Fiscal Decentralization, Electoral Accountability, and Fiscal Uniformity¤

Jean Hindriks[y] Ben Lockwood[z]

First Version: March 2004
This Version: November 2006

Abstract

This paper studies the relationship between …scal decentralization and electoral accountability, in a setting where decentralization impacts upon incentive and selection e¤ects of elections. The e¤ect of …scal centralization on voter welfare works through two channels: (i) via its e¤ect on the probability of pooling by the bad incumbent; (ii) conditional on the probability of pooling, the extent to which, with centralization, the incumbent can divert rents in some regions without this being detected by voters in other regions (selective rent diversion). Both these e¤ects depend on the information structure; whether voters only observe …scal policy in their own region, in all regions More voter information does not necessarily raise voter welfare. Under some conditions, voters would choose uniform …scal policy ex ante (uniform tax rates and also possibly expenditures across regions) to constrain selective rent diversion.

JEL Classi…cation: D72, D73, H41

Keywords: …scal federalism; decentralization; elections; accountability

¤Pervious versions of this paper were circulated under the title: "Fiscal Decentralization and Electoral Accountability: Incentives, Separation, and Voter Welfare".

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

[z]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. For example, nearly all the large Latin American countries have initiated some form of …scal decentralization in the last decade e.g. Bolivia (Faguet(2004)), as have Indonesia, the Philippines, and Pakistan, to name just a few. China and Russia's transition from socialism involves various aspects of decentralization. Moreover, it is actively promoted as a development strategy by organizations such as the World Bank[1]. There have also been similar reforms in high-income countries, e.g. devolution of tax and spending powers to Scotland in the UK in 1999, and in Italy, starting in 1993 with the introduction of a municipal property tax.

What are the costs and bene…ts of …scal decentralization? One can distinguish two[2] quite di¤erent approaches in the literature. The …rst approach, which could be called the standard theory, derives from the classic work of Oates(1972), and has subsequently generated a large literature[3]. According to this perspective, the cost of decentralization is that inter-regional spillovers are not internalized, and the bene…t is that there is better preference-matching: that is, public goods provided by governments in localities will be better matched to the preferences of the residents of those localities. Oates(1972), and many subsequent contributions, have modelled the preference-matching bene…t by the assumption of policy uniformity i.e. that is, under centralization, local public goods are provided at the same level across localities, irrespective of local preferences. Some more recent contributions (Lockwood(2002), Belsey and Coate(2003)), have relaxed this assumption, modelling prefer-matching is a more sophisticated way.

The second, newer approach, is a political economy one, and focusses on whether decentralization increases the accountability of policy-makers via elections[4]. The key con-

[1] For more details on country decentralization programs, and the World Bank's view of the costs and bene…ts, see http://www1.worldbank.org/publicsector/decentralization/, or World Bank (2000).

[2] There are also now some papers which study the way that lobbying activity varies across …scal regimes. These are discussed in Section 8.1.

[3] Theoretical work includes 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),(1999), Oberholzer-Gee and Strumpf(2002). Empirical work includes Azfar et al.(2001), Faguet(2004), Oberholzer-Gee and Strumpf(2002).

[4] There are also now a number of empirical studies that show that across countries, measures of …scal decentralization are generally negatively correlated with low accountability outcomes, such as corruption and poor governance (s Huther and Shah(1998), Fissman and Gatti(2002), Mello and Barenstein(2001), , Fissman and Gatti(2002a), Henderson and Kuncoro(2004), Treisman (2000),(2002). But, these studies

tributions here are Seabright(1996), and Persson and Tabellini(2000), Chapter 9. These models do not have inter-regional public good spillovers and preference variation between regions, so they do not rely on the traditional spillover and preference-matching arguments. Speci…cally, they consider multi-region, two-period models where the incumbent (whether national or regional) must supply a regional public good in each period. The amount of the public good supplied in any region depends on (…xed) tax revenue, the amount of rent diverted by the incumbent in that region, and on a stochastic region-speci…c productivity shock, which is not known to the incumbent when he chooses …scal policy and rent-diversion. At the end of the …rst period, here is an election where the incumbent faces a challenger. When voters within a region vote, they observe tax and public good provision but cannot distinguish rent-diversion from the productivity shock. So, they re-elect the incumbent if public good provision is su¢ciently high.

In this setting, voter welfare is determined by the amount of …rst-period rent-diversion (in the second period, the incumbent will always take maximum rent). In turn, …rst-period rent-diversion can be higher or lower with decentralization than with centralization. The factor driving it lower is that the bene…t of re-election is higher for the incumbent with centralization: he can extract rent in every region in the second period, rather than just one (the scale e¤ect). A factor driving it higher is that with centralization, the amount of rent-diversion in any given region has a smaller negative impact on the incumbent's re-election probability. This is because with decentralization, voters in any one region are always pivotal in determining the outcome of the election in that region, so there is a strong link between the rent diverted in that region by the incumbent and his re-election probability. With centralization, the voters in a region are pivotal in determining the outcome of the national election only in the event that the particular region is pivotal i.e. half the other regions vote for the incumbent, and half against. We call this the reduced pivot probability e¤ect of centralization.

But, while these papers, in our view, provide a distinctive story of the costs and bene…ts of centralization, they are not without problems. A weakness of Seabright's model is that all policy-makers are identical[5], implying that the voters have no strictly positive incentive to vote out the incumbent. In that case, as is well-known (e.g. Ferejohn(1986)) there are a continuum of equilibrium performance cuto¤s i.e. critical values of public good provision below which the voters will vote for the challenger. This makes it essentially

are not sufceintly structural to distinguish between the lobbying and political agency mechanisms.

[5] In Seabright's model, productivity shocks are characteristics of the regions, not the incumbents, and are i.i.d. across regions.

impossible to make welfare comparisons between …scal regimes, due to multiple equilibria. Persson and Tabellini(2000, Chapter 9.1) resolve this indeterminacy by supposing that the productivity shock is an inherent competence characteristic of the incumbent. Then, voters are not indi¤erent about a performance cuto¤ ex post, because the higher the cuto¤, the more likely it is that the incumbent who passes it is competent. But, they retain Seabright's assumption that the …rst-period incumbent does not observe the shock.

So, in the terminology of Besley and Smart(2003), the work of Seabright(1996) and Persson and Tabellini(2000) only focuses on the incentive[6] e¤ects of elections, not the selection e¤ects[7]. In Seabright, there are no selection e¤ects, as all policy-makers are identical. In Persson and Tabellini(2000), because the incumbent does not know his own competence when he sets …rst-period policy, the probability that an incumbent of given competence wins the election is the same with centralization and decentralization[8].

A second, more minor, weakness of Persson and Tabellini(2000, Chapter 9.1) is that it relies on a second period non-monetary rent (ego-rent) from o¢ce in order to generate conditions under which decentralization could ever be preferred by voters. This is because the scale e¤ect and reduced pivot probability e¤ect enter the rent diversion formula in a ratio, and the former dominates the latter[9]. This is in fact true much more generally, as is also demonstrated in Lockwood(2006). So, their theory is really an incomplete account of the costs and bene…ts of decentralization, because ego-rents are inherently hard to measure - they may be large or small.

This paper addresses the key question of what e¤ect …scal (de)-centralization will have

[6] Elections provide accountability in two senses. First, they allow voters to de-select bad incumbents (selection e¤ects). Second, the selection e¤ect provides an incentive for incumbents to change their behavior in order to increase the probability of re-election (incentive or discipline e¤ects)

[7] it is remarkable that the complete contracting principal-agent theory also ignores the selection e¤ect to consider only incentives. One notable exception is Banks and Sundaram (1998) who study the optimal retention rule in agency problems, and show that in equilibrium the chosen retention rule disciplines the agents (incentive e¤ect) and the retained agents are more productive on average (selection e¤ect).

[8] Speci…cally, as both the incumbent's and challenger's competence levels are random draws from the same distribution, and voters have rational expectations, in equilibrium, the probability that the initial incumbent has a competence level above the expected level of the challenger is simply 0.5. So, with both centralization and decentralization, by construction, the incumbent loses o¢ce with probability 0.5. Lockwood(2006) makes this point more formally, where it is also shown that this is true not only of Persson and Tabellini's model, which has a rather special structure, but essentially all models where in the …rst period, the incumbent is equally well-informed about his own characteristics as the voters.

[9] Speci…cally, setting $R = 0$ in equation (9.8) of Persson and Tabellini(2000), we see that rent-version is lower under centralization as the ratio of the two e¤ects is $3/2$: there are three regions, but the probability that any region is pivotal is $1/2$.

on both discipline and selection e¤ects of elections. As just argued, by the invariance result, for selection e¤ects to be truly endogenous (and thus vary between centralization and decentralization), there must be asymmetric information: the incumbent must be better-informed about his own competence (or some other characteristic) than the electorate. Moreover, there are no exogenous ego-rents from o¢ce. This paper provides a comprehensive analysis of such a model. Our focus is both normative - on the costs and bene…ts for voters of a change in …scal regime - and positive - on the impact of a change in …scal regime on incumbent turnover. We claim that our analysis provides a coherent count of the costs and bene…ts of centralization that is distinct from the standard theory.

