. Aggregate voter welfare is therefore $W_H + W_L + \delta 2EW$.

With probability $(1-\pi)^2$, there are two bad incumbents. The first will separate, generating $-\overline{r}$, and will be replaced by a new incumbent, who will generate expected voter welfare of $W^0 = \pi EW - (1-\pi)\overline{r}$. The second will pool, generating W_H followed by $-\overline{r}$ .Aggregate voter welfare is therefore $W_H - \overline{r} + \delta(W^0 - \overline{r})$.

With probability $\pi(1-\pi)$, there will be a good incumbent in the high-cost region and a bad in the low-cost region : the good generates W_H in period 1 and EW in the second, and the bad will pool, generating W_H in period 1 and $-\overline{r}$ in the second. Aggregate voter welfare is therefore $2W_H + \delta(EW - \overline{r})$.

With probability $\pi(1-\pi)$, there will be a bad incumbent in the high-cost region and a good in the low-cost region: the bad will separate, generating $-\overline{r}$ in period 1 and being replaced, so generating W^0 in the second period; and the good will provide W_L in period 1 and EW in the second. Aggregate voter welfare is therefore $W_H - \overline{r} + \delta(W^0 + EW)$.

So, expected voter welfare with decentralization is

$$\begin{aligned} EW_D &= \pi^2 \left[W_H + W_L + 2\delta EW\right] + (1-\pi)^2 \left[W_H - \overline{r} + \delta(W^0 - \overline{r})\right] + \\ &\quad \pi(1-\pi)\left[2W_H + \delta(EW - \overline{r})\right] + \pi(1-\pi)\left[W_L - \overline{r} + \delta(W^0 + EW)\right] \end{aligned}$$

Now consider expected voter welfare with centralization. Recall that the cost state is HL. With probability π, there is a good incumbent, who will provide voter welfare $W_H + W_L$ in period 1. This incumbent will be re-elected, producing expected voter welfare

$$\begin{aligned} & (q^2 + \rho q(1-q))2W_H + ((1-q)^2 + \rho q(1-q))2W_L + 2(q(1-q) - \rho q(1-q))(W_H + W_L) \\ = \; & 2EW \end{aligned}$$

in the next period. With probability $1-\pi$, there is a bad incumbent, who will pool and provide $2W_H$, and who will be re-elected, taking maximum rent $2\overline{r}$. So, expected payoff under centralization is

$$EW_C = \pi(W_H + W_L + \delta 2EW) + (1-\pi)(2W_H - \delta 2\overline{r})$$

So, decentralization is preferred if $EW_D \geq EW_C$.

Consider first the difference in first-period payoffs only- this is the *incentive effect:*

$$\begin{aligned}\Delta_1 &= \pi^2 [W_H + W_L] + (1-\pi)^2 [W_H - \overline{r}] + \pi(1-\pi)2W_H + \pi(1-\pi)[W_L - \overline{r}] - \\ &\quad \pi(W_H + W_L) - (1-\pi)(2W_H) \\ &= -(W_H + \overline{r})(1-\pi) < 0\end{aligned}$$

The incentive effect is negative, as there is less pooling with decentralization, and thus more rent-extraction. Now consider the the difference in second-period payoffs only- this is the *selection effect:*

$$\begin{aligned}\Delta_2 &= \pi^2 2EW^* + (1-\pi)^2(W^0 - \overline{r}) + \pi(1-\pi)(EW^* - \overline{r}) + \pi(1-\pi)(EW^* + W^0) \\ &\quad -\pi 2EW^* - (1-\pi)(-2\overline{r}) \\ &= (1-\pi)(W^0 + \overline{r}) > 0\end{aligned}$$

This is positive, as de-selection of bad incumbents is more likely with decentralization. Overall, we get:

$$EW_D - EW_C = \Delta_1 + \delta\Delta_2 = (1-\pi)[-(W_H + \overline{r}) + \delta(W^0 + \overline{r})]$$

So, decentralization dominates if

$$(W_H + \overline{r}) \leq \delta(W^0 + \overline{r}) = \delta\left(\pi q W_H + \pi(1-q)W_L - (1-\pi)\overline{r} + \overline{r}\right)$$

which reduces to

$$\delta\pi \geq \frac{W_H + \overline{r}}{EW + \overline{r}}$$

(ii) The proof for *partial pooling case* closely follows that of the full pooling case, and so is omitted.