Our model has two periods and n regions. In the …rst period, the type of the incumbent policy-maker is determined by random draw: the incumbent may be benevolent or rent-seeking. With decentralization, the policy-maker in region i, knowing his type, then chooses a tax and a level of public good provision in their region. The cost of public good provision is random and not observed by voters, only the incumbent. Voters observe this choice and then vote for the incumbent or the challenger. The type of the challenger is also determined by random draw. In the second period, the winner again chooses a tax and a level of public good provision in their region. Note that in our model, there is no preference variation across regions, and no local public good spillovers,

In this setting, incentive and selection e¤ects are ultimately determined by the incentives the bad incumbent has to pool with the good one - the greater the pooling incentive, the stronger the discipline e¤ect, and the weaker the selection e¤ect. In turn, we show that, with a change in …scal regime, the change in the pooling incentive is determined by two e¤ects. First, the statistical distribution of costs the incumbent of a given type will face generally di¤ers between …scal regimes (the cost distribution e¤ect). For example, with decentralization, an incumbent has a single region and cost can be either high or low. But with centralization, unless costs are assumed perfectly correlated across regions, an incumbent of the same type will face either 0,1,2 or 3 high-cost regions with varying probabilities. This will change the incentives for pooling in a mechanical - if ambiguous - way.

The second e¤ect is the selective rent diversion e¤ect, which comes into play with centralization only when voters cannot observe …scal policy in other regions- partial voter information. Speci…cally, with centralization, if the incumbent wishes to win the election and stay in o¢ce, he can do so most e¢ciently by only imitating the good incumbent in a minimum majority of $m = (n + 1)/2$ regions, and can take unconstrained rent in the other regions. In this sense, he is less accountable to the electorate with centralization than with decentralization. Selective rent diversion has two implications. First, it tends

to decrease the separation probability relative to decentralization, and second, it unambiguously decreases voter welfare with centralization, for a given separation probability.

It does not follow from this, however, that voter welfare is always lower with centralization and partial information than it is with either centralization and full information, or decentralization. This is because the separation probability also changes with the …scal regime due to the cost distribution e¤ect. Because of this, with centralization, it is not generally true that giving voters more information will make them better o¤. This result is comparable to Proposition 5 of Besley and Smart(2003), who demonstrate that yardstick competition between regions (which can only occur when voters are fully informed in our sense) does not necessarily increase voter welfare. The mechanism at work is quite di¤erent, however: in our case, statistical correlation in the cost of producing the public good in each region is not needed.

A second objective of the paper is to revisit the important and controversial assumption of policy uniformity, which as already remarked, is very widely used in the literature on …scal decentralization. This assumption, …rst introduced by Oates(1972), is that the level of local public good provision is the same across localities[10]. This plays several key roles. In earlier models where government is a benevolent social planner, such as Oates(1972), it captures in a crude way the idea that decentralized government is more responsive to local preferences - indeed, without this assumption, centralization would always dominate decentralization. In political economy models, where local public good provision is decided by majority vote, such as Alesina, Angeloni, and Etro(2005), or Bolton and Roland(1998), with centralization, it collapses the policy space down to one dimension thus ensuring that a voting equilibrium exists. But, the puzzle is why voters would commit ex ante to a policy that constrains the actions of incumbents. This paper shows how policy uniformity can emerge endogenously in equilibrium[11].

Speci…cally, we show that ex ante, at a constitutional stage, voters may unanimously prefer to impose uniformity across regions on both per capita expenditures and tax rates, when voter information is partial. This commitment has the cost that it prevents good incumbents from adapting public good supply to variations in regional costs (loss of ‡exibility). The advantage, however, is that it eliminates the possibility of selective rent-diversion when he incumbent is bad. A third e¤ect, is that it also increases the separation

[10] Empirically, the evidence for policy uniformity is mixed. In the US, for example, it is well-known that per capita expenditure via Appropriations Bills varies widely across states (Knight(2006)). In the UK, by contrast, nearly all central government expenditure is determined by formula, and designed to equalize per capita provision of goods across regions, appropriately adjusted for di¤ering needs.

[11] Harstad(2007) also has an explanation as to why this can occur, further discussed in Section 8.1.

probability, which voters may or not like.

We also consider a constitutional commitment to uniform taxes only. This is an interesting case, because in reality, uniform taxes are very widely observed, whereas uniform expenditures are not. Here, the advantages of uniformity are stronger. Speci…cally, there is now no loss of ‡exibility, but there is still a gain as selective rent-diversion is constrained - although no longer eliminated. So, our model can explain the stylized fact that tax rate uniformity is much more commonly observed than expenditure uniformity.

The layout of the remainder of the paper is as follows. Section 2 sets up the model. Section 3 studies the case of decentralization for the benevolence model. Sections 4, 5, and 6 study the cases of centralization with full voter information, partial voter information, and uniform policy respectively for the benevolence model. Section 7 makes the case that most of the key results are robust as they also hold for a model where politicians di¤er in competence, and also considers spatial cost correlation. Section 8 discusses related literature and concludes.

2. The Model

2.1. Preliminaries

There are two time periods $t = 1, 2$ and an odd number of regions $i = 1, ..n,$ with $n \geq 3$. In each region in each time period, an incumbent politician makes 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 …rst-period …scal policy. In each region, there are $v \geq 3$ identical voters[12] who do not vote for weakly dominated alterative, implying - with two alteratives - that they vote or their more preferred alterative. With decentralization, there are n incumbents and n challengers: one in each region. With centralization, there is one incumbent and challenger.

In each region, there are a 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

[12] The reason for this assumption is that it ensures that in equilibrium, any individual voter is never pivotal. The problem that can arise otherwise is the following. Suppose that there are three regions, with one voter in each, and that voter information is partial. Then, the equilibrium described in Proposition 3 is in fact not an equilibrium. Suppose w.l.o.g. that the bad incumbent pools in regions 1 and 2. Then, he equilibrium requires that voters 1 and 2 vote for the incumbent, and voter 3 against i.e. voters 1,2 are pivotal. On the other hand, if the incumbent is good, all voters will vote or the incumbent. So, voters 1,2 are pivotal i¤ the incumbent is bad. In this case, it is clearly rational for 1,2 to vote against the incumbent.

good, normalized to unity. The public good is …nanced 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 $0 \cdot \tau_t^i \cdot 1$ so the endowment can be fully taxed. The incumbent can also divert tax revenue of amount r_t^i up to a maximum level of $r \cdot 1$ per region in period t. Both voters and politicians have the same discount factor, $0 < \delta < 1$.

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. So, the budget constraint is

$$\sum_{i=1} c_t^i g_t^i + r_t = \sum_{i=1} \tau_t^i$$

where r_t are the rents (if any) diverted from aggregate tax revenue.[13]

It is a widely observed "stylized fact" that centrally set tax rates are uniform across regions, and consequently, almost all the literature on …scal centralization assumes that the tax rate is uniform with centralization i.e. $\tau_t^i = \tau_t$. We do not wish to impose the assumption ex ante, for reasons discussed at the end of this section.

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

Benevolence. A "good" politician derives utility only from the welfare of the voters in his jurisdiction: in particular, he maximizes the sum or average of these utilities. A bad politician cares only about discounted sum of rents diverted. Either type is equally competent in producing the public good. The cost of the public good is high in any region and period with probability $q = 0.5$[14], and c_t^i is uncorrelated across regions and periods.

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

[14] Imposing this constraint on q rules out the "hybrid" equilibrium of Besley and Smart (2003). The reason for this is discussed further in Section 3 below.

Competence. Any politician maximizes the discounted sum of rents diverted; conditional on this, he has a lexicographic secondary preference for supplying the public good at its optimal level. The cost c_t^i is uncorrelated across time and regions, but the probability q_t^i that the unit cost is high in region i at time t is conditional on the competency of the incumbent. A "good" politician is more competent than the bad. In particular, if the incumbent is good, then $q_t^i = 0$, and if the incumbent is bad, $q_t^i = q$, with $1 > q > 0$.

Finally, we state our assumptions about the information voters have about …scal policy in other regions. we study three possible scenarios:

1. Full voter information; at time t, the voters in i can observe $(g_t^i, \tau_t^i)_{i=1,..n}$:
2. Partial voter information; at time t, the voters in i can observe only (g_t^i, τ_t^i):
3. Uniform policy: at time t, the voters in i can observe only (g_t^i, τ_t^i), but the constraint is imposed that $\tau_t^i = \tau_t,\ g_t^i = g_t$ all i.

A variant of 1 and 3 studied in Section 6.3 is where voters only have full information about taxes, or where only taxes are constrained to be uniform. This is a particularly realistic case, as argued below.

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 di¤erence between centralization and decentralization in the established literature: there are no economies of scale, there are no spillovers between regions, voters do not di¤er in tastes for the public good, either within or between regions[15]. So, the di¤erence in outcome between centralization and decentralization is entirely due to the di¤erence 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 in the benevolence model 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,[16] and in either region and period, the incumbent will provide the public good e¢ciently, conditional on cost c_L or c_H. This is true whether there is centralization

[15] Also, by assuming a …xed number of incumbents and challengers, we rule out rule out the district magnitude e¤ect bias in favor of centralization, that larger electoral districts lower barriers to entry and favor competition improving political discipline and selection (Myerson (1993), Persson et al (2000)).

[16] There can be other equilibria where the incumbent is no re-elected, as all voters are always indi¤erent

or decentralization. Of course, e¢cient public good provision, denoted $g_k, k = L, H$ is implicitly de…ned by the Samuelson rule $H^0(g_k) = c_k$. 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 benevolence is an n ¡ region version of Besley and Smart (2003). We should stress that in their paper, they do not consider centralized decision-making as we have de…ned it here: their benchmark is decentralization without "competition" between regions, and then the impact on selection and incentive e¤ects of introducing either tax or yardstick competition is studied. Finally, Kotsogiannis and Schwager (2004) consider a two-region model of policy innovation and elections, which is more loosely related to this one[17].

3. Decentralization

We solve backward to obtain a unique Bayes-Nash equilibrium in either region.[18] 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 by setting a tax of $\tau = r$, and providing no public good i.e. $g = 0$. So, all voters prefer the honest policy-maker.

In the …rst-period, assume for the moment that good incumbent in either region will be elected with probability 1 if he behaves non-strategically i.e. makes exactly the same policy choices as in the second period. We will shortly verify when this is equilibrium behavior for the voters. In this case, the best strategy for the good incumbent is to behave non-strategically.

As for the bad type, when cost is high, he always prefers to take maximum rent in

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

[17] In their model, there are two regions, but the two regimes studied are not …scal centralization and decentralization in the conventional sense. Rather, they compare a unitary system, where a national policy-maker chooses whether to innovate in policy or not (a binary choice) in either region, to federal system, where two incumbents initially choose policy innovation at the regional level, and then run in an election to be national policy-maker (president) in the second period.

[18] Obviously, the results in this section recapitulate Section 3 of Besley and Smart (2003): the reader is referred to that paper for deeper discussion of the issues.

the …rst period, rather than imitate 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. When cost is low, the bad type has only two options that may potentially be optimal. First, he can set (g_H, τ_H) and take $\hat{} = g_H(c_H - c_L)$ in the form of rents: call this the pooling strategy. Second, he can set $g = 0$, and take maximal rents, by setting $\tau = r$: call this the separating strategy. We assume throughout that $\hat{} < r$.

We are assuming for the moment that any incumbent who chooses (g_H, τ_H) will be re-elected. So, when cost is low, the payo¤s to separating and pooling for the bad incumbent are $r + \delta.0$ and $\hat{} + \delta.r$ respectively. There is therefore a pooling equilibrium[19], where the bad politician imitates the good one when the cost of public good provision is low, and is re-elected with probability 1 in that event if $\hat{} + \delta.r \geq r$, i.e. $\hat{} \geq (1 - \delta)r$, and a separating equilibrium where bad politician does not imitate the good one even when the cost of public good provision is low if $\hat{} < (1 - \delta)r$.

To con…rm that the pooling equilibrium exists, we only need con…rm that voters are willing to re-elect the incumbent if they observe[20] (g_H, τ_H). A voter's posterior belief that the incumbent is good i.e. benevolent, conditional on observing (g_H, τ_H) is

$$q(\tau_H, g_H) = \frac{\pi q}{\pi q + (1 - \pi)(1 - q)} \tag{3.1}$$

Note from (3.1) that as $q \geq 0.5$, $q(\tau_H, g_H) \geq \pi$, so the voters are indeed willing to re-elect the incumbent after observing (g_H, τ_H).

So, in any region, the ex ante probability that a bad incumbent separates (the selection e¤ect) is

$$s_D = \begin{cases} q & \text{if } \hat{} > (1 - \delta)r \\ 1 & \text{if } \hat{} < (1 - \delta)r \end{cases} \tag{3.2}$$

It is convenient for what follows to show s_D as a function of the discount factor, δ. This is done in Figure 1. It is clear that δ is a key parameter here, as the higher δ, the greater the incentives for pooling, and thus the lower is the separation probability.

[insert …g 1 here]

Finally, note the role of the assumption that $q \geq 0.5$. This rules out the scenario where the incumbent wants to pool by setting τ_H, g_H, assuming that he can be re-elected, but

[19] We assume w.l.o.g. a tie-breaking rule where if the incumbent is indi¤erent between separating and pooling, he pools.

[20] As a bad incumbent will never set τ_L, g_L, then voters are always willing to re-elect the incumbent having observed τ_L, g_L.

the voters place a low probability on $c_i = c_H$, and thus will not be willing to re-elect the incumbent if he sets τ_H, g_H. In this case i.e. when $q < 0.5$, and $\hat{} > (1 - \delta)r$, Besley and Smart construct a hybrid equilibrium, where both the bad incumbent and voters randomize. However, for some parameter values this equilibrium does not satisfy the Cho-Kreps stability criterion (Lockwood(2005)). The reason is that the "good " type has an incentive to strategically distort public good provision when cost is high to signal his type to the electorate, in order to avoid being replaced by a (possibly) bad challenger. In this case, a stable fully separating equilibrium can be constructed. We wish to avoid these rather technical issues, and do so by assuming $q \geq 0.5$.

4. Centralization with Full Voter Information

4.1. Equilibrium

We solve backward. In the second-period, the benevolent policy-maker will provide the optimal public good level in each region given local costs and charge a tax equal to the cost. The non-benevolent policy-maker will provide no public good and take maximum rent, regardless of the cost con…guration. So, all voters prefer the benevolent policy-maker in period 2. In the …rst-period, the benevolent incumbent behaves non-strategically and so will make exactly the same policy choices as in the second period.

So, it remains to characterize the …rst-period behavior of the non-benevolent incumbents. At the end of the …rst period, all voters observe $(g_i, \tau_i)_{i=1,...n}$. Now, if an incumbent extracts maximum rents in one region (by setting $g^i = 0, \tau^i = r$) this will be observable by the voters in all regions, and the incumbent will thus reveal his type and lose the election.[21] This means that there are only two …rst-period strategies that are potentially optimal: pooling, which is $(g_i, \tau_i) = (g_H, \tau_H), i = 1, ..n$, and separating, which is $(g_i, \tau_i) = (0, r),\ i = 1, ..n$. Finally, say a region is high-cost (low-cost) if $c_1^i = c_H$ ($c_1^i = c_L$). We then have the following result:

Proposition 1. Assume that $q \geq (1/2)^{1/n}$. Suppose that $k \in \{0, 1, ..n\}$ of the regions are high cost. If $k = n$, the incumbent always separates. If $k < n$, the incumbent pools if $\hat{} \geq \frac{n}{n-k}(1 - \delta)r = r_k$ and separates otherwise.

Note the key feature of Proposition 1: the more high-cost regions there are, the higher …rst-period rent relative to maximum rent $\hat{}/r$ has to be to induce the bad incumbent to pool. Note also that we make an assumption that $q \geq (1/2)^{1/n}$: this plays the role of

[21] We call this the "information consolidation" e¤ect of centralization.

ruling out a possible hybrid equilibrium, as in the decentralization case.

4.2. Separation Probabilities

Note that r_k is strictly increasing in k, and strictly so when the r_k are strictly positive, with r_n = + 1 . So, we can write down a formula for the ex ante probability of separation by the bad incumbent. Let p_k be the probability that k or fewer regions are high-cost[22]. Then, if r_k · ^ $< r_{k+1}$, the incumbent pools only if there are k or fewer high-cost regions, which occurs with probability p_k so he separates with complementary probability 1 ¡ p_k. If ^ $< r_0$, the incumbent separates no matter what k is. That is:

$$s_F = \begin{cases} 1 - p_k, & r_k \cdot \hat{} < r_{k+1},\ k = 0, 1..n - 1 \\ 1, & \hat{} < r_0 = (1 - \delta)r \end{cases} \tag{4.1}$$

How does s_F compare to s_D? It is convenient to use the Figure 1 above to illustrate this. The separation probability s_F as a function of δ, is superimposed on Figure 1 to give Figure 2.

[Insert Fig 2 here]

. When δ is low, i.e. below 1 ¡ $\frac{\hat{}}{r}$, separation always occurs, even if all regions are low-cost. When δ is high, i.e. above 1 ¡ $\frac{\hat{}}{nr}$, separation never occurs, unless all regions are high-cost, which occurs with probability q^n. Generally, s_F is monotonically decreasing in δ. Note that s_F can be above or below s_D; indeed, it is apparent from Figure 2 that as long as there is some pooling in either regime i.e. 1 ¡ $\frac{\hat{}}{r} < \delta$, there is a critical value of δ below which the separation probability is higher with centralization, and above which it is lower. In particular, when δ ¸ 1 ¡ $\frac{\hat{}}{nr}$, the bad incumbent is harder to "detect" with centralization than with decentralization, as he only reveals himself when all regions are high-cost (whereas the bad incumbent with decentralization reveals himself whenever his own region only is high-cost).

The intuition for this di¤erence is simply the cost distribution e¤ect referred to in the introduction. Note that with k high-cost regions, the opportunity cost (per region) of pooling with centralization is simply one nth of maximum rent with separation, rn, minus the maximum rent with pooling ^(n ¡ k) i.e. r ¡ ^(1 ¡ $\frac{k}{n}$) This increases quite smoothly with k, especially when n is large. This is to be contrasted with the decentralization case,

[22] Note p_k = Pr(X · k), where X is a random variable with a Binomial distribution with parameters q, n.

where the opportunity cost of pooling, i.e. r ¡ ^(1 ¡ k), k 2 f0, 1g, changes discontinuously when k rises from zero to 1.

4.3. Voter Welfare

We now turn to welfare analysis. Let EW_F, EW_D be the expected present value of welfare to the voter of any region calculated at the beginning of period 1, before the type of the incumbent and the cost shocks are determined, under full-information centralization and decentralization respectively.

It is useful to develop the formulae for EW_F, EW_D as they will make clear that the welfare ranking of centralization and decentralization depends entirely on the separation probabilities. De…ne

$$W_k = H(g_k) \text{ ¡ } c_k g_k, \; EW_G = qW_H + (1 \text{ ¡ } q)W_L, \; EW = \pi EW_G \text{ ¡ } (1 \text{ ¡ } \pi)r$$

where EW_G and EW denote the expected per period welfare produced by a good incumbent and by an incumbent of unknown type, given that the bad incumbent extracts maximum rent, respectively. Then, a very convenient and insightful way of writing voter welfare with decentralization is

$$EW_D = (1 + \delta)EW + (1 \text{ ¡ } \pi)(1 \text{ ¡ } s_D)(¢_I \text{ ¡ } \delta\pi ¢_S) \tag{4.2}$$

where $¢_I = r + W_H$, and $¢_S = r + EW_G$. The explanation is as follows. First, assume that the bad incumbent separates with probability 1 in period 1. Then, he will take maximum rent in period 1. So, in either period, a bad incumbent will generate voter welfare ¡ r. So, in either period, expected welfare (before the incumbent type is known) is EW. This generates baseline present value of welfare $(1 + \delta)EW$.