(iii) For *hybrid equilibrium* we consider first difference in first-period welfares (incentive effects). If the state is HH incumbents act the same with centralization and decentralization and there is no welfare difference. When the state is HL or LH (with probability $2\Pr(HL)$) the bad incumbent separates with centralization but pools under decentralization in low cost region with probability λ_D producing incentive gain of $W_H + \overline{r}$ in the low cost region. When the state is LL, the probability of pooling differs: from proposition 5, there is less pooling with centralization $\lambda(\rho) < \lambda_D$ producing incentive gain of $(\lambda_D - \lambda(\rho))$ $(W_H + \overline{r})$ in each region. So first-period welfare difference is

$$\begin{aligned}\Delta_1 &= 2\Pr(HL)(1-\pi)\lambda_D(W_H+\overline{r}) \\ &\quad + \Pr(LL)(1-\pi)(\lambda_D-\lambda(\rho))2(W_H+\overline{r}) \\ &= 2(1-\pi)[\Pr(HL)\lambda_D+\Pr(LL)(\lambda_D-\lambda(\rho)](W_H+\overline{r})>0\end{aligned}$$

Turning to second-period, when state is HH incumbents act the same with the same chance of reelection and second-period welfare is the same. When state is LL incumbents pool with different probabilities ($\lambda(\rho) < \lambda_D$) but have the same chance of reelection σ and so again there is no second-period welfare difference. When state is HL or LH, with decentralization bad incumbent in low cost region pools and is reelected with probability σ whereas bad incumbent separates with centralization and is voted out. So centralization is more likely to vote out bad incumbent producing selection gain from centralization of $\sigma(W^0+\overline{r})$. Also, when state is HL or LH good type chooses intermediate tax τ_{HL} with centralization and is reelected, whereas with decentralization good type in high cost region is only re-elected with probability σ. So centralization is also less likely to vote out good incumbent producing selection gain of $(1-\sigma)(EW-W^0)$. It follows that second-period welfare difference is

$$\begin{aligned}\Delta_2 &= -2\Pr(HL)[(1-\pi)\sigma(W^0+\overline{r})+\pi(1-\sigma)(EW-W^0)] \\ &= -2\Pr(HL)[(1-\pi)\sigma\pi(EW+\overline{r})+\pi(1-\sigma)(1-\pi)(EW+\overline{r})] \\ &= -2\Pr(HL)\pi(1-\pi)(EW+\overline{r})<0\end{aligned}$$

Overall, we get that decentralization dominates if :

$$EW_D - EW_C = \Delta_1 + \delta\Delta_2 \geq 0$$

which requires[29]

$$\delta\pi \leq \left[\lambda_D + \frac{\Pr(LL)}{\Pr(HL)}(\lambda_D-\lambda(\rho))\right]\frac{W_H+\overline{r}}{EW+\overline{r}}$$

Using $\frac{\Pr(LL)}{\Pr(HL)} = \frac{1-(1-\rho)q}{(1-\rho)q}$, $\lambda_D = \frac{q}{1-q}$ and from proposition 5 $\lambda(\rho) = \frac{q^2+\rho q(1-q)}{(1-q)^2+\rho q(1-q)}$, it is easily checked that

$$\lambda_D + \frac{\Pr(LL)}{\Pr(HL)}(\lambda_D-\lambda(\rho)) = 1$$

[29] For $\Pr(HL) \neq 0$.

and the condition reduces to

$$\delta\pi \leq \frac{W_H + \overline{r}}{EW + \overline{r}}$$

which completes the proof. □

Proof of Proposition 5. First, note that the options for the bad incumbent as in the independent case. As before, let the probability of pooling in asymmetric state HL (or LH) be ξ, and the probability of pooling in symmetric state LL be λ. Then, using formulae (5.1), voters' posterior belief that the incumbent is benevolent, conditional on the tax rate and public good level are the same as without cost correlation, except that now

$$\Pr(b\,|\tau_H, g_H) = \frac{\pi[q^2 + \rho q(1-q)]}{[q^2 + \rho q(1-q)]\pi + (1-\pi)[\lambda((1-q)^2 + \rho q(1-q)) + 2\xi(q(1-q) - \rho q(1-q))]} \tag{A.1}$$

Consider now the condition for an equilibrium with the maximum degree of pooling i.e. $\lambda = \xi = 1$ (the *full pooling equilibrium).* In this equilibrium, from (A.1),

$$\Pr(b\,|\tau_H, g_H) = \frac{\pi[q^2 + \rho q(1-q)]}{\pi[q^2 + \rho q(1-q)] + (1-\pi)[1 - q^2 - \rho q(1-q)]} \tag{A.2}$$