Next, we must add to this baseline the change in expected voter welfare given that the bad incumbent pools rather than separates, multiplied by the probability that he is bad and he pools. The latter is of course$(1 \text{ ¡ } \pi)(1 \text{ ¡ } s_D)$. The change in expected voter welfare given that the bad incumbent pools rather than separates is composed of two terms. The …rst is a selection loss $\delta\pi ¢_S$. This is due to the fact that if the bad incumbent is replaced, voters can expect a second-period payo¤ of EW rather than ¡ r. So the expected gain to voters in that region of selection is $EW + r = \pi(r + EW_G)$, which of course has present value $\delta\pi(r + EW_G) = \delta\pi ¢_S$. So, if the bad incumbent pools rather than separates, the voters lose amount $\delta\pi ¢_S$.

The other side of the coin is that voters have a …rst-period incentive or discipline bene…t when the bad incumbent pools rather than separates. If the bad incumbent pools,

he will always do so by setting …scal policy (g_H, τ_H) generating voter welfare W_H. If he separates, voters of course get ¡r. So, the incentive or discipline bene…t of pooling is ¢$_I$ = W_H + r.

By exactly the same reasoning, the equilibrium welfare with centralization is

$$EW_F = (1 + \delta)EW + (1 ¡ \pi)(1 ¡ s_F)(¢_I ¡ \delta\pi ¢_S) \tag{4.3}$$

where s_F replaces s_D, but otherwise, the formula is the same. So, the di¤erence is:

$$EW_F ¡ EW_D = (s_F ¡ s_D)(\delta\pi ¢_S ¡ ¢_I) \tag{4.4}$$

So, the welfare comparison depends entirely on (i) whether the separation probability is smaller or larger with decentralization than with centralization, and (ii) whether the selection bene…t of separation, $\delta\pi$¢$_S$, is smaller or larger than the incentive bene…t of pooling i.e. ¢$_I$.

Proposition 2. With either …scal arrangement, voter welfare is increasing in the separation probability if ¢$_I$/¢$_S$ $> \delta\pi$. In this case, voter welfare is higher with whichever arrangement gives the higher separation probability. With either …scal arrangement, voter welfare is decreasing in the separation probability if ¢$_I$/¢$_S$ $< \delta\pi$. In this case, voter welfare is lower with whichever arrangement gives the higher separation probability. If ¢$_I$/¢$_S$ = $\delta\pi$, voters are always indi¤erent between centralization and decentralization.

The condition determining voter preference over separation is intuitive. The bene…ts of separation come in the second period, and only occur with probability π. So, δ and π must be su¢ciently high for voters to prefer separation.

5. Centralization with Partial Voter Information

5.1. Equilibrium

Second-period behavior of an incumbent of a given type (good or bad) is the same as with full voter information. So, all voters prefer the benevolent policy-maker in period 2. In the …rst-period, the benevolent incumbent behaves non-strategically and so will make exactly the same policy choices as in the second period.

To analyze the …rst-period behavior of the non-benevolent incumbents, we introduce the following terminology. The incumbent separates in region i if he chooses g^i, τ^i 6= (g_L, τ_L) or (g_H, τ_H), and pools in region i otherwise. As voters only observe …scal policy in their own region, all voters in i vote for the incumbent if he pools, and for the challenger

otherwise. So, w.l.o.g, we can assume that if the incumbent separates in region i, he sets $(g^i, \tau^i) = (0, r)$.

Now consider what the bad incumbent has to do to be elected. Clearly, because the voting rule is majority rule, and because all voters in a region vote the same way, all he has to do is "satisfy" the voters in majority of regions, i.e. pool in a majority of regions. In other words, compared to the case of full information, the incumbent has the option of selective pooling, and he will clearly make use of this option.

Moreover, note that whenever the incumbent does choose to selectively pool, in order to be re-elected, he will do so by setting (g_H, τ_H) whenever he pools in region i. Also, he will choose to pool in low-cost regions wherever possible. That is, if the number of high-cost regions is $k < m$, he will pool in m randomly selected low-cost regions. If $k \geq m$, he will pool in all the available $n - k$ low cost regions, plus $m - (n - k)$ high-cost regions. Given these observations, it is easy to prove:

Proposition 3. If $\delta \geq \frac{m}{n}$, the incumbent always selectively pools, and is re-elected. If $\delta < \frac{m}{n}$, the incumbent will selectively pool if $\hat{r} \geq \frac{m - n\delta}{\min\{m, n-k\}} r - r_k$, in which case he is re-elected. Otherwise, he sets $g_i = 0$, $\tau = r$ in all regions, and is not re-elected.

Note that as in the case of full information, the more high-cost regions there are, the higher …rst-period rent $\hat{r}$ has to be to induce the bad incumbent to pool. however, the critical value of $\hat{r}$ is lower than in the case of full voter information (for a formal proof, see Section 5.4 below), as the incumbent now has the option of selective rent-diversion. Note now that we only need the condition $q \geq 0.5$ as now the inference problem facing the voter in any region is the same as with decentralization.

5.2. Separation Probabilities

We assume in this section that $\delta < \frac{m}{n}$, because if not, the separation probability is always zero, from Proposition 3. If $\delta < \frac{m}{n}$, note that given the formula for r_k in Proposition 3, $r_k = (1 - \frac{n}{m}\delta) r - r_{m-1}$, for $k \leq m - 1$, and is increasing in k otherwise, with $r_n = +1$. Again, de…ne p_k to be the probability that the number of high-cost regions is less than or equal to k. Then the ex ante probability of separation in any region is :

$$s_P = \begin{cases} 1, & 0 \leq \hat{r} < r_{m-1} \\ 1 - p_k, & r_k \leq \hat{r} < r_{k+1},\ k \geq m - 1 \end{cases} \tag{5.1}$$

The explanation is as follows. If $\hat{r} < r_{m-1}$, then $\hat{r} < r_k$ for all k, so the incumbent always separates. If $r_k \leq \hat{r} < r_{k+1}$, then the incumbent only separates if the number of high-cost regions is greater than k, which occurs with probability $1 - p_k$.

How does s_P compare to s_D? Again, it is convenient to use Figure 1 above to illustrate this. The separation probability s_P as a function of δ, is superimposed on Figure 1 to give Figure 3.

[insert …g 3 here]

Here, for clarity, we have assumed n = 3. When δ is low, i.e. below $\frac{2}{3}$ ¡ $\frac{\hat{}}{r}$, separation always occurs. When δ is high, i.e. above $\frac{2}{3}$ ¡ $\frac{\hat{}}{3r}$, separation does not occur unless all regions are high-cost, which occurs with probability q^3. In between these two values of δ, separation only occurs with partial information if at least two regions are high-cost, an event which occurs with probability q^3 + $3q^2(1$ ¡ $q)$. In that case, it is possible that q^3 + $3q^2(1$ ¡ $q) > q$: for example, if q = $\frac{3}{4}$, q^3 + $3q^2(1$ ¡ $q)$ = $\frac{9}{8}\frac{3}{4}$.

Thus, separation can be more likely with centralization even with the possibility of selective pooling. On the other hand, as we will see below, separation is unambiguously less likely under partial information than it is under full information.

The intuition is simply that there are now two determinants of s_P. As in the case of s_F, the cost distribution e¤ect is still at work. But now, overlaid on this e¤ect is the selective pooling e¤ect which implies that $s_P < s_F$. But. the cost distribution e¤ect can still dominate, implying that we can get $s_P > s_D$, as above.

5.3. Voter Welfare

The only di¤erences in the analysis of voter welfare - relative to the full-information case - is that the incentive gain from pooling is di¤erent. Speci…cally,

$$EW_P = (1 + \delta)EW + (1 \text{ ¡ } \pi)(1 \text{ ¡ } s_P)(\frac{m}{n}¢_I \text{ ¡ } \delta\pi¢_S) \tag{5.2}$$

So, the expected gain in …rst-period voter welfare from pooling is now $\frac{m}{n}¢_I$ rather than $¢_I$. This is because of selective pooling: if the bad incumbent wins the election, he will do so only by pooling in m of the n regions - the others will still be exploited, getting a payo¤ of ¡r. So, only m of the n regions can expect a gain $¢_I$: as these regions are selected randomly in equilibrium, ex ante, a voter in any given region can expect an incentive gain of $\frac{m}{n}¢_I$.

So, it is straightforward to calculate, using (4.2), (5.2) that

$$EW_P \text{ ¡ } EW_D = (s_P \text{ ¡ } s_D)(1 \text{ ¡ } \pi)(\delta\pi¢_S \text{ ¡ } ¢_I) \text{ ¡ } (1 \text{ ¡ } s_P)(1 \text{ ¡ } \pi)(1 \text{ ¡ } \frac{m}{n})¢_I \tag{5.3}$$

As is clear from (5.3), there are now two e¤ects on welfare of moving to centralization:

1. A change in the separation probability:

2. A reduction in welfare at a given separation probability, because limits on rent-diversion are only needed in a majority of regions (instead of all regions) to be reelected - the selective pooling e¤ect,

In general, these two e¤ects could go either way. However, we have:

Proposition 4. If $\delta > \max\{\frac{¢_I}{\pi ¢_S}, \frac{m}{n}\}$, then $s_P = 0$, and $EW_D > EW_P$. If $\delta \cdot \max\{\frac{¢_I}{\pi ¢_S}, \frac{m}{n}\}$, then examples can be found when $EW_D < EW_P$.

Proof. (i) If $\delta > \max\{\frac{¢_I}{\pi ¢_S}, \frac{m}{n}\}$, then $\delta > \frac{m}{n}$, so from Proposition 3, $s_P = 0$. Also, $\delta\pi > \frac{¢_I}{¢_S}$, so $¡ ¢_I + \delta\pi ¢_S > 0$. But then as $s_D \, , \, s_P$, the result follows from (5.3).

(ii) See Example 1 below. ¤

The intuition for the general result is as follows. When δ is high, voters prefer a higher separation probability, because the bene…ts of separation come in the second period. But when δ is high, the incentive to pool with centralization is very strong, as the policy-maker only need sacri…ce rent-extraction in $\frac{m}{n}$ of the regions to be elected in the …rst period, thus gaining second-period rents in all regions. So, voters are worse o¤ with centralization both because the separation probability is lower, and because they prefer decentralization at a given separation probability, due to the selective pooling e¤ect.

To generate an example where centralization is preferred, a necessary condition is that voters dislike separation (i.e. $\delta\pi$ is low enough). But that is not su¢cient: we require also that the gain from greater pooling under centralization o¤sets the loss from selective pooling e¤ect. But this is possible if δ is low enough, as the following example shows.

Example 1. Let $n = 3$, $\delta \cdot \min\left\{\frac{2}{3\pi}\frac{¢_I}{¢_S}, \frac{1}{2}\right\}$. Then, (5.1) gives the relevant separation probabilities. Assume $\hat{r}$ is such that $(1 ¡ \frac{3}{2}\delta)r \cdot \hat{r} < (1 ¡ \delta)r$. Then, $s_D = 1$, $s_P = q^3 + 3q^2(1 ¡ q)$. Further, let $q = 0.5$. Then, $s_P = \frac{1}{2}$. Then from (5.3),

$$\begin{aligned} EW_P ¡ EW_D &= (1 ¡ \pi)\left[¡\frac{1}{2}[¡ ¢_I + \delta\pi ¢_S] ¡ \frac{1}{2}\frac{1}{3}¢_I\right] \\ &= (1 ¡ \pi)\left[\frac{1}{3}¢_I ¡ \frac{1}{2}\delta\pi ¢_S\right] \\ &> 0 \text{ as } \delta \cdot \frac{2}{3\pi}\frac{¢_I}{¢_S}. \end{aligned}$$

which is the required result. ¤

5.4. Comparing Partial and Full Voter Information

We are now in a position to ask what the e¤ects on separation probabilities and voter welfare are of switching from partial to full voter information, given centralization. In

an incomplete contracting framework such as this, one should not presume that more information is better, and indeed that is not the case. However, it is possible to establish that conditional on a …xed separation probability, voter prefer full information. Our results here are:

Proposition 5. (i) A change from partial to full voter information always increases separation probabilities (ii) A change from partial to full voter information always increases voter welfare, conditional on a …xed s : (iii) A change from partial to full voter information will always increase voter welfare unconditionally if $\delta\pi > \frac{¢_I}{¢_S}$, but if $\delta\pi < \frac{¢_I}{¢_S}$, examples can be found where a move from partial to full voter information will decrease voter welfare.

The reason of part (i) of Proposition 5 is that pooling is more pro…table with partial information as the incumbent has to restrain rent diversion in only a majority of regions (instead of all regions). As already remarked, part (iii) of Proposition 5 is comparable to Proposition 5 of Besley and Smart(2003). In the context of decentralized …scal policy, they demonstrate that yardstick competition between regions (which can only occur when voters are fully informed in our sense) does not necessarily increase voter welfare relative to no yardstick competition. However, in our case, the mechanism at work is quite di¤erent.

6. Policy Uniformity

6.1. Equilibrium and Voter Welfare with Policy Uniformity

Now consider the policy uniformity scenario i.e. partial voter information, but with the constitutional constraint that $g_{it} = g_t,\ \tau_{it} = \tau_t,\ t = 1, 2$. The implementation of this constraint is discussed below. Conveniently, the second-period behavior of an incumbent of a bad type (to set $g = 0, \tau = r$ in every region) satis…es policy uniformity. So, voter welfare is $-r$ in period 2, as before. If there are k high-cost regions, the good type will choose $g_2 = \overline{g}_k$ where $\overline{g}_k$ equates the average of the marginal bene…ts across regions from uniform provision to the average marginal cost i.e.

$$H'(\overline{g}_k) = \frac{kc_H + (n - k)c_L}{n} = \overline{c}_k$$

The uniform tax will be chosen just to …nance this expenditure i.e. if there are k high-cost regions, the benevolent incumbent sets $\overline{\tau}_k = \overline{c}_k\overline{g}_k$. Then, expected voter welfare in period 2 from a good incumbents is now

$$EW_G^U = \sum_{k=0}^{\mathbf{X}} q_k \left(H(\overline{g}_k) - \overline{c}_k\overline{g}_k\right)$$

where $q_k = p_k - p_{k-1}$,using the above definition of p_k, is the probability of exactly k high-cost regions, and is given by the usual Binomial formula.

Moving to the first period, assume for the moment that the incumbent is re-elected if and only if he chooses some $(g_1, \tau_1) \in \{\overline{g}_k, \overline{\tau}_k\}_{k=0}^n$. Then, the benevolent incumbent will choose as in the second period. To analyze the first-period behavior of the non-benevolent incumbent, note first that the overall rent from choosing $(g_1, \tau_1) = (\overline{g}_l, \overline{\tau}_l)$ when there are k high-cost regions is the present value of second period rent δnr, as he is re-elected, plus first-period rent

$$\Phi_{k,l} = \overline{c}_l \overline{g}_l - \overline{c}_k \overline{g}_l = \frac{(l-k)(c_H - c_L)}{n} \overline{g}_l$$

So, if the bad incumbent decide to pool, he will do so by choosing l as high as possible i.e. $l = n$, in which case his first-period rent from pooling is, using the fact that $\overline{g}_n = g_H$

$$\Phi_{n,l} = \frac{(n-k)(c_H - c_L)}{n} g_H$$

So, pooling will take place if $\Phi_{n,l} + \delta r \geq r$: using Proposition 1, it is easily seen that the incentives to pool are exactly the same as with full information. This means that the separation probability s_U is exactly the same as with full information i.e. is also given by (4.1). It remains to verify that a voter in i will re-elect the incumbent even if he observes $g_1 = g_H$, $\tau_1 = \tau_H$. But an argument identical to that in the proof of Proposition 1 indicates that this requires $q \geq (\frac{1}{2})^{1/n}$.

Now we turn to calculate voter welfare under policy uniformity. This is not equal - and is generally lower - than welfare under full information, as the good incumbent's flexibility to tailor local public good provision to local costs is constrained. In particular, we can write

$$\begin{aligned} EW_G^U &= EW_G - \Gamma, \\ \Gamma &= \sum_{k=0} q_k \{H(\overline{g}_k) - \overline{c}_k \overline{g}_k\} - q\{H(g_H) - c_H g_H\} - (1-q)\{H(g_L) - c_L g_L\} > 0 \end{aligned}$$

where Γ measures the benefit of flexibility i.e. the ability of the good incumbent to tailor puli good provision to local cost conditions.

We now turn to welfare analysis. Now, the formula (4.3) applies, except that EW_G is replaced by $EW_G - \Gamma$. After straightforward simplification, using $s_F = s_U$, we get:

$$EW_U = (1+\delta)EW + (1-\pi)(1-s_U)(\Phi_I - \delta\pi\Phi_S) = EW_F - \pi[1 + \delta(\pi + (1-\pi)s_U)]\Gamma \quad (6.1)$$

So, even though $s_F = s_U$, welfare under policy uniformity is lower than full information, as there is loss of flexibility for the good incumbent, captured by the last term in (6.1). So,

to summarize, we have the important insight that as far as control of the bad incumbent is concerned, policy uniformity is a perfect substitute for full information for voters, but it also has negative consequences as far as the good incumbent is concerned.

One might then ask why voters do not agree ex ante to implement full information rather than impose policy uniformity. Indeed, one could also ask the more basic question of how policy uniformity can even be implemented without voters having full information. The answer is simply that implementing policy uniformity can be done using a fraction of the resources than may be required to keep the voters informed. For example, several countries have a commission or branch of the judiciary that is responsible for overseeing the constitution. If such a body has veto power over …scal policy, and is not susceptible to bribery, and the constitution requires uniformity, then it will approve a …scal policy proposal i¤ in any period $g_{it} = g_t,\ \tau_{it} = \tau_t$. All that is required here is that the court be able to observe $g_{it},\ \tau_{it}$ not the citizens.