This can be an equilibrium only if $\Pr(g\,|\tau_H) \geq \pi$ (otherwise, the bad politician would prefer to take maximum rents, anticipating not being re-elected), which requires which requires from (A.2) after some simplification that

$$q^2 + \rho q(1-q) \geq 1 - q^2 - \rho q(1-q) \Longrightarrow 2(1-\rho)q^2 + 2\rho q \geq 1$$

So, the critical value of q above which a pooling equilibrium exists solves

$$2(1-\rho)q^2 + 2\rho q = 1 \Longrightarrow q = q(\rho) \tag{A.3}$$

It is easily checked that $1/2 \leq q(\rho) \leq 1/\sqrt{2}$ and is decreasing in ρ with $q(0) = 1/\sqrt{2}$ and $q(1) = 1/2$. We also require that whatever the cost realizations, the bad politician prefers to take the rent from pooling rather than go for maximum rent extraction. But this condition is unaffected by the degree of cost correlation, and so is $\hat{r} \geq 2(1-\delta)\overline{r}$ as before.

Now consider the *partial pooling equilibrium.* This requires $\lambda = 1,\ \xi = 0$. Then, from (A.1),

$$P(b\,|\tau_H, g_H) = \frac{\pi[q^2 + \rho q(1-q)]}{\pi[q^2 + \rho q(1-q)] + (1-\pi)[(1-q)^2 + \rho q(1-q)]}$$

Again, this can be an equilibrium only if $P(b\,|\tau_H, g_H) \geq \pi$ which requires

$$q^2 + \rho q(1-q) \geq (1-q)^2 + \rho q(1-q)$$

which reduces to $q \geq 1/2$. We also require $\hat{r} \geq (1-\delta)\overline{r}$ as before.

Turning now to the conditions for *separating equilibrium*, it must be the case that *whatever* the cost realizations, the bad PM prefers to take maximum rents $2\overline{r}$ now rather than pool. This occurs when $2\overline{r} > 2\hat{r} + \delta 2\overline{r}$, or $\hat{r} < (1-\delta)\overline{r}$, a condition which is clearly independent of cost correlations, so the conditions for this equilibrium are unaffected.

Finally, the *hybrid equilibrium*. When $q < 1/2$ and $\hat{r} \geq (1-\delta)\overline{r}$, the bad incumbent wants to pool but cannot be re-elected if he pools with probability 1. Hence, for such a equilibrium, we need $\lambda, \xi < 1$ such that

$$\begin{aligned} P(b\,|\tau_H, g_H) &= \frac{\pi[q^2 + \rho q(1-q)]}{\pi[q^2 + \rho q(1-q)] + (1-\pi)[\lambda((1-q)^2 + \rho q(1-q)) + 2\xi(q(1-q) - \rho q(1-q))]} \\ &= \pi \end{aligned}$$

This requires

$$q^2 + \rho q(1-q) = \lambda((1-q)^2 + \rho q(1-q)) + 2\xi(q(1-q) - \rho q(1-q))$$

Because the bad incumbent prefers to pool when the state is LL than when HL or LH, he will set $\lambda = 1$ if possible. But with $\lambda = 1$ we get

$$2q - 1 = 2\xi(q(1-q) - \rho q(1-q))$$

which is impossible given $q < 1/2,\ \xi > 0$. So $\xi = 0$ and

$$\lambda = \frac{q^2 + \rho q(1-q)}{(1-q)^2 + \rho q(1-q)} = \lambda(\rho)$$

as required. □

Proof of Proposition 6. By definition, a re-election rule with yardstick competition has the following feature, $\sigma(\tau_H, \tau_H) > \sigma(\tau_H, \tau_L)$. The associated levels of public good provision is suppressed for clarity. Since only benevolent type will choose τ_L, we also have $\sigma(\tau_L, x) = 1$ for $x = \overline{\tau}, \tau_H, \tau_L$. So, there are two possibilities for pooling by the bad incumbent.

1. $\Pr(b\,|\tau_H, \tau_H), \Pr(b\,|\tau_H, \tau_L) \geq \pi$. In this case, the re-election rule has no yardstick competition, $\sigma(\tau_H, \tau_H) = \sigma(\tau_H, \tau_L) = 1$ and the incumbent in each region will be re-elected if he pools.