6.2. Comparing Di¤erentiated and Uniform Policy

In our setting, the only case where precommitment to policy uniformity may be advantageous is when voters only have partial information at the voting stage i.e. they only observe expenditure and the tax rate in their own region. In our model, a shift from differentiated to uniform policy will have three e¤ects on voter welfare. First, it will reduce the ‡exibility of good incumbents, as just argued. Second, at a given separation probability, it will increase voter welfare as it constrains the ability of the incumbent to extract rents without being detected. Third, it will raise the equilibrium separation probability, as - for the same reason - pooling is less pro…table for the incumbent. These three e¤ects are illustrated in the following decomposition:

$$
\begin{aligned}
EW_U - EW_P &= EW_U - EW_F + EW_F - EW_P \qquad (6.2)\\
&= \underbrace{-\pi[1+\delta(\pi+(1-\pi)s_U)]\mathbb{C}}_{\text{loss of ‡exibility (-)}} + \underbrace{(1-s_U)(1-\pi)(1-\frac{m}{n})\mathbb{C}_I}_{\text{improved incentives (+)}}\\
&\quad + \underbrace{(1-\pi)(s_U-s_P)(\delta\pi\mathbb{C}_S-\frac{m}{n}\mathbb{C}_I)}_{\text{increased separation (?)}}
\end{aligned}
$$

As indicated, the …rst e¤ect is negative, the second positive, and the third ambiguous , even though $s_U > s_P$ - it depends on the size of selection and incentive e¤ects.

So, we can make the following observations. First, obviously but importantly, if there is no political agency problem (i.e. the incumbent is always good, $\pi = 1$), policy uni-

formity is always bad. This is of course the puzzle raised by Oates' original assumption. Second, if there is a political agency problem, then under some conditions, a constitutional commitment to policy uniformity can be good. Explicit conditions on parameters under which this is the case can easily be calculated (see also below for the case of a uniform tax only). For example, assume for simplicity that $\delta\pi \¢_S = \frac{m}{n} \¢_I$, so that there is no separation e¤ect, and assume that π is small: then the incentive e¤ect will dominate the loss of ‡exibility e¤ect.

6.3. Uniform Taxes Only

As remarked in the introduction, empirically, it is much more common for taxes set by central government to be uniform across regions than it is for expenditures to be uniform. Our model can o¤er an explanation of this stylized fact - in fact, several explanations. The …rst, which is quite plausible, is that taxes in other regions are observable by voters, but expenditures are not. This is plausible, as taxes -.e.g. state income taxes in the US - are simple in structure and are in the public domain (e.g. on the web at /www.taxadmin.org/fta/rate/!), whereas data on expenditures, much less physical quantities of goods supplied, are typically multi-dimensional and not in the public domain.

In this case, an argument much like in the previous section shows that when $q \leq \frac{1}{2} \¢_n$, there is an equilibrium where the bad incumbent pools by choosing τ_H in all regions (he cannot discriminate now by choosing $\tau = r$ in some regions, as this will be observed), and sets $g = g_H$ in a majority of regions, and $g = 0$ in a minority. Conditional on a …xed number of high-cost regions, the gains to pooling are greater than with full voter information, but less than with partial voter information, which implies that the separation probability, s_T, is higher than s_P but lower than s_U, weakly in both cases. (These results are proved in an earlier version of the paper, Hindricks and Lockwood(2005)). Now suppose that voters only know about taxes in their own regions, but that there is a constitutional constraint for tax uniformity. Then, the equilibrium is exactly the same as with full tax information.

So, in either case, the gain from uniform taxation relative to fully di¤erentiated policy is

$$EW_T - EW_P = \underbrace{(1-s_T)(1-\pi)(1-\frac{m}{n})(\¢_I - H(g_H))}_{\text{improved incentives (+)}} + \underbrace{(1-\pi)(s_T - s_P)(\delta\pi \¢_S - \frac{m}{n}\¢_I)}_{\text{increased separation (?)}} \tag{6.3}$$

Comparing (6.2) and (6.3) indicates three di¤erences between the gain to tax uniformity and full uniformity. First, with tax uniformity, there is no loss of ‡exibility in public good provision. Second, the incentive e¤ect is weaker ($¢_I$ ¡ $H(g_H)$ = r ¡ $\tau_H > 0$ rather than $¢_I$) because the bad incumbent still has the option of supplying zero public good in regions where he does not need political support. Third, the separation probability is di¤erent.

Overall, these e¤ects are not unambiguously positive, but given s_T ¸ s_P, if voters like increased separation i.e. $\delta\pi¢_S$ ¸ $\frac{m}{n}¢_I$, then the gain to uniform taxation - or full information about taxation - is positive from (6.3) . If, moreover, π is high, the gain to uniform taxation and expenditure is negative from ((6.2). So, our model can explain the "stylized fact" that tax uniformity is much more common than expenditure uniformity.

7. Some Extensions

7.1. Spatial Cost Correlation

So far, we have assumed that $c_t^1, c_t^2 ... c_t^n$ are uncorrelated. What happens if they are (positively) correlated? This is an interesting case as it is very likely in practice to be some common aspect to any shocks hitting the cost of regional public good provision. One example would be national-level bargaining for workers employed by local government, which certainly occurs in the UK. Another would be a national regulatory change in e.g. health and safety regulations in the workplace.

There are two implications of spatial correlation. First, if voters have full information, the decentralization outcome will be di¤erent, as now there will be yardstick competition, as pointed out by Besley and Smart(2003). Second, the cost distribution e¤ect is weakened. This simplest way to see this is to assume perfect spatial correlation. Then, the cost distribution e¤ect disappears altogether. That is, under either decentralization or centralization, the proportion of regions under the incumbent's jurisdiction is either 0 or 1.

We have a formal analysis of how these changes a¤ect equilibrium separation probabilities and voter welfare, under the simplifying assumption[23] of perfect spatial correlation. For reasons of space, we just report the main …ndings, but the details are available on request. Two …ndings emerge which are robust in that they hold, no matter what the

[23] Besley and Smart(2003) provide a analysis of yardstick competition with less than perfect correlation under the special assumptions of two regions, and q = 0.5. An analysis of the general case (q ¸ 0.5, n odd, arbitrary positive correlation) would be very di¢ cult and in any case beyond the scope of this paper.

information of voters. First, the ambiguity in relative separation probabilities disappears: that is, separation probabilities under decentralization are weakly higher than under centralization. The reason for this as follows. Under full voter information, if the two separation probabilities di¤er, the decentralization one is higher, because yardstick competition deters bad incumbents from pooling. Under partial voter information, if the two separation probabilities di¤er, the centralization one is lower because the possibility of selective rent-diversion encourages pooling. The second …nding is that when separation probabilities under decentralization are strictly higher than under centralization, centralization is preferred by voters i¤ the incentive e¤ects are su¢ciently large relative to the selection e¤ects.

7.2. Politicians who Di¤er in Competence

Here, we brie‡y sketch results for the competence model. The purpose of this section is simply to emphasize that for the most part, the results already obtained carry over to the competence case. Consider …rst decentralization. In the second period, both good and bad incumbents will extract maximum rent r. Having done that, they wish to supply the good as close as possible to the e¢cient level in each region, whatever the cost. We will assume that $g_k > (1 - r)/c_k = \bar{g}_k$ i.e. having extracted maximum rent, the maximum possible public good supply -given the maximum tax rate of unity - is less than the e¢cient level. So, an incumbent with cost c_k supplies $\bar{g}_k = (1 - r)/c_k$, $k = H, L$. So, all voters prefer a competent to an incompetent incumbent, since public good supply is higher.

Now consider the …rst period. Assume for the moment that good incumbent in either region will be elected with probability 1 if he behaves non-strategically i.e. makes exactly the same policy choices as in the second period. We will shortly verify when this is equilibrium behavior for the voters. In this case, the best strategy for the good incumbent is to behave non-strategically.

Now consider the bad incumbent in i. Note that the bad incumbent can always be re-elected by imitate the tax and expenditure of the good incumbent i.e. set $(g_i, \tau_i) = (\bar{g}_L, 1)$, because given this observation, the voters'posterior probability that the incumbent is good is at least π, from the fact that the good incumbent always sets $(\bar{g}_L, 1)$. So, if his cost is low ($c_i = c_L$) the bad incumbent cannot do better than imitate (pool with) the good incumbent, because if he imitates, he will be re-elected while extracting maximum rent. If his cost is high, ($c_i = c_H$) he has two possible options. The …rst is to which leaves him

with reduced rent of $\hat{\alpha} < r$, where

$$\frac{(1 - \hat{\alpha})}{c_H} = \frac{1 - r}{c_L} = g_L$$

giving him payoff $\hat{\alpha} + \delta r$. The second is to separate by taking maximum rent and thus setting $(g_i, \tau_i) = (g_H, 1)$, thus losing the election and giving him payoff $r + \delta.0$. So, will pool if $\hat{\alpha} > (1 - \delta)r$, and separate otherwise.

So, in any region, the ex ante probability that a bad incumbent separates (the selection effect) is

$$s_D = \begin{cases} 0 & \text{if } \hat{\alpha} > (1 - \delta)r \\ q & \text{if } \hat{\alpha} < (1 - \delta)r \end{cases}$$

Comparing with the benevolence model, the separation probability is lower in the competence model because when cost is high the incumbents do not separate if $\hat{\alpha} > (1 - \delta)r$; and when cost is low the incumbent never separates, no matter what $\hat{\alpha}$ is. In the benevolence model there is always separation when cost is high and separation is also possible when cost is low if $\hat{\alpha} < (1 - \delta)r$.

With centralization, again two different possibilities can be analyzed: full information and partial information. The qualitative results are very similar to the case of benevolence, with two key exceptions. So, to avoid excessive re-statement of very similar propositions (full details are available on request from the authors), we will just note these exceptions.

First, in both of the cases of full and partial information., unlike in the benevolence case, voter welfare is always decreasing in the separation probability. The reason for this is straightforward. Let $W_k = H(g_k) - 1$ be the payoff to a voter from an incumbent who behaves as if he has cost c_k, $k = L, H$. Then, in period 1, if the bad incumbent separates rather than pools in a region, he delivers W_H rather than W_L to the voters, so the current loss from separation is $W_L - W_H$. The expected future gain is that next period's incumbent will be high-cost with probability only $(1 - \pi)q$, rather than q, implying a gain in voter welfare of $\delta\pi q(W_L - W_H)$. So, it is not surprising that voters are worse off with separation as it involves a cost in the first-period that is certain against an equal-size benefit in the next-period that is uncertain and discounted. All the propositions above hold, appropriately restated, taking into account the differences just noted.

Second, on voter welfare. With full voter information, voter welfare is higher with whichever arrangement gives the lower separation probability: this is the analog of Proposition 2. With partial information, given the same separation probability, decentralization dominates centralization, due to selective pooling effect which is also present in the competence model with partial information. Nevertheless, examples can be found where

centralization dominates; these involve a lower separation probability with centralization. This is the analog of Proposition 4. Note in this case, the cost and bene…t of decentralization is particularly clear: it prevents selective pooling, to the bene…t of voters, but will usually increase the separation probability, to the cost of voters.

8. Related Literature and Conclusions

8.1. Related Literature

A …rst related literature is that on lobbying and decentralization (Bardhan and Mookherjee(2000), Bordignon, Colombo, and Galmarini (2003), Redoano(2003), Ruta(2006)). When lobbying is introduced to a model of local public good provision, there can be di¤erences in the outcome under …scal regimes even in the absence of inter-regional spillovers and preference heterogeneity. For example, Bardhan and Mookherjee(2000) show that even if districts are homogeneous is all respects, lobbies have more in‡uence over the outcome with national elections as opposed to local elections, and so policy will be less distorted away from the Utilitarian optimum. This is because in their model, there are district-level shocks to electoral outcomes which are averaged away at the national level, giving the lobby a greater incentive to donate to the dominant party. But, the remaining literature in his area uses one or both of the assumptions of inter-regional spillovers and preference heterogeneity, and so can be regarded as an extension of the standard theory, not an alternative to it.

A second related paper is Harstad(2007), who provides an explanation for endogenous policy uniformity based on asymmetric information about willingness to pay for public goods. A special case of his model has two regions bargaining over their distribution of a …xed budget, where each region uses its budget allocation to buy a public good, whose value is known only to that region. In this bargaining problem with asymmetric information, there is a unique Cho-Kreps stable equilibrium, which exhibits delay; a region with a low willingness to pay can signal its type by rejecting an o¤er and delaying a counter-o¤er. So, preference-matching is achieved only at the cost of some delay. On the other hand, if uniformity is imposed i.e. the budget has to be equally divided, there is trivially, immediate agreement. Under some conditions, the cost of delay outweighs the bene…ts of preference-matching, and then uniformity is best.

8.2. Conclusions

This paper has considered the e¤ects of …scal decentralization on both incentives and selection of policy-makers. The main message is that (except in the probably unrealistic case where voters have full information about local public good provision across the economy), bad incumbents can pool worth good ones at lower cost to themselves (but at a higher cost to voters) with centralization. This has two consequences. First, at a given separation probability, voter welfare is lower with centralization. But, equilibrium separation probabilities can be higher or lower with centralization: the forces at work on the separation probabilities are quite subtle.

Our model presents a …rst step toward addressing the question of the expected e¤ect of decentralization on government e¢ciency in a systematic fashion by studying jointly the incentive and selection e¤ects. One possible direction for future work is to study .empirically separation rates for policy-makers (e.g. legislators or governors) at the national and sub-national levels. While there is an existing political science literature on the determinants of job tenure of politicians (see e.g. Finnoccario and Lin(2000)), to our knowledge, there has been no investigation of whether expected tenure is signi…cantly di¤erent at di¤erent levels of government.

Acknowledgements. We are grateful to the participants of the conferences: "Political Economy and Institutional Performance" (Brussels, June 11-12, 2004), "Social Choice and Welfare Economics" (Malaga, June 2-4, 2005), PET 2005 (Marseilles, 16-18 June 2005) and seminar participants at Bristol, Essex, Warwick, Nottingham, Milan and Louvain for their comments. We would also like to thank Massimo Bordignon, Micael Castanheira, Kai Konrad, Nicolas Porteiro, Christian Schultz and Jean Tirole for helful comments on a previous draft.

References

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

[2] Alesina, A. , I.Angeloni and F.Etro (2005), "International Unions", American Economic Review, 95, 602-615.

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

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

[5] Baron, D. and D. Besanko (1992), "Information, Control and Organizational Structure", Journal of Economics and Management Strategy

[6] Banks, J.S. and R.K. Sundaram (1998), "Optimal Retention in Agency problems", Journal of Economic Theory 82, 293-323.

[7] Belle‡amme, P. and J. Hindriks (2005), "Yardstick Competition and Political Agency Problems", Social Choice and Welfare 24, 155-70.

[8] 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.

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

[10] Bordignon, Colombo, and Galmarini (2003), "Fiscal Federalism and Endogenous Lobbies' Formation", (CESIfo Working Paper 1017)

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

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

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

[14] Faguet, J-P, (2004), "Does Decentralization Increase Government Responsiveness to Local Needs? Evidence from Bolivia" Journal of Public Economics 88, 667-893

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

[16] Finnoccario, C.J. and T.Lin(2000), "The Hazards of Incumbency: An Event-History Analysis of Congressional Tenure", (unpublished paper, University of Michigan)

[17] Fissman, R., and R.Gatti (2002) "Decentralization and Corruption: Evidence Across Countries," Journal of Public Economics, 83(3), 325-45.

[18] Fissman,R. and R.Gatti(2002a), ”Decentralization and Corruption: Evidence from US Transfer Programs”, Public Choice, 113, 25-35

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

[20] Henderson, J.V, and A.Kuncoro (2004) ”Corruption in Indonesia”, (NBER Working Paper W10674)

[21] Hindriks, J and B. Lockwood (2005) ”Decentralization and Electoral Accountability: Incentives, Separation, and Voter Welfare”, (Warwick Working Paper in Economics, 717)

[22] Harstad, B (2007), "Harmonization and Side-Payments in Political Cooperation", American Economic Review, forthcoming

[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] Kotsogiannis,C. and R.Schwager (2004), "Policy Innovation in Federal Systems", unpublished paper

[25] La¤ont, J.-J., and J. Tirole (1998), ”The Dynamics of Incentive Contracts”, Econometrica, 56, 1153-75.

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

[27] Lockwood, B., (2005), ”A Note on the Hybrid Equilibrium in the Besley-Smart Model””, (Warwick Economics Working Paper 727).

[28] Lockwood, B. (2006) "A Note on Some Political Agency Models of Fiscal Decentralization", unpublished paper, University of Warwick

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

[30] Myerson, R. (1993), ”E¤ectiveness of Electoral Systems for Reducing Government Corruption: a Game Theoretic Analysis”, Games and Economic Behavior 5, 118-32.

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

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

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

[34] 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.

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

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

[37] Redoano, M. (2003) ""Does Centralisation A¤ect the Number and Size of Lobbies?", CSGR Discussion Paper 604, University of Warwick

[38] Rogo¤, K. (1990), Equilibrium political budget cycles, American Economic Review 80: 21-37.

[39] Ruta,M (2006), "Special Interests and the Gains from Political Integration", Economics and Politics, 18, 191-218

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

[41] Treisman, D, (2000), "The Causes of Corruption: A Cross-National Study", Journal of Public Economics, 76: 399-457.

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

[43] 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).

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

A. Proofs of Propositions

Proof of Proposition 1. Suppose for the moment that the voters are willing to re-elect the incumbent whenever he pools. Only two strategies can possibly be optimal for the incumbent. The …rst is to separate, and lose the election which gives him a payo¤ $nr + \delta.0$. The second is to pool and win the election. This gives payo¤ $(n ¡ k)\hat{r} + \delta nr$. The second strategy is better i¤ $\hat{r} ¸ \frac{n}{n ¡ k}(1 ¡ \delta)r = r_k$. As $r_n = 1$, it is always best to separate when $k = n$.

Now we verify that the voters are willing to re-elect the incumbent whenever he pools. By Bayes' rule, the voters' posterior belief that the incumbent is good when $(g_i, \tau_i) = (g_H, \tau_H), i = 1, ..n$ is

$$\pi^{0} = \frac{q^n \pi}{q^n \pi + (1 ¡ \pi)\sigma}$$

where σ is the probability that in equilibrium, a bad incumbent plays the pooling strategy. For the voters to be willing to re-elect the incumbent, we require $\pi^{0} ¸ \pi$, or $q^n ¸ \sigma$. As the incumbent always separates when $k = n$, which occurs with probability q^n, the maximum possible value of σ is $\sigma = 1 ¡ q^n$ (this occurs when the incumbent pools in all other cost states, which occurs when $\hat{r} ¸ n(1 ¡ \delta)r$). So, for the voters to be willing to re-elect in equilibrium, we require $q^n ¸ 1 ¡ q^n$, or $q ¸ (1/2)^{1/n}$. ¤

Proof of Proposition 3. If $k < m$, there is a majority of low cost regions and only two strategies can possibly be optimal for the incumbent. The …rst is to take maximum rent in all regions, thus separating in all regions and lose the election which gives him a payo¤ $nr + \delta.0$. The second is to limit his rent to $\hat{r}$ in m low-cost regions (thus pooling in those regions), and take maximum rent in all other regions. This is pooling overall, and thus the incumbent wins the election. This gives payo¤ $\hat{r}m + (n ¡ m)r + \delta nr$. The second strategy is better i¤ $\hat{r} ¸ (1 ¡ \frac{n}{m}\delta)r = \underline{r}$.