2. $\Pr(b\,|\tau_H,\tau_H) \geq \pi > \Pr(b\,|\tau_H,\tau_L)$. In this case, the re-election rule is with yardstick competition, $\sigma(\tau_H,\tau_H) = 1 > \sigma(\tau_H,\tau_L) = 0$, the incumbent in each region will not be re-elected if he pools when tax is low in the other region.[30]

$$\begin{aligned}
\Pr(\tau_H,\tau_H\,|b) &= \pi[q^2+\rho q(1-q)] + (1-\pi)\lambda q(1-q)(1-\rho) \\
\Pr(\tau_H,\tau_H\,|nb) &= \lambda\pi q(1-q)(1-\rho) + \lambda^2(1-\pi)((1-q)^2+\rho q(1-q)) \\
\Pr(\tau_H,\tau_L\,|b) &= \pi q(1-q)(1-\rho) \\
\Pr(\tau_H,\tau_L\,|nb) &= \pi(1-q)(1-(1-\rho)q)\lambda
\end{aligned}$$

So,

$$\Pr(b\,|\tau_H,\tau_H) = \frac{\pi}{\pi+(1-\pi)l},\ l = \frac{\Pr(\tau_H,\tau_H\,|nb)}{\Pr(\tau_H,\tau_H\,|b)}$$

where l is the non-benevolent likelihood ratio. For an equilibrium to be pooling we need $l \leq 1$ at $\lambda = 1$. This requires[31]

$$\pi \geq 0.5 + \frac{(0.5-q)}{(1-2q)^2+4\rho q(1-q)} = \pi(q,\rho)$$

Define $\tilde{q}(\pi,\rho) = \pi^{-1}(q,\rho)$. The condition for an equilibrium to be pooling becomes

$$q \geq \tilde{q}(\pi,\rho)$$

To get a pooling equilibrium with yardstick competition, we must also have $\pi > \Pr(b\,|\tau_H,\tau_L)$ which requires the non-benevolent likelihood ratio

$$l' = \frac{\Pr(\tau_H,\tau_L\,|nb)}{\Pr(\tau_H,\tau_L\,|b)} \geq 1$$

which reduces to

$$q \leq \widehat{q}(\rho) = \frac{1}{2(1-\rho)}$$

which completes the proof.□

[30] It is easily checked with pooling, necessarily $\Pr(b\,|\tau_H,\tau_H) \geq \Pr(b\,|\tau_H,\tau_L)$ for all $q \in (0,1)$ and $\rho \in (0,1)$.

[31] This has the following properties: $\pi(0,\rho) = 1, \pi(1,\rho) = 0,\ \pi(0.5,\rho) = 0.5$, and $\pi(q,\rho)$ decreasing in ρ for $q < 0.5$, and increasing in ρ for $q > 0.5$, and finally $\pi(q,1) = 1-q$. Also, $\pi(q,\rho)$ decreasing in q for $\rho \in (0,1)$.

Proof of Proposition 8.

The proof follows closely the proof of Proposition 4 for hybrid equilibrium; where it is shown that decentralization dominates if :

$$\delta\pi \leq \left[\lambda_D + \frac{\Pr(LL)}{\Pr(HL)}(\lambda_D - \lambda(\rho))\right]\frac{W_H + \overline{r}}{EW + \overline{r}}$$

where $\lambda_D - \lambda(\rho)$ is the difference in the probability of pooling between decentralization and centralization. It was then established that the expression in bracket is equal to one. The effect of yardstick competition on the hybrid equilibrium is to reduce the probability of pooling from λ_D to λ_{YC} so that the condition above becomes

$$\delta\pi \leq \left[\lambda_{YC} + \frac{\Pr(LL)}{\Pr(HL)}(\lambda_{YC} - \lambda(\rho))\right]\frac{W_H + \overline{r}}{EW + \overline{r}}$$

Let $\gamma(\rho) \equiv \lambda_{YC} + \frac{\Pr(LL)}{\Pr(HL)}(\lambda_{YC} - \lambda(\rho))$. Then $\gamma(\rho) < \lambda_D + \frac{\Pr(LL)}{\Pr(HL)}(\lambda_D - \lambda(\rho)) = 1$ and so the condition for decentralization to dominate is more stringent with yardstick effect. It is easily checked that $\gamma(\rho)$ is decreasing with $\gamma(\rho) < 1$ for all ρ and $\gamma(\rho) < 0$ above a critical level $\rho^* < 1$,as required $\square$