If $k ¸ m$, there is a majority of high cost regions and only two strategies can possibly be optimal for the incumbent. The …rst is to take maximum rent in all regions, thus separating in all regions and lose the election which gives him a payo¤ $nr + \delta.0$. The second is to limit his rent to $\hat{r}$ in all $n ¡ k$ low-cost regions (thus pooling in those regions), and limit his rent to zero in $l = m ¡ (n ¡ k)$ high-cost regions, and take maximum rent in all other regions. This is pooling overall, and thus the incumbent wins the election. This gives payo¤ $\hat{r}(n ¡ k) + l.0 + (n ¡ (n ¡ k) ¡ l)r + \delta nr = \hat{r}(n ¡ k) + (n ¡ m)r + \delta nr$. The second strategy is better i¤ $\hat{r} ¸ r_k$. ¤

Proof of Proposition 5. (i) We must show that separation probability is higher with full voter information. With full information the separation probability s^F is as follows. Let $r_k^F = \frac{n}{n ¡ k}(1 ¡ \delta)r; k · n$. If $\hat{r} < r_0^F$, the incumbent separates no matter what k is and $s^F(\hat{r}) = 1$.

If $r_k^F \cdot \hat{r} < r_{k+1}^F$, the incumbent separates with probability $s^F(\hat{r}) = 1 - F(k)$. With partial information the separation probability s^P is as follows. Let $r_k^P = \max\{(1 - \frac{n}{m}\delta)r, \frac{m}{n-k}(1-\delta)r\}$; with $k \cdot n$. If $\hat{r} < (1 - \frac{n}{m}\delta)r = \underline{r}$, the incumbent separates no matter what k is and $s^P(\hat{r}) = 1$. If $r_k^P \cdot \hat{r} < r_{k+1}^P$, the incumbent separates with probability $s^P(\hat{r}) = 1 - F(k)$. Thus separation probability is decreasing step function with the same downward jump of $F(k) - F(k-1)$ around r_k^P and r_k^F . Since for all $k \cdot n$, $r_k^F = \frac{n}{n-k}(1-\delta)r > r_k^P = \max\{(1 - \frac{n}{m}\delta)r, \frac{m}{n-k}(1-\delta)r\}$ then the step function $s^F(\hat{r})$ decreases less rapidly than the step function $s^P(\hat{r})$. Thus given that $s^F(0) = s^P(0) = 1$ we must have $s^F(\hat{r}) \geq s^P(\hat{r})$ for all $\hat{r}$, with strict inequality for $\hat{r} > \underline{r}$. Hence we conclude that separation probability is higher with full information.

(ii) Given s …xed, the fact that voter welfare is higher with full information follows from (??) and the fact that $EW^F(s) = EW^D(s)$.

(iii) Let $EW_F(s), EW_P(s)$ be voter expected welfares with full and partial information conditional on a …xed s. Then from (4.3), (??) above,

$$EW_F(s) - EW_P(s) = (1-s)(1-\pi)(1-\frac{m}{n})¢_I$$

i.e. conditional on a given s, voters prefer full information. So, if $\delta\pi > \frac{¢_I}{¢_S}$, the voters also prefer a higher separation probability and the result follows immediately.

If $\delta\pi < \frac{¢_I}{¢_S}$, an example where partial information is preferred to full information is the following. Assume $\hat{r}$ is such that $(1 - \frac{3}{2}\delta)r \cdot \hat{r} < (1-\delta)r$. Then, from the formulae (4.1), (5.1), we see that $s_F = 1$, $s_P = q^3 + 3q^2(1-q)$. Then, noting that at a given s, $EW_D = EW_F$, the example is exactly the same as in Example 1 above. ¤

Figure 1: Separation Probabilities,
Decentralisation

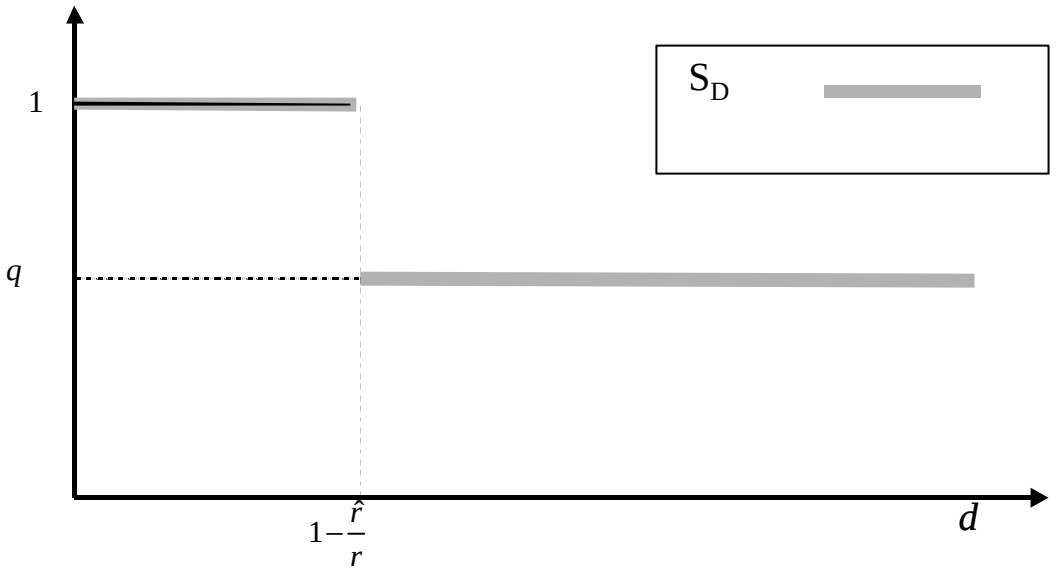

Figure A.1:

Figure 2: Separation Probabilities,
Centralization with Full Information

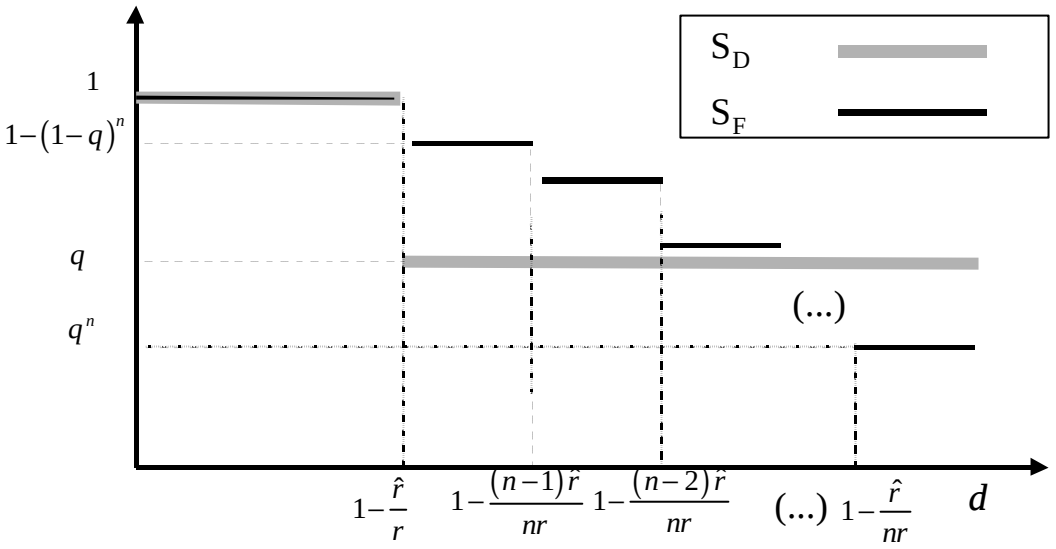

Figure A.2:

Figure 3: Separation Probabilities, Centralization with Partial Information

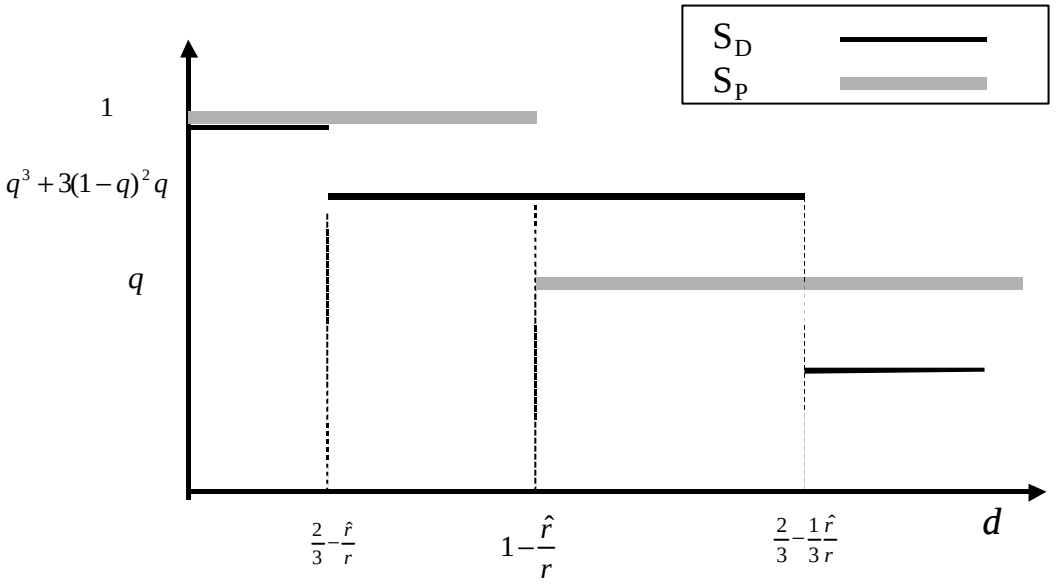

Figure A.3: