The Assignment of Powers in Federal and Unitary States*

Ben Lockwood†

First Version: January 1997
This version: June 2000

Abstract

This paper studies a model where the power to set policy (a choice of project) may be assigned to central or regional government via either a federal or unitary referendum (constitutional rule, CR). The benefit of central provision is an economy of scale, while the cost is political inefficiency. The relationship between federal and unitary CRs is characterized in the asymptotic case as the number of regions becomes large, under the assumption that the median project benefit in any region is a random draw from a fixed distribution, G. Under some symmetry assumptions, the relationship depends only on the shape of G; not on how willingnesses to pay are distributed within regions. The relationship to Cremer and Palfrey's (1996) "principle of aggregation" is established. Asymptotic results on the efficiency of the two CRs are also proved.

Keywords: decentralisation, referenda, federal and unitary states.
JEL Classification Numbers: D72, H70

*Address for Correspondence: Department of Economics, University of Warwick, Coventry CV4 7AL, UK, b.lockwood@warwick.ac.uk

†I would like to thank Martin Cripps for very valuable advice.

1. Introduction

The issue of assignment of tax and spending powers between different levels of government is receiving increasing attention amongst economists, perhaps because many countries are moving in the direction[1] of greater decentralization (Bird(1993)). All countries have constitutional rules or procedures, implicit or implicit, for choosing the level of decentralization of a tax or spending power. These rules differ significantly between federal and unitary states. In federal states, the allocation of powers is usually specified in the constitution and may require[2] a constitutional amendment. In all major federal states, rules for constitutional amendment require that at least a majority of sub-central governments must approve the amendment (Wheare(1963)). For example, in the US, any amendment must be approved by at least three-quarters of all state legislatures. In a unitary state, such as the UK, reallocation of powers is achieved either by legislation in the national parliament, or by national referendum[3]- the agreement of any sub-central level of government per se is not required.

While there is now a large and growing theoretical literature on decentralization, remarkably, there is only one[4] paper that addresses directly the different decision-making procedures of federal and unitary states, Cremer and Palfrey(1996). In their model, regional or central governments choose some value of a policy variable (a real number) via majority voting. With centralized choice, it is assumed that the value of the policy variable must be the same for all regions. Thus, the cost of centralization is policy uniformity. Moreover, they assume that voters are incompletely informed about the preferences of other voters, both in their regions and in other regions. It turns out in this set-up that the benefit of

[1] For example, Bird cites the "new federalism" in the US, and moves to federalism in Spain and Belgium. In the UK, recent referenda on devolution of powers to Scotland and Wales will result in the establishment of Scottish and Welsh parliaments.

[2] However, the degree to which reallocation of powers leads to constitutional amendment varies considerably across federal countries. In the US, there has only been one constitutional amendment for this purpose (in 1913, to allow a Federal income tax), whereas in Switzerland there have been a large number of amendments over the last 100 years, enhancing the tax powers of central government (Wheare(1963), Chapter 6).

[3] Again, if the reallocation of powers requires a constitutional amendment, a national referendum is sometimes required, e.g. in France (Curtis(1997)). In countries without a well-defined constitution, such as the UK, reallocation of powers requires only a majority in parliament.

[4] Cremer and Palfrey(1999) use the same model to study the implications of unit- and population-proportional representation.

centralization is policy moderation[5].

In this setting, Cremer and Palfrey study two referenda, which we call the unitary and federal two constitutional rules[6] respectively. Under the unitary constitutional rule (CR), a choice between centralization and decentralization is made by national referendum. Under the federal CR, every region chooses between centralization and decentralization using a regional referendum, and then the arrangement preferred by a majority of regions is chosen. These CRs capture in a simplified way the distinguishing features of federal and unitary states mentioned above. They obtain a remarkable result[7]: as the number of (equal-sized) regions become large, whenever the unitary CR selects centralization, the federal CR also selects centralization (but not necessarily vice versa), so federal CRs unambiguously lead to more centralization. They call this result the principle of aggregation.

This paper takes an alternative approach to the same question, and in doing so, addresses three limitations of Cremer and Palfrey's paper. The first limitation is that (as remarked above) policy uniformity with centralized decision-making is assumed, rather than derived endogenously as an equilibrium of the political process. Second, as they say themselves, they do not model any efficiency gains from centralization; in their setting policies are costless (or equivalently, equally costly). This means that decentralization is always the efficient choice, as it allows for policy diversity[8]. Third, due to the information structure, their model is only tractable if very specific assumptions on the distribution of preferences within regions and between regions are made[9], and indeed, Cremer and Palfrey assume

[5] That is, when the number of regions becomes large, the subjective probability for any particular voter that the policy variable will, in voting equilibrium, take on an extreme value (i.e. far from that voter's most preferred value) is lower with centralization.

[6] Here, following Buchanan(1975,1987,1988) we refer to all proceedures such as referenda, parliamentary votes, etc as constitutional rules.

[7] This follows from Figure 1 in their paper, where it is clear that if the proportion of voters preferring centralization is greater than 0.5, then the proportion of regions preferring centralization must also be greater than 0.5.

[8] More precisely, as is shown in the Appendix of this paper, in their model, the sum of utilities across all voters is always strictly greater with decentralization than with centralization.

[9] Specifically, voters do not observe the median voter's ideal point m_d in their region directly, but know the distribution G from which m_d is drawn, and their own ideal point, t; which is a draw from a distribution F with median m_d. From this information, they have to calculate the mean and variance of m_d conditional on t: These calculations are only really tractable if G is a natural conjugate prior for F:

for the most part[10] that both these distributions are Normal. Then question then arises as to whether their principle of aggregation is an artefact of the Normal distribution, or whether it holds more generally.

The …rst contribution of this paper is to present a political economy model[11] of the costs and bene…ts of centralization that addresses all these issues. We consider a model where voters have preferences over a private good and a discrete public good in their region (a "project"); these preferences are characterized by the bene…t the voter gets from the project. In this model, there are no constraints on the distribution of project bene…ts within regions, and between regions. There are economies of scale in joint production of projects. Following the …scal federalism literature, we assume that regional governments cannot cooperate to exploit these economies of scale. So, the bene…t of central provision is that a given number of projects can be undertaken more cheaply.

More importantly, the cost of centralization is modelled explicitly as political ine¢ciency, that is, ine¢ciency arising from the legislative process. With centralization, the process of decision-making about projects is modelled as a legislative game, following Baron and Ferejohn(1989), Baron (1991). Moreover, legislative bargaining is embedded in a citizen-candidate model where any citizen in a region can stand for election as regional delegate. The equilibrium of this model has the property that the set of projects that are funded is insensitive to the median project bene…t in the region, but is driven by cost considerations. In our simple model, every region's project is funded with the same probability i.e. there is endogenously derived policy uniformity.

This model of the costs and bene…ts of decentralization then allows us to study constitutional rules for allocating powers. In Section 3, we show that with a …xed and …nite number of regions, and no restrictions on the distribution of project bene…ts, either within or across regions, there is no particular reason to think that the federal CR will be systematically more centralized (or indeed decentralized) than the unitary rule.

In Section 4, we establish the main (asymptotic) results of the paper, under the assumptions, also made by Cremer and Palfrey, namely; (i) regional medians

[10] In Section 4.3 of their paper, they present a uniform distribution example where 100% of regions prefer centralization, in which case the principle of aggregation certainly holds, but this example is not general.

[11] See also Lockwood(1998), Besley and Coate(1998) for political economy models of decentralization, and Caillaud, Julien, and Picard(1996), Cremer and Palfrey(1996), Gilbert and Picard(1996), Klibano¤ and Poitevin(1996)) for models based on asymmetric information considerations.

project benefits are random draws from a fixed distribution; (ii) conditional on the regional median, the distribution of tastes within any region is the same; (iii) the number of regions is large.

We first have a key benchmark result. Say that the federal and unitary CRs are *asymptotically equivalent* if, in the limit as the number of regions becomes large, the federal CR will choose decentralization if and only if the unitary CR does. Then, under some symmetry assumptions on preferences, we show that the federal and unitary CRs are asymptotically equivalent if the distribution of median project benefits across regions is uniform. This result holds irrespective of how preferences are distributed within regions.

So, the uniform is obviously the borderline case. We then have two more results. First, if the distribution of median project benefits across regions is positively single-peaked (i.e. has a quasi-concave density) then the federal CR is asymptotically more centralized than the unitary CR - i.e. in the limit as the number of regions becomes large it chooses centralization whenever the unitary CR does. Second, if the distribution of median project benefits across regions medians is negatively single-peaked (i.e. has a quasi-convex density) then the federal CR is asymptotically less centralized than the unitary CR. The intuition for these results is that the federal CR is more sensitive to changes in the distribution of regional medians away from the uniform than is the unitary CR. For example, a mean-preserving spread of the distribution of regional medians may convert a uniform distribution into a negatively single-peaked one. In this case, the proportion of median voters preferring decentralization rises by more than the proportion of voters in total preferring decentralization.

These findings relate to Cremer and Palfrey's "principle of aggregation" as follows. The two cases analyzed in their model were when preferences were Normal. But the Normal distribution is single peaked, in which case our result is that the federal CR is more centralized, consistently with their principle of aggregation.

As argued above, our set-up also allows us to analyze the efficiency of CRs. Buchanan argues that while policy acts (acts flowing from CRs) may well be inefficient in particular cases, we should expect society as a whole to choose constitutional rules that are in some sense efficient. Constitutional decisions are long-run ones, and so the performance of any constitution should be evaluated from behind a Rawlsian veil of ignorance, where citizens are not sure about what their position in "society" will be. In Section 5, we study the efficiency of federal and unitary CRs in this sense.

In general, both CRs will be inefficient, for the usual reason that majority voting does not take account of intensity of preferences. However, in the asymptotic case, under the same assumptions as before, a number of results can be proved. Again, the benchmark case is where the distribution of median project benefits across regions is uniform. In this case, both federal and unitary CRs are fully efficient. Also, deviations of both rules from full efficiency can again be characterized when the distribution of median project benefits across regions is either positively or negatively single-peaked.

2. The Costs and Benefits of Decentralization

2.1. Preliminaries

We develop the simplest possible model for our purpose. There are an odd number $i = 1, \ldots n$ of regions, with equal populations of measure 1. The assumption of equal populations is made as is it is easy (and not very interesting) to generate differences between federal and unitary CRs when regional populations differ[12]. In each region there is a discrete project $x_i \in \{0, 1\} = X$. The payoff of a resident of region i is

$$u_i = b_i x_i + y_i \tag{2.1}$$

where b_i is a benefit parameter, and y_i is consumption of a numeraire private good. In region i, b_i is a continuously distributed random variable with median b_{mi}, support $\left[\underline{b}_i, \overline{b}_i\right]$, and distribution function F_i.

The project in region i may be provided by regional government i (decentralization), or by central government (centralization): In either case, the relevant government is assumed to finance the public good by levying a proportional income tax. Every citizen in region i has an endowment of 1 unit of the private good, and pays a proportional income tax t_i, so $y_i = 1 - t_i$.

2.2. Decentralization

In this case, the cost to any regional government of funding its project is c. So, the regional budget constraint is $t_i = c$ where t_i is the regional income tax rate. Substituting personal and regional budget constraints into the utility function

[12] For example, suppose that there are three regions, and no intra-regional variation in tastes, and that regions $i = 1, 2$ prefer decentralization. If region 3 has more than 50% of the total population, the federal CR will select D, and the unitary CR will select C.

(2.1), we get

$$u_i = x_i(b_i - c) + 1$$

Then, in region i, x_i is determined by majority voting over the space of alternatives, X. This implies that

$$x_i^d = \begin{cases} 1 & \text{if } b_{mi} \geq c \\ 0 & \text{otherwise} \end{cases}$$

So, the utility from decentralized provision for a citizen of i with benefit parameter b_i is

$$u^d(b_i) = \begin{cases} b_i - c + 1 & \text{if } b_{mi} \geq c \\ 1 & \text{otherwise} \end{cases} \tag{2.2}$$

2.3. Centralization

We assume that there are economies of scale with centralized provision i.e. central government can produce a vector of projects more cheaply than can regional governments, reflecting the assumption that regional governments cannot cooperate to exploit economies of scale in such activities as research and development.

We model economies of scale in the simplest possible way by supposing that the cost per project of financing to central government of providing some quantity of the good in every region is $c - k;\ k > 0$. A more sophisticated approach would allow the economies of scale to depend on the number of regions in which projects are provided, but this would not change the main results, while adding considerably to the complexity of the algebra.

Following the large literature on distributive politics, (see e.g. Persson (1998)), we assume that it is a constitutional restriction that the cost of public good provision is financed out of a proportional income tax levied nationally at rate t. So, the national budget constraint is

$$nt = \mu(x)(c - k)$$

where $\mu(x) = \#\{i \in N \mid x_i = 1\}$. Substituting personal and national budget constraints into the utility function (2.1), we see that the utility from project vector $x = (x_1, \ldots x_n)$ for a type-b_i individual in region i is

$$u_i(x; b_i) = b_i x_i - \frac{\mu(x)}{n}(c - k) + 1 \tag{2.3}$$

So, in choosing x, the central government faces a problem of distributive, or "pork barrel" politics: expenditures are specific to particular regions, whereas the tax is national. The simplest form of social choice in this case would be to have a national referendum over pairs of alternatives in X^n. The problem with this procedure is that it is well-known that in this setting, there is no policy x^* which is a Condorcet winner (Ferejohn, Fiorina and McKelvey(1987)), and so a voting cycle would emerge.

Several resolutions of this problem have been suggested, by placing some structure on the agenda-setting and voting rules that a legislature may use. One of the most influential is the legislative bargaining model of Baron and Ferejohn(1981), whose model we adapt for use here. As our model is one where the identities of the legislators are not exogenously given, we must also specify a procedure by which legislators are selected from regions, and here, we make use of the citizen-candidate model of Besley and Coates (1997) and Osborne and Slivinski (1996). The order of events is as follows.

1. Election of Delegates. (i) any citizen in a region can stand for election (at some positive cost, $\sigma > 0$); (ii) those citizens who stand are voted on; (iii) the winner is selected by the plurality rule[13] and is that region's delegate in the national legislature; (iv) if no delegate stands for election, the region is unrepresented in the legislative process.

2. The Legislative Process. Suppose that $k \leq n$ delegates are elected. In the first session, each delegate is selected with probability $\frac{1}{k}$ to make a proposal: A proposal is a vector $a_i \in X^n$ of projects to be funded. It is then voted on. If it is accepted by a strict majority of delegates, it is implemented, but if it is not accepted by a strict majority, then the legislature continues to the next session in which a member is selected to make another proposal and so on. Sessions take time, and delegates have a per-session discount factor of $\delta < 1$.

A political equilibrium is (i) a subgame perfect equilibrium to the legislative game, conditional on the set of delegates; (i) a voting equilibrium in each region, conditional on the set of candidates in that region, and the delegates elected by other regions; (iii) a candidate set for each region, where in every region, every candidate who stands for election does so only if the benefit of doing so is at least σ: A more formal description of each of these three stages, plus a proof of the proposition below[14], is given in Appendix A.

[13] If k candidates get equal numbers of votes, then each candidate is selected with equal probability $1/k$:

[14] All subsequent propositions are proved in Appendix B if a proof is required.

Proposition 1. There is a political equilibrium where (i) exactly one resident of region i with $b_i \geq n\sigma/m$ stands for election; (iii) this resident is unanimously elected as the delegate from region i; (ii) when selected as proposer, the delegate from region i proposes an a_i consisting of a project for region i and $m - 1$ regions in $N/\{i\}$ selected at random; (iii) the first proposal is accepted by the legislature.

Then from (2.3) and Proposition 1, expected payoff to any citizen of region i in this equilibrium is

$$\begin{aligned} u^c(b_i) &= \frac{m}{n}\left[b_i - \frac{m}{n}(c-k)\right] + \left(1 - \frac{m}{n}\right)\left[-\frac{m}{n}(c-k)\right] + 1 \qquad (2.4) \\ &= \frac{m}{n}[b_i - c + k] + 1 \end{aligned}$$

The first term on the right-hand side is the expected payoff in the event that region i's legislator is either selected as proposer, or is randomly selected to be "bribed" to vote for the proposal. The second term in the expected payoff otherwise.

One way to interpret (2.4) is that under centralization[15], any region gets a project with probability $\frac{m}{n}$, regardless of the region's willingness to pay for the project, as measured by b_i. We call this ex ante uniformity; prospects are only identical for every region before (every round of) legislative bargaining begins. So, while it is traditional to assume policy uniformity under centralization, we have generated policy uniformity endogenously, as the outcome of a political equilibrium[16]. It is probably fair to say that Oates(1972) had in mind ex post uniformity with centralization. However, from the point of view of constitutional design, as long as the choice of centralization or decentralization is prior to the legislative bargaining process (and this is a reasonable assumption), it does not matter whether the uniformity is ex ante or ex post.

[15] In other words, with centralization, provision of projects is entirely insensitive to regions' willingness to pay.

[16] The political equilibrium described above is not unique, nor are the equilibrium payoffs described in (2.4) unique, as there exist multiple equilibria to the legislative subgame. For example, with three regions, there is an asymmetric equilibrium where the delegate from 1 always makes a proposal to the delegate from 2, and vice versa, and where region 3 never gets a project, even if it elects a delegate, and consequently does not even both to elect a delegate. However, the equilibrium generating payoffs (2.4) seems an excellent candidate for a "focal" equilibrium, any region is chosen with equal probability to join the "minimum winning coalition" with the agenda-setter.

2.4. Pivotal Citizens

In region i, there may be a pivotal citizen with taste parameter $\hat{b}_i$ who is indifferent between C and D i.e. $u^c(\hat{b}_i) = u^d(\hat{b}_i)$: Writing this out in full using (2.2) and (2.4) and solving for $\hat{b}_i$, we get[17]

$$\hat{b}_i = \begin{cases} c + \frac{n+1}{n-1}k & \text{if } x_i = 1 \\ c - k & \text{if } x_i = 0 \end{cases} \tag{2.5}$$

The importance of the pivotal citizen is that she determines how all citizens in region i vote on a choice of C versus D; as described in the following Proposition:

Proposition 2. If $x_i^d = 1$, then all residents of region i with $b_i > \hat{b}_i$ strictly prefer D, and all residents of region i with $b_i < \hat{b}_i$ strictly prefer C. If $x_i^d = 0$, the reverse is true.

This is intuitive. As all residents in a region bear the same share of cost of provision, those who value the project more than (resp. less than) the pivotal citizen will prefer whichever arrangement gives the higher probability (lower probability) of the project taking place.

3. Constitutional Rules for the Assignment of Powers

As argued in the introduction, we view the choice of centralization or decentralization as generated by constitutional rules (CRs). Here, we specify two CRs that we wish to study.

The Unitary CR. centralization or decentralization is selected by national referendum.

The Federal CR. centralization or decentralization is selected by two-stage referendum. All citizens within a region vote on centralization or decentralization, and the alternative that has the support of the majority of regions is selected.

The unitary CR captures the idea that a vote in the national parliament, or national referenda, are used in unitary states to (re)assign powers. The federal CR is intended to capture the idea that (re) assignment of powers in a federal

[17] Of course, $\hat{b}_i$ may not lie in B_i i.e. there may be no citizen that is indifferent between C and D: So we extend the definition to these "corner" cases as follows. For example, when $x_i^d = 1$; then if $c + \frac{mk}{n-m} \leq b_i^-$, then we set $\hat{b}_i = b_i^-$; and if $c + \frac{mk}{n-m} \geq b_i^+$, then we set $\hat{b}_i = b_i^+$:

state usually requires the approval of at least a simple majority of the regions, and it is this that is being modelled[18].

We turn now to specify the conditions under which each CR selects centralization or decentralization. First, the federal CR. Note[19] from (2.5) and Proposition 2 that the median voter in region i strictly prefers centralization to decentralization iff

$$b_{mi} \in \left(c - k,\ c + \frac{n+1}{n-1}k\right) = B_C \tag{3.1}$$

Note that as the number of regions becomes large, the interval B_C becomes symmetric around c, with length approximately $2k$. To simplify the statement and proof of subsequent results, we assume in what follows that no median voter is indifferent between centralization and decentralization i.e. $b_{mi} \neq c - k,\ c + \frac{n+1}{n-1}k$.

Now with the federal CR, the region votes for the median voter's most preferred alternative. So, the above analysis implies that under the federal CR, the fraction of votes in favor of centralization is

$$\pi_F = \frac{\#\{i \in N \mid b_{mi} \in B_C\}}{n} \tag{3.2}$$

Then the federal CR selects centralization iff $\pi_F > 0.5$, and decentralization otherwise.

If the unitary CR is used, from Proposition 2, the fraction of votes in favor of centralization is

$$\pi_U = \frac{1}{n}\sum_{\{i \mid b_{mi} \geq c\}} F_i\left(c + \frac{n+1}{n-1}k\right) + \frac{1}{n}\sum_{\{i \mid b_{mi} < c\}} \left[1 - F_i(c - k)\right] \tag{3.3}$$

and the unitary CR selects centralization iff $\pi_U > 0.5$, and decentralization otherwise.

---

[18] An important caveat here is that in practice, rules for constitutional amendment are more complex than this (Wheare(1963)). For example, the approval of a super-majority of regions may be required, as in the US, where 3/4 of states must approve. Or, as in the case of Switzerland and Australia, a majority of voters, as well as regions, must apporove. These amendment rules are difficult to analyse, as they give a privileged position to the status quo: that is C (resp. D) is more likely to be selected by the rule if C (resp. D) is the initial position. Study of such rules is a topic for future work.

[19] The proof of this is simple. If $c \leq b_{mi} < c + \frac{n+1}{n-1}k = \hat{b}_i$, then clearly $x_i^d = 1$, and so from Proposition 2, the median voter strictly prefers C. Again, if $\hat{b}_i = c - k < b_{mi} < c$, then clearly $x_i^d = 0$, and so from Proposition 2, the median voter again strictly prefers C.

How do the two CRs compare? Generally, there will be a minority of voters in a region who disagree with the decision of the median voter of the region. We call the voters who disagree dissenting voters. It is clear that the way in which dissenting votes are distributed across regions determines whether or not the unitary CR conflicts with the federal CR. For example, if a majority of regional median voters prefer decentralization, but in those regions, there are large numbers of dissenting voters who prefer centralization, then the unitary CR may choose centralization. Of course, this argument works in reverse, so there is no presumption that the federal rule will choose decentralization more often, or indeed less often, than the unitary rule.

To understand how dissenting voters may determine the difference between the two rules, it is very helpful to start with benchmark conditions under which the rules are equivalent. Say that federal and unitary CRs are equivalent if the federal CR selects decentralization iff the unitary CR selects decentralization. Two simple sufficient conditions for equivalence are the following.

Proposition 3. If there is either (i) no intra-regional variation in tastes ($B_i \equiv \{b_{mi}\}, \forall i \in N$), or (i) no inter-regional variation in tastes ($b_i \equiv b_j, \forall\, i, j \in N$), or both, federal and unitary CRs are equivalent.

The intuition for this result is clear. First, condition (i) implies that there is no dissenting vote in any region. Condition (ii) implies that if (de)centralization is chosen by the federal CR, all regions must vote for this option. So, as at least 50% of the electorate in each region prefers the option, so must at least 50% of the electorate overall. So, any difference between the two CRs will appear when both intra-regional and inter-regional variances in tastes are present, as the following example shows.

Example 1

There are three regions, where regional medians are $b_{m1} = b_{m2} - \varepsilon$, $b_{m2} = c + \frac{k}{2}$, $b_{m3} = b_{m2} + \varepsilon$. Also, F_1 is uniform with support of length 2ϕ, and F_2, F_3 are uniform with support of length 2θ. So, θ, ϕ measure intra-regional variation in tastes, and ε measures inter-regional variation in tastes. $\|$

We then have the following fact, proved in the Appendix:

Fact 1. Assume $\phi = 0$ in Example 1. Then, when $\varepsilon < \frac{3}{2}k$, centralization is chosen by the federal CR, while decentralization is chosen otherwise. When $\theta \geq 2\varepsilon - 3k$, C is chosen by the unitary CR, while decentralization is chosen otherwise. $\|$

So, when $\phi = 0$, there exist parameter values where decentralization is chosen by the federal CR, and centralization by the unitary CR (but not vice versa), as shown in panel (a) of Figure 1. The intuition is as follows. When there is no intra-regional variance (µ = 0), federal and unitary CRs agree, as predicted by Proposition 3 above. Now suppose that " > 3k=2, so decentralization is chosen by both. As µ rises from zero, a dissenting votes in favor of centralization develop in both high-taste region 3, and low-taste region 1 (namely those residents with high b^0s in the low-taste regions, and low b^0s in the high-taste regions.) With a federal CR, these dissenters are ignored (the tyranny of the majority), but with a unitary CR, these voters' preferences count. When there are enough of these dissenting voters (when µ is high enough), the unitary CR chooses centralization when the federal CR chooses decentralization. (The disagreement is in the region RD; UC).

Figure 1 in here

However, we can also establish the opposite, using a di¤erent variant of Example 1.

Fact 2. Assume Á > 3k in Example 1. Then, as before, when " < $\frac{3}{2}$k, centralization is chosen by the federal CR, while decentralization is chosen otherwise. When µ < $\frac{3k ¡ 2"}{1 ¡ 3k=Á}$, centralization is chosen by the unitary CR, while decentralization is otherwise.k

So, in this variant of the example, there exist parameter values where centralization is chosen by the federal CR, and decentralization by the unitary CR. This is shown in panel (b) of Figure 1 above, where the federal and unitary decisions are compared. When there is no intra-regional variance (µ = 0), federal and unitary CRs agree, as predicted by Proposition 3 above. As µ rises, dissenting votes accumulate in regions 2 and 4, and so the unitary CR eventually chooses decentralization when the federal CR chooses centralization. (The disagreement is in the region RC; UD).

4. Asymptotic Results

Proposition 3, plus Example 1, indicates that without imposing some more structure on the problem, we are unlikely to be able to make general statements comparing unitary and federal rules. In this Section, we study the asymptotic behavior of the two rules as that when the number of regions is "large", under some symmetry assumptions on the distribution of preferences both with and across regions.

In this case, it turns out, somewhat surprisingly, that a comparison of the two rules can be based only on the distribution of regional median project benefits i.e. $\{b_{mi}\}_{i\in N}$. In particular, if this distribution is uniform (in the limit, as defined below), then the two rules are equivalent, no matter how the project benefits are distributed within regions. Starting from this benchmark, we can then develop simple conditions on the limiting distribution of regional median benefits for the federal rule to be either "more" or "less" centralized than the unitary rule. One of these results is a significant generalization of Cremer and Palfrey's "Principle of Aggregation" in the context of our model.

To conduct an asymptotic analysis as the number of regions becomes large, we assume the following structure: (i) regional median project benefits are random draws from a known distribution; (ii) conditional on the regional median, the distribution of tastes within any region is the same. Specifically, we assume:

A0. Every b_{mi} is a random draw from a common distribution G; where G is absolutely continuous with support $[\underline{b}_m, \overline{b}_m]$.

A1. The distribution of project benefit b in any region with median b_m; net of the median; i.e. $y = b - b_m$; is given by $F(y)$ on $[\underline{y}, \overline{y}]$; with $F(0) = 0.5$ by definition.

Now let π_F^n be the proportion of n regions that prefer centralization, given A0 and A1. From A0, for fixed n, π_F^n is a random variable. Moreover, from 3.1) above, region i chooses centralization iff $b_{mi} \in (c - k, c + \frac{n+1}{n-1}k) = B_n$. So, as $n \to \infty$; the probability limit of the proportion of regions choosing centralization is

$$\text{plim}_{n\to\infty}\pi_F^n = \text{plim}_{n\to\infty}\frac{\#\{i \in N \mid b_{mi} \in B_n\}}{n} = G(c+k) - G(c-k) = \pi_F(k) \qquad (4.1)$$

where $\pi_F(k)$ is strictly increasing in k. This is intuitive, the higher the cost saving from centralization, then ceteris paribus, the larger the fraction of regional median voters who will be in favor of centralization. For future reference, let the unique solution to $\pi_F(k) = 0.5$ be k_F. Then, the federal CR selects centralization iff the cost saving from centralization is sufficiently high i.e. $k > k_F$.

Now consider the unitary rule. In all regions with a median project benefit $b_m \geq c$, from Proposition 2, the proportion of residents who prefer centralization is $F(c + k - b_m)$; and in all regions with a median $b_m < c$, from Proposition 2, the proportion of residents who prefer centralization is $1 - F(c - k - b_m)$. Now let π_U^n be the proportion of citizens in n regions that prefer centralization, given A0 and A1. Again, from A0, for fixed n, π_U^n is a random variable. Its probability limit is

$$\operatorname{plim}_{n\to\infty}\pi_U^n = \int_c^{\bar{b}_m} F(c+k-b_m)g(b_m)db_m + \int_{\underline{b}_m}^{c}[1-F(c-k-b_m)]g(b_m)db_m = \pi_U(k) \tag{4.2}$$

Again, note that $\pi_U(k)$ is increasing in k. For future reference, let the unique solution to $\pi_U(k) = 0.5$ be k_U. Then, the federal CR selects centralization iff $k > k_U$.

We say that the federal CR is more centralized (decentralized) than the unitary CR if, when centralization is chosen by the unitary rule, it is also chosen by the federal rule (vice-versa). Using the above arguments, these two cases can be expressed as

$$k_U > k_F;\ k_U < k_F \tag{4.3}$$

For example, if the federal CR is more centralized, the cost advantage to centralization has to be higher (ceteris paribus) for centralization to be chosen under the unitary CR. Finally, the federal and unitary CRs are equivalent when $k_U = k_F$.

The main results of this section show that when certain symmetry assumptions are made about the distributions F, G, then whether the federal CR is more or less decentralized than the unitary CR depends only on the shape of G. These assumptions are the following:

A2. F, G are symmetric.

A3. In the limit, half the regions choose projects with decentralization i.e. $\operatorname{plim}_{n\to\infty}\frac{\#\{i\in N|b_{mi}\geq c\}}{n} = 0.5$.

These assumptions impose two forms of symmetry. A2 requires that the within-region project benefits and median benefits across regions are both symmetrically distributed. A3 ensures that the choices of regions under decentralization are symmetric.

Assumptions A2, A3 imply the following very useful simplifications. First, A3 plus A2 imply that b_m has a mean value of c. Therefore, it follows that $G(b_m) \equiv \Gamma(b_m - c)$, where Γ is a symmetric mean-zero distribution. So, from (4.1),

$$\pi_F(k) = \Gamma(k) - \Gamma(-k) = 2\Gamma(k) - 1 \tag{4.4}$$

where the second equality follows from the symmetry of Γ. Second, defining $x = b_m - c$, $\bar{x} = \bar{b}_m - c$, $\gamma(x) = \Gamma'(x)$, we get:

$$\pi_U(k) = \int_c^{\bar{b}_m} F(c+k-b_m)g(b_m)db_m + \int_{\underline{b}_m}^{c}[1-F(c-k-b_m)]g(b_m)db_m \tag{4.5}$$

$$
\begin{aligned}
&= \int_{c}^{\bar{b}_m} F(c+k-b_m)\gamma(b_m-c)db_m + \int_{\underline{b}_m}^{c}[1-F(c-k-b_m)]\gamma(b_m-c)db_m \\
&= \int_{0}^{\bar{x}} F(k-x)\gamma(x)dx + \int_{-\bar{x}}^{0}[1-F(-k-x)]\gamma(x)dx \\
&= \int_{0}^{\bar{x}} F(k-x)\gamma(x)dx + \int_{-\bar{x}}^{0} F(k+x)\gamma(x)dx \\
&= \int_{0}^{\bar{x}} F(k-x)\gamma(x)dx + \int_{0}^{\bar{x}} F(k-z)\gamma(-z)dz \\
&= 2\int_{0}^{\bar{x}} F(k-x)\gamma(x)dx
\end{aligned}
$$

Here, we have used the definition of Γ in the second line, a change of variables in the third, the symmetry of F (around zero, by definition) in the fourth, a change of variables in the second integral in the fifth, and finally the symmetry of γ around zero in the sixth.

We are now in a position to state and prove our first, benchmark, result.

Theorem 1. If A0-A3 hold, and in addition, the regional medians are uniformly distributed across regions i.e. Γ is uniform, then the federal and unitary CRs are equivalent i.e. $k_F = k_U$.

So, we see that in the borderline case is where the distribution of regional median project benefits is uniform, irrespective of how project benefits are distributed within regions. What happens when we move away from the uniform? Let H be any absolutely continuous distribution function with support $[a; b]$. Then we have the following definition:

Definition. H is strictly positively (negatively) single-peaked if the density $h(.)$ is strictly quasi-concave (quasi-convex) on $[a; b]$.

Note that if a density function is positively (negatively) single-peaked and is symmetric around zero, then it must have a global maximum (minimum) at zero. We now have,

Theorem 2. Assume that A0-A3 hold, that $|b - b_m| \leq \Gamma^{-1}(0.75)$ and in addition, Γ is strictly positively single-peaked. Then, the federal CR is more centralized than the unitary CR i.e. $k_F < k_U$.

The intuition is that the proportion of median voters in each region preferring centralization under the federal CR, π_F, is more responsive to changes in Γ away from the uniform distribution than the proportion of all voters preferring centralism under the unitary CR, π_U. In turn, this is because π_F does not take account of the views of dissenting voters. The following example may help clarify this argument. Suppose that Γ is initially uniform on $[-1, 1]$ and it is changed to a positively single-peaked distribution Γ^+ by moving some probability weight δ from the tails to the centre i.e. so that Γ^+ has a mass-point of δ at zero, and the remaining fraction $1 - \delta$ of regional means are distributed uniformly on $[-(1-\delta/2), 1-\delta/2]$. Then, as long as the median voters in the regions in the tails of the distribution initially preferred decentralization ($1 - \delta/2 > c + k$), a fraction δ more median voters will prefer centralization with Γ^+. So, π_F rises by δ.

Now consider a region i whose median voter was initially in the positive tail of the distribution i.e. where $b_{mi} \simeq 1$, assuming δ small. Initially, a fraction $F(c+k-b_{mi}) \simeq F(c+k-1) < 0.5$ of the voters in i already prefer centralization (the dissenting voters). After the switch, $F(k) > 0.5$ of the voters now prefer centralization. So, in this region, the net increase in the number of voters preferring centralization is $F(k) - F(c+k-1) < 1$. So, π_U rises by approximately $\delta[F(k) - F(c+k-1)] < \delta$ following the switch.

The same intuition explains our next result:

Theorem 3. *Assume that A0-A3 hold, that $|b - b_m| \le \bar{b}_m - c - \Gamma^{-1}(0.75)$ and in addition, Γ is strictly negatively single-peaked. Then, the federal CR is more decentralized than the unitary CR i.e. $k_F > k_U$.*

Note that the upper bound on the support of $b - b_m$ is strictly positive[20]. This strong characterization of the relationship between federal and unitary CRs has used the symmetry assumptions A2 and A3. We now present two examples which show that these assumptions cannot be relaxed.

Example 2. First, $x = b_m - c$ is distributed according to Γ, with density

$$\gamma(x) = \begin{cases} \frac{1}{2\gamma} & 0 \le x \le \gamma \\ \frac{1}{2\pi} & -\pi \le x < 0 \end{cases}$$

Also, F is given by the density[21]

$$f = \frac{1}{2\varepsilon}, \quad -\varepsilon \le y \le \varepsilon$$

Note that if $\gamma \neq \pi$, Γ is asymmetric, violating assumption A2, but that A3 is always satisfied. So, from (4.2), the proportion of regions preferring centralization

[20] $\bar{b}_m - c - \Gamma^{-1}(0.75) = \bar{x} - \Gamma^{-1}(0.75) > 0 \iff \Gamma(\bar{x}) > 0.75$, but by definition, $\Gamma(\bar{x}) = 1$.

[21] To ensure that $b \ge 0$, we need $c \ge \pi + \varepsilon$.

is

$$\pi_F(k) = \Gamma(k) - \Gamma(-k) = \frac{k}{2\gamma} + \frac{k}{2\pi} = \frac{k}{2}\left(\frac{\pi\gamma}{(\pi+\gamma)}\right)$$

So, as $\pi_F(k_F) = 0.5$,

$$k_F = \frac{\pi\gamma}{(\pi+\gamma)}$$

Now, from (4.3), and changing the variable of integration, see that the proportion of voters preferring centralization is:

$$\begin{aligned}
\pi_U(k) &= \int_0^{\gamma} \frac{k - x + \varepsilon}{2\varepsilon}\frac{1}{2\gamma}dx + \int_{-\pi}^{0}\left[1 - \left(\frac{-k - x + \varepsilon}{2\varepsilon}\right)\right]\frac{1}{2\gamma}dx \\
&= 0.5 + \frac{1}{8\varepsilon}\left[(2k - \gamma + 2\varepsilon) + (2k - \pi - 2\varepsilon)\right]
\end{aligned}$$

So, as $\pi_U(k_U) = 0.5$,

$$k_U = \frac{\gamma + \pi}{4}$$

So, the two CRs are only equivalent if $\frac{\gamma+\pi}{4} = \frac{\pi\gamma}{(\pi+\gamma)}$, which holds iff $\pi = \gamma$. So, as long as the distribution of Γ asymmetric i.e. $\pi \neq \gamma$, Theorem 1 no longer holds. ■

Now we present an example which shows that the benchmark result does not hold either when A3 does not hold.

Example 3. First, $x = b_m - c$ is distributed uniformly on $[-\gamma + \beta, \gamma + \beta]$ with $|\beta| < \gamma$ i.e. $\Gamma(x) = \frac{x+\gamma-\beta}{2\gamma}$. So, as long as $\beta \neq 0$, the mean of b_m is no longer c and consequently, A3 is violated. It is easily checked that $k_F = 0.5\gamma$. Now, also assume that F is distributed uniformly on $[-0.5, 0.5]$ i.e. $F(y) = y + 0.5$. Then

$$\begin{aligned}
\pi_U(k) &= \frac{1}{2\gamma}\left[\int_0^{\beta+\gamma} F(k - x)dx + \int_{\beta-\gamma}^{0}(1 - F(-k - x))dx\right] \\
&= \frac{1}{2\gamma}\left[\int_0^{\beta+\gamma}(0.5 + k - x)dx + \int_{\beta-\gamma}^{0}(0.5 + k + x)dx\right] \\
&= \frac{1}{2\gamma}\left[2\gamma(0.5 + k) - 0.5(\beta + \gamma)^2 - 0.5(\beta - \gamma)^2\right] \\
&= 0.5 + k - 0.5\beta^2/\gamma - 0.5\gamma
\end{aligned}$$

Now, as $\pi_U(k_U) = 0.5$, we see that

$$k_U = 0.5\beta^2/\gamma + 0.5\gamma$$

So, as long as A3 is not satisfied, i.e. $\beta \neq 0$, then $k_U > k_F$ i.e. the unitary CR is more decentralized than the federal CR, and consequently Theorem 1 fails. Interestingly, the unitary CR is more decentralized whether $E b_m$ is greater or less than c. ∎

Finally, we can state some comparative-statics results that describes how $k_F; k_U$ vary as the dispersion of preferences for project benefits increases, both across and within regions. We model an increase in the dispersion of median project benefits across (within) regions as a mean-preserving spread in the distribution of G (F): It is fairly obvious from (4.4) that (i) following a mean-preserving spread in the distribution of G, median voters in more regions will prefer decentralization, and so the cost saving from centralization at which half the median voters prefer centralization, namely k_F; will rise, and (ii) following a mean-preserving spread in the distribution of F, k_F is unchanged. The following theorem also establishes some less obvious results about what happens to k_U:

Theorem 4. If A0-A3 hold, then following a symmetric mean-preserving spread in G;both $k_F; k_U$ rise. If A0-A3 hold, then following a symmetric mean-preserving spread in F; (i) k_F is left unchanged; (ii) if, in addition, the hypotheses of Theorem 2 (Theorem 3) hold, k_U rises (falls).

These results establish that an increase in the dispersion of median project benefits across and within regions affects our two constitutional rules in quite different ways. An increase in the dispersion of median project benefits across regions makes both rules unambiguously more "likely" to choose decentralization, whereas an increase in the dispersion of project benefits within regions has an ambiguous effect on the unitary rule - it may make centralisation more likely.

5. Efficiency of Constitutions

We are now in a position to assess the relative efficiency of federal and unitary constitutions. As utility is linear in income, the model is one of transferable utility, and so the natural measure of efficiency is the aggregate surplus, or sum of utilities. If the aggregate surplus is greater under the federal CR, then the federal

CR is unambiguously potentially Pareto-preferred[22]. The aggregate surplus is

$$W^k = E \sum_{i=1} u^k(b_i)$$

in the case of centralization ($k = c$) and decentralization ($k = d$), and where the expectation is taken with respect to variables $(b_1, \ldots, b_n)$. An alternative way if justifying the use of aggregate surplus as a measure of efficiency is to suppose following Buchanan that choice between constitutions should be thought of as taking place behind a Rawlsian "veil of ignorance" (Dixit (1996)). If we suppose that the veil is complete i.e. every citizen, ex ante, believes that is equally likely that he will be resident in any region and if resident in region i, will have characteristics b_i drawn at random from the distribution F_i. In this case, the expected utility of the citizen behind the veil of ignorance can be calculated as W^k/n.

Now, using (2.4), (2.3), we see that the gain from a move to decentralization for a resident of region i with taste parameter b_i is:

$$u^D(b_i) - u^C(b_i) = \Delta_i = \phi(b_i; b_{mi}) - \frac{m}{n}(b_i - c) - \frac{m}{n}k$$

with

$$\phi(b_i; b_{mi}) = \begin{cases} b_i - c & \text{if } b_{mi} \geq c \\ 0 & \text{otherwise} \end{cases}$$

For simplicity, we assume from now on that the distribution of b_i is symmetric ($Eb_i = b_{mi}$). Then, taking expectations over the b_i, it is easy to show that

$$E\Delta_i = \left[\max\{b_{mi} - c, 0\} - \frac{m}{n}(b_{mi} - c) \right] - \frac{m}{n}k$$

That is, the expected gain from decentralization across all residents of region i is just the gain to decentralization for the median voter in that region. The first term of $E\Delta_i$ in the square brackets captures the gain of being able to respond more flexibly to regional preferences. The term $-mk/n$ is the loss implied by the inability to exploit economies of scale. So the gain to decentralization for a fixed number n of regions ins

$$\Delta^n = \sum_{i=1} E\Delta_i$$

---

[22] Of course, the Kaldor-Hicks criterion is only of interest here if lump-sum transfers between regions are possible at the point where the choice between centralization and decentralization is made.

So, the efficient CR selects decentralization iff $\Delta^n \geq 0$. Note that with b_i symmetrically distributed, the efficient CR is independent of the distribution of the taste parameter b_i within each region, as is the federal CR.

We now turn to discuss efficiency of federal and unitary CRs against this benchmark. Define a CR to be efficient if it makes the same selection of C or D as the efficient CR. Say that the federal rule is more efficient than the unitary rule, if whenever the unitary rule leads to an efficient decision, the federal rule does also, and vice versa. Also, say that a rule is inefficiently (de)centralized if when it makes an inefficient choice, it chooses (de) centralization.

Both rules may be inefficient, for the usual reason that majority voting does not take into account intensity of preference. However, in general, neither CR is biased in any particular direction; i.e. neither rule is inefficiently centralized or decentralized. The following example illustrates this.

Example 4

Suppose that there are three regions, with $b_{m1} = b_{m2} = c = 1$; $k = 0.5$; and $b_{m3} = 9$. Then, it is easy to calculate that $E\Delta_1 = E\Delta_2 = -1$, but $E\Delta_3 = 3$. Then, as $\sum_{i=1}^{3} E\Delta_i = 1$, the efficient CR will choose D, but as a majority of regions have $E\Delta_i < 0$, the federal CR will choose C.

On the other hand, suppose that $b_{m3} = 0$, $b_{m1} = b_{m2} = 3$, with the other parameters as before. Then, $E\Delta_1 = E\Delta_2 = 1/3$, but $E\Delta_3 = -1$. Then, as $\sum_{i=1}^{3} E\Delta_i = -1/3$, the efficient CR will choose C, but as a majority of regions have $E\Delta_i > 0$, the federal CR will choose D. $\square$

However, in the asymptotic case as n becomes large, as shown above, there are a number of conditions (A0-A3), under which we can say compare the federal and unitary rules quite straightforwardly. It turns out that under the same assumptions, we can obtain a quite tight characterization of how both rules compare to the efficient benchmark.

First, we can calculate the probability limit of the gain from decentralization under conditions A0-A3 when the principle of aggregation holds;

$$\begin{aligned} \text{plim}_{n\to\infty} \Delta^n &= \int_c^{\bar{b}_m} (b_m - c) g(b_m) db_m - 0.5(Eb_m - c) - 0.5k \\ &= \int_c^{\bar{b}_m} (b_m - c) g(b_m) db_m - 0.5k \\ &= \int_0^{\bar{x}} x g(x) dx - 0.5k \end{aligned} \tag{5.1}$$

We have used the fact that $Eb_m = c$ from A3 in the second line, and a change of variable in the last line. Clearly, from (5.1) centralization is strictly more efficient iff $\int_0^{\overline{x}} xg(x)dx - 0.5k > 0$, or

$$k > 2\int_0^{\overline{x}} xg(x)dx = k_E$$

We now have the following result.

Theorem 5. Assume A0-A3. If G is uniform, then both rules are efficient ($k_F = k_U = k_E$). If G is strictly positively single-peaked, then the federal rule is inefficiently centralized ($k_F < k_E$). If G is strictly negatively single-peaked, then the federal rule is inefficiently decentralized ($k_F > k_E$).

The intuition for this result is the following. When γ deviates from the uniform by (say) becoming strictly positively single-peaked, a proportion $x\%$ of median voters will change their preference from decentralisation to centralisation. But some of these "switchers" will only gain a very small amount, as they were nearly indifferent, so that the increase in expected benefit from centralisation relative to decentralisation (i.e. the percentage change in Δ^n) will be less than $x\%$.

It remains to say something about the unitary rules in the non-uniform case. Let $\mathcal{F}$ be the set of symmetric single-peaked zero-mean distributions on $[-\overline{y}, \overline{y}]$, and let $\mathcal{A} \subset \mathcal{F}$ have the property that for any two distributions F, F' in $\mathcal{A}$, one distribution is a mean-preserving spread of another. Then, due to the symmetry of members of $\mathcal{A}$, if F' is a mean-preserving spread of F, the variance of F' is greater than that of F. Suppose we denote the variance of F by σ_F^2; this number[23] uniquely defines any F in $\mathcal{A}$. and so the line is negatively sloped. Then we can state the following:

Theorem 6. Assume $F \in \mathcal{A}$. Under the assumptions of Theorem 2, there exists an $\overline{\sigma}^2$ such that (i) for any $\sigma_F^2 < \overline{\sigma}^2$, the unitary rule is more efficient than the federal rule, but is inefficiently centralized; (ii) when $\sigma_F^2 = \overline{\sigma}^2$, the unitary rule is efficient. (iii) when $\sigma_F^2 > \overline{\sigma}^2$, the unitary rule is inefficiently decentralized.
Under the assumptions of Theorem 3, there exists an $\overline{\sigma}^2$ such that (i) for any $\sigma_F^2 < \overline{\sigma}^2$, the unitary rule is more efficient than the federal rule, but is inefficiently decentralized; (ii) when $\sigma_F^2 = \overline{\sigma}^2$, the unitary rule is efficient. (iii) when $\sigma_F^2 > \overline{\sigma}^2$, the unitary rule is inefficiently centralized.

So, not surprisingly, when σ_F^2 is small, the unitary CR behaves in a similar

[23] Note that $\sigma_F^2 \leq \overline{y}^2$, as $\overline{y}^2$ is the maximum possible variance of all distributions in $\mathcal{A}$.

may to the federal CR, but less obviously, when σ_F^2 is large enough, the unitary CR may exhibit a different direction of inefficiency than the federal CR.

6. Related Literature and Conclusions

This paper has attempted both a positive and normative analysis of two "constitutional rules" for choosing the degree of decentralization, in the setting of a particular model of the costs and benefits of decentralization. In the asymptotic case, we have obtained a number of results about how these rules differ; the key determinant of the difference seems to be how median voter preferences are distributed across regions. The efficiency of both CRs has also been investigated. One of the main limitations of the current study is that the CRs studied are very oversimplified. In particular, one key feature of real-world CRs is that they are "rigid" in the sense that they favor the status quo e.g. rules for constitutional amendment in federal states. One topic for future research is to investigate the positive and normative properties of these rigid CRs.

The results of this paper can be compared to Cremer and Palfrey(1996). In the basic model of their paper, where taste parameters were Normally distributed, they showed[24] that - in our terminology - whenever a unitary CR chooses centralization, the federal CR does also. Cremer and Palfrey call this result "the principle of aggregation". They also argued that it was robust to several extensions, including other possible statistical distributions for tastes, such as the uniform, but do not provide any general conditions under which it holds[25]. Bearing in mind that the two papers build quite different models of the costs and benefits of decentralization, we can note the following. If ϵ has a Normal distribution with mean zero, (truncated so that b_m is positive), it is symmetric and positively single-peaked, so Theorem 2 applies, and Theorem 2 states the "principle of aggregation" for this

[24] Cremer and Palfrey obtained formulae for the proportion of regions, and the proportion of voters, who prefer centralization, as a function of only one parameter σ^2; the ratio of the inter-regional variance in tastes to the intra-regional variance in tastes. They established two main facts in their paper, both in the limiting case as the number of regions went to infinity. First, when this ratio was below some $\bar{\sigma}^2$, the proportion of regions preferring centralization was greater than the proportion of individual voters preferring centralization: Second; for all values of the ratio below $\bar{\sigma}^2$, the proportion of individual voters preferring centralization was greater than 0.5 (and so the unitary CR chose centralization).

[25] In Section 4.3 of their paper, they present a uniform distribution example where 100% of regions prefer centralization, in which case the principle of aggregation certainly holds, but this example is not general.

model.

There are several other recent papers that are more loosely related to this one. Besley and Coate(1998) and Lockwood(1998) build "political economy" models of the choice between centralization and decentralization, but with a di¤erent focus, namely to study the political ine¢ciency of centralization. Our model has elements of both models, but is deliberately stylized, so as to enable an analysis of federal and unitary CRs. Bolton and Roland (1995) study a model quite similar to the one of this paper[26], but the purpose of their analysis is rather di¤erent, namely, to analyze possible secession by one of the regions, and how the threat of secession changes the tax policy of central government.

7. References

Baron, D.P, "Majoritarian incentives, pork barrel programs, and procedural control", American Journal of Political Science 35 (1991), 57-90

Baron, D.P, and J.Ferejohn, "Bargaining in legislatures", American Political Science Review, 83 (1989), 1181-1206

Besley,T, and S.Coate, Quarterly Journal of Economics 62, (1997), 84-114

Besley, T. and S.Coate(1998), \The Architecture of Government; Centralized versus Decentralized Provision of Public Goods", unpublished paper

Bird,R.M. "Threading the …scal labyrinth: some issues in …scal decentralization", National Tax Journal, 46, (1993), 207-227

Bolton,P. and G.Roland, "The break-up of nations: a political economy analysis", CEPR Discussion Paper No. 1225, 1995

Buchanan,J.M., "A contractarian paradigm for applying economic theory", American Economic Review, 65 (Papers and Proceedings), (1975), 225-230

Buchanan,J.M., "The constitution of economic policy", American Economic Review, 77, (1987), 243-250

Buchanan,J.M., "Contractarian political economy and constitutional interpretation", American Economic Review, 78 (Papers and Proceedings, 1978), 135-139

Caillaud, B., B.Julien, and P.Picard, "National vs. European incentive policies: bargaining, information, and coordination", European Economic Review, 40 (1996), 91-111

[26] In their model, there are two, rather than n, regions, agents di¤er in incomes, rather than preferences for the public good (in their model, the public good can be interpreted as a lump-sum transfer). Also, the economies of scale from centralised public good provision are modelled rather di¤erently.

Cremer,J and Palfrey "In or out? Centralization by Majority vote", European Economic Review, 40 (1996), 43-60

Cremer,J and Palfrey "Political Confederation", American Political Science Review, 93 (1999), 69-83

Curtis,M.(1997), Western European Government and Politics, (Longman, 1997)

Dixit,A. The Making of Economic Policy, (Cambridge, MIT Press, 1996)

J.Ferejohn, M.Fiorina, and R.D.McKelvey, "Sophisticated voting and agenda-independence in the distributive politics setting", American Journal of Political Science (1997) 31, 169-194

Gilbert,G., and P.Picard, "Incentives and the optimal size of local jurisdictions", European Economic Review, 40 (1996), 19-41

Klibano¤,P. and M.Poitevin "A theory of (de)centralization", unpublished paper, University of Montreal, 1996

Lockwood,B (1998) "Distributive politics and the bene…ts of decentralisation", Warwick Research Paper in Economics 513, University of Warwick

Oates,W., Fiscal Federalism (Harcourt, Brace, Jovanovitch,1972)

Osborne,M.J. and A.Slivinski, "A model of political competition with citizen candidates", Quarterly Journal of Economics 61 (1996), 65-96

Persson,T. (1998), "Economic policy and special interest politics", Economic Journal 108, 529-543

Rothschild,M. and J.E.Stiglitz,"Increasing risk I: a de…nition", Journal of Economic Theory (1970), 225-243

Wheare,K.C. Federal Government, (Oxford, Oxford University Press, 1963)

A. Political Equilibrium with Centralization

Legislative Equilibrium

Some details of the legislative game are as follows. Let K be the set of regions that elect delegates. Assume for the moment that $K = N$. A time index $t = 0, 1, 2, ...$ tracks the number of "rounds" of legislative bargaining. The strategies of the delegates are as follows. If at time t delegate i is proposer, he must choose a proposal $a_i = (x_1^i, ...x_n^i) \in X^n$ conditional on the history of play up to that point, H_t. If at time t delegate i is a responder, he must choose a response $r_i \in \{yes, no\}$ conditional on the history of play up to that point (H_t, a_j) where j is proposer. If a strict majority of delegates choose r_i = yes, the proposal is approved and the game terminates. Otherwise, the game continues to the next round.

Following Baron, we restrict attention to stationary equilibria. A stationary equilibrium has the property that whenever g and h are structurally identical subgames, the continuation values of any player i (denoted V_i) are the same in both subgames, no matter what the time period i.e. $V_i(\tau, g) = V_i(\tau', h)$ when $g = h$.

Now let V_i denote the payoff in the subgame beginning with the random selection of candidates. In a stationary equilibrium, delegate i will vote for proposal j iff

$$x_i^j b_i^d - \frac{k(a_i)}{n}(c - k) \geq \delta V_i \tag{A.1}$$

where $\delta < 1$ is the discount factor, and b_i^d is the taste parameter of the delegate. Say that j offers i a project if $x_i^j = 1$. Conjecture then that if delegate i is proposer, he chooses $m - 1$ delegates at random and offers them projects, (as well as choosing a project for himself) and these delegates accept. Given these strategies, delegate i's continuation payoff at the proposer selection stage is

$$V_i = \frac{m}{n}(b_i^d - c + k) \tag{A.2}$$

It is then clear from (A.2) that if i is offered a project, (A.1) reduces to

$$b_i^d - \frac{m}{n}(c - k) \geq \delta \frac{m}{n}(b_i^d - c + k)$$

which certainly holds. So, all delegates will accept projects if offered. It remains to show that this is the best strategy for the proposer. This follows from two observations.

First, if delegate i is proposer, he will never offer projects to more than $m-1$ delegates, as only m delegates are needed to approve a proposal, and additional offers raise the cost to the proposer through the tax rule. Second, if delegate i is not offered a project, he will never accept a proposal as $-\frac{m}{n}(c-k) < \delta\frac{m}{n}(b_i^d - c + k)$ as long as $\delta < 1$, $b_i^d \geq 0$.

Now consider the case where region i does not elect a delegate i.e. $K = N \backslash \{i\}$. If no delegate from region i is elected, then there is a legislative game with $n-1$ delegates excluding region i, but the residents of region i continue to pay tax. There is a stationary equilibrium of the game with $n-1$ delegates with the same structure as that described above, where at each round the proposer j selects at random $m-1$ regions in $N \backslash \{i; j\}$ and offers them projects, and the first proposal is accepted. Consequently, the payoff to any resident of i in this case is

$$W = -\frac{m}{n}(c-k) \tag{A.3}$$

Voting Equilibrium

At the voting stage, every resident of i can vote for one of the candidates, or abstain. First, assume that only one candidate stands. As $V_i - W > 0$, for all $b_i \in B_i$, all voters strictly prefer to be represented rather than not, so a single candidate is always elected.

Second, assume that $l > 1$ candidates are standing for election. We have established that whatever the b_i^d of the delegate from region i, he will adopt the same strategy at the legislative stage. So, all delegates yield any citizen the same payoff. This means that all voters are indifferent between candidates. Assuming w.l.o.g. that indifferent voters randomize over candidates with equal probabilities, all candidates are elected with equal probability.

Candidate Entry

It is clear from (A.2), (A.3) that as long as

$$V_i - W = \frac{m}{n} b_i \geq \sigma$$

it pays a resident of i with taste parameter σ to stand for election, given that no other resident is standing. So, there is always a one-candidate equilibrium where one candidate stands with probability one, as claimed. $\square$

B. Proofs of Propositions

Proof of Proposition 2. We only give the proof for the case where $\hat{b}_i \in B_i$; the proof in the "corner" case is similar. It is clear that if $u^c(\hat{b}_i) = u^d(\hat{b}_i)$ and $x_i^d = 1$

then $u^c(b) < u^d(b)$; $b > \hat{b}_i$, and $u^c(b) < u^d(b)$; $b < \hat{b}_i$. The argument is the same for a region where $x_i^d = 0$. $\square$

Proof of Proposition 3. (i) In this case, all citizens in a given region have identical preferences, so again there are no dissenting voters. It follows immediately from (3.2) and (3.3) that the federal and unitary constitutions are equivalent.

(ii) Let centralization = C; decentralization = D. It is clear that with $b_i \equiv b_j$, $i, j \in N$, regions are unanimous in their choice of some $A \subset \{C, D\}$. As the total net dissenting vote is bounded below 0.5, if the federal CR chooses A, then so must the unitary CR. $\square$

Proof of Fact 1. (a) First consider the choice of a federal CR between C and D. In the example, b_{m2} is constructed to be in the centre of the interval $B_C = (c - k, c + 2k)$, so that b_{m1}, b_{m3} are in this interval iff $\varepsilon < \frac{3}{2}k$.

(b) Now the unitary CR. As $\phi = 0$, all voters in region 2 choose C. So, the proportion of all voters choosing C is

$$\pi = \frac{1}{3}[1 + (1 - F_1(c - k)) + F_3(c + 2k)] \tag{B.1}$$

where

$$F_i(x) = \frac{x - b_{mi} + \theta}{2\theta}; \; i = 1, 3, \; b_{mi} - \theta \leq x \leq b_{mi} + \theta \tag{B.2}$$

So, it is easy to compute

$$1 - F_1(c - k) = F_3(c + 2k) = 0.5 + \frac{\frac{3}{2}k - \varepsilon}{2\theta} \tag{B.3}$$

Combining (B.1) and (B.3), the proportion of voters preferring C is greater than one half ($\pi > 0.5$) iff

$$\theta > 2\varepsilon - 3k$$

As $\theta \geq 0$, the Fact follows. $\square$

Proof of Fact 2. In this case, we suppose that $\phi > \frac{3}{2}k$, so that some voters in region 2 prefer D. Then, the proportion of voters preferring C is then

$$\pi = \frac{1}{3}[F_2(c + 2k) - F_2(c - k) + (1 - F_1(c - k)) + F_3(c + 2k)]$$

Also, using the fact that F_2 is defined as in (B.2) with θ replacing ϕ, it is easy to compute that

$$F_2(c + 2k) - F_2(c - k) = \frac{3k}{2\phi}$$

So, after substitution,

$$\pi = \frac{1}{3}\left[\frac{3k}{2\phi} + 1 + \frac{\frac{3}{2}k - \varepsilon}{\theta}\right]$$

Now assume that $\phi > 3k$. Then $\pi > 0.5$ iff $\theta < \frac{3k-2\varepsilon}{1-3k/\phi}$, as required. $\square$

Proof of Theorem 1. First, we calculate an explicit formula for k_F. By assumption, $\Gamma(x) = \frac{x+\overline{x}}{2\overline{x}}$. So, from (4.4), $\pi_F(k) = k/\overline{x}$, so $k_F = 0.5\overline{x}$. To prove the result, it is sufficient to show that $\pi_U(k_F) = 0.5$. Now note that as Γ is uniform, from (4.5), we have;

$$\pi_U(k) = \frac{1}{\overline{x}}\int_0^{\overline{x}} F(k-x)dx$$

So,

$$\pi_U(k_F) = \frac{1}{\overline{x}}\int_0^{\overline{x}} F(0.5\overline{x}-x)dx = \frac{1}{\overline{x}}\int_{-0.5\overline{x}}^{0.5\overline{x}} F(y)dy;\ y = 0.5\overline{x}-x$$

Now, as $F(0) = 0.5$, and F is symmetric around zero, it is easy to see that $\int_{-0.5\overline{x}}^{0.5\overline{x}} F(y)dy = 0.5\overline{x}$. So, $\pi_U(k_F) = 0.5$ as required. $\square$

Proof of Theorem 2. (i) Write $\pi_U(k;F)$ in (4.5) with the dependence of π_U on F made explicit. Define $k_U(F)$ implicitly by

$$\pi_U(k_U(F);F) \equiv 0.5 \tag{B.4}$$

Let $\mathcal{F}$ be the class of symmetric, single-peaked zero mean distributions on $[-\overline{y},\overline{y}]$. Then, then it is sufficient to show that:

$$k_U(F) > k_F,\ \text{all } F \in \mathcal{F} \tag{B.5}$$

Note also that $k_U(F_0) = k_F$, where F_0 is the degenerate distribution[27] with all the probability mass at $y = 0$. Also, any non-degenerate $F \in \mathcal{F}$ is a symmetric MPS of F_0. Then, to establish (B.5), all we need to prove is that

$$k_U(F') > k_U(F) \tag{B.6}$$

where F' is a symmetric MPS of F. For then from (B.6), $k_U(F) > k_U(F_0) = k_F$, any $F \in \mathcal{F}$ as required.

[27] That is, $F_0(y) = \begin{cases} 0 & y < 0 \\ 1 & y \geq 0 \end{cases}$.

(ii) To prove (B.6), as π_U is increasing in k, it suffices to show that if F' is a mean-preserving spread of F, then $\pi_U(F'; k_F) < \pi_U(F; k_F)$. But from (4.5), this is equivalent to

$$\Delta = \int_0^{\bar{x}} [F(k_F - x)]\gamma(x)dx - \int_0^{\bar{x}} [F'(k_F - x)]\gamma(x)dx > 0$$

Now note that as $y = b - b_m$, $|b - b_m| \le \Gamma^{-1}(0.75)$ implies $\bar{y} \le \Gamma^{-1}(0.75)$. But $\pi_F(k_F) = 2\Gamma(k_F) - 1 = 0.5$, implying $\Gamma(k_F) = 0.75$. But then, $\bar{y} \le k_F$, so we have;

$$\int_0^{\bar{x}} F(k_F - x)\gamma(x)dx = \gamma(k_F - \bar{y}) - \gamma(0) + \int_{k_F - \bar{y}}^{\min\{\bar{x}, k_F + \bar{y}\}} F(k_F - x)\gamma(x)dz$$

So,

$$\Delta = \int_{k_F - \bar{y}}^{\min\{\bar{x}, k_F + \bar{y}\}} [F(k_F - x) - F'(k_F - x)]\gamma(x)dx$$

But now from Rothschild and Stiglitz(1971) if F' is a MPS of F, we must have

$$F'(z) > F(z),\ z < 0;\ F'(z) < F(z),\ z > 0;\ \int_{-\bar{y}}^{\bar{y}} F(z)dz = \int_{-\bar{y}}^{\bar{y}} F'(z)dz \qquad \text{(B.7)}$$

If the MPS is symmetric, then both F, F' must also be symmetric. These conditions in fact imply that $F' - F$ is a symmetric function round 0 in the sense that .

$$F'(-z) - F(-z) = F(z) - F'(z) = \psi(z),\ z > 0 \qquad \text{(B.8)}$$

So, we have;

$$\begin{aligned}
\Delta &= \int_{k_F - \bar{y}}^{k_F} [F(k_F - x) - F'(k_F - x)]\gamma(x)dx + \int_{k_F}^{\min\{\bar{x}, k_F + \bar{y}\}} [F(k_F - x) - F'(k_F - x)]\gamma(x)dx \\
&\ge \int_{k_F - \bar{y}}^{k_F} [F(k_F - x) - F'(k_F - x)]\gamma(x)dx + \int_{k_F}^{k_F + \bar{y}} [F(k_F - x) - F'(k_F - x)]\gamma(x)dx \\
&= \int_0^{\bar{y}} [F(z) - F'(z)]\gamma(k_F - z)dx - \int_0^{\bar{y}} [F'(z) - F(z)]\gamma(k_F + z)dx \\
&= \int_0^{\bar{y}} \psi(z)[\gamma(k_F - z) - \gamma(k_F + z)]dx \\
&> 0
\end{aligned}$$

In this sequence of inequalities, we have used: (i) in the second line, the fact that $F(k_F - x) < F'(k_F - x)$ when $x > k_F$, from (B.7); (ii) change of variables in the third line; (iii) (B.8) in the fourth line; (iv) $\gamma(k_F - z) > \gamma(k_F + z)$ from single-peakedness of Γ and $k_F > \overline{y} > z$ in the final line. So, we have proved $\Delta > 0$, as required. ∎

Proof of Theorem 4. (i) The results concerning k_F are obvious. (ii) Note that it is established in the proof of Theorem 2 that $k_U(F') > k_U(F)$, where F' is any symmetric MPS of F. It follows immediately that when the hypotheses of Theorem 2 hold, k_U rises following a symmetric MPS in F. A similar argument shows that the hypotheses of Theorem 3 hold, k_U falls following a symmetric MPS in F.

(iii) It remains to prove that following a symmetric mean-preserving spread in G, k_U rises. But recall that

$$\pi_U(k; F) = 2\int_0^{\overline{x}} [F(k - x)]\gamma(x)dx$$

Now, consider a symmetric MPS in Γ. This is a sequence of simple symmetric MPSs. From Rothschild and Stiglitz(1971), Each simple symmetric MPS in the sequence (from Γ to Γ') can be characterized as follows

$$\begin{aligned} \gamma'(x) &= \gamma(x) + s(x) \\ s(x) &= \alpha[I_{[-\delta-\epsilon-t;-\delta-\epsilon]} - I_{[-\delta-t;-\delta]} + I_{[\delta+\epsilon;\delta+\epsilon+t]} - I_{[\delta;\delta+t]}, \ \epsilon > t; \ \delta + \epsilon + t < \overline{x} \end{aligned}$$

where $I_{[a;b]}$ is the indicator function on $[a;b]$. So, s is a function that moves a probability mass $t\alpha$ to the tails of the distribution of γ, while keeping the distribution symmetric.

But then

$$\begin{aligned} \pi'_U(k; F) - \pi_U(k; F) &= 2\int^{\overline{x}} [F(k - x)](\gamma'(x) - \gamma(x))dx \\ &= 2\int_0^{\overline{x}} [F(k - x)](I_{[\delta+\epsilon;\delta+\epsilon+t]} - I_{[\delta;\delta+t]})dx \\ &\leq 2t\alpha[F(k - \delta - \epsilon) - F(k - \delta - t)] \\ &< 0 \end{aligned}$$

So, if $\pi_U(k; F) = 0.5$, $\pi'_U(k; F) < 0.5$. Consequently, k must rise to achieve $\pi'_U(k; F) = 0.5$, implying that the curve shifts outward. ∎

Proof of Theorem 5. Denote by H the distribution of x, conditional on x being positive (i.e. $H(x) = 2\Gamma(x) - 1$; so h is given by $2g(x)$, $x \in [0, \overline{x}]$). Then, k_E is the mean of this distribution i.e.

$$k_E = 2\int_0^{\overline{x}} x\gamma(x)dx = E[x|x \geq 0]$$

Also, note that k_F is the median of H; as $H(k_F) = 2G(k_F) - 1 = 0.5$, where the last inequality follows from the definition of k_F

Now if G; and therefore Γ is uniform, then H us uniform also and therefore symmetric, so $k_F = k_E$. If Γ is positively single-peaked, then H is skewed to the right and so $k_E > k_F$. If Γ is negatively single-peaked, then H is skewed to the left and so $k_E < k_F$. $\square$

Proof of Theorem 6. First note that if $F', F \in A$, then F' has higher variance than F iff it is a MPS of F. Now suppose that the assumptions of Theorem 2 hold. Then from the proof of Theorem 2, it is clear that (a) $k_F < k_U(F)$, and (b) $k_U(F)$ is increasing in σ_F^2. Then, as $k_E > k_F$ from Theorem 6; either; (i) there must be some critical variance $\hat{\sigma}^2$ for which $k_U(F) = k_E$, or (ii) $k_U(F) < k_E$, all F. To over both these cases, recall that the maximum possible variance of F is $\overline{\sigma}^2 = \overline{y}^2$; and set $\bar{\sigma}^2 = \min\{\hat{\sigma}^2, \overline{y}^2\}$. In the case where the assumptions of Theorem 2 hold, the proof is similar. $\square$

C. Proof of the Optimality of Decentralization in the Cremer-Palfrey Model

One can calculate the expected utility of all voters in a district (in their terminology) under centralization or decentralization as $EU^C = -E[E[(t-M)|t; d]]$ and $E^{UD} = -E[E[(t-m_d)|t; d]]$. respectively. So, assuming large numbers of equal-sized districts, we get, applying the law of iterated projection and the relevant formulae for $E[(t-M)|t; d]$, $E[(t-m_d)|t; d]$ in Cremer and Palfrey's paper that

$$\begin{aligned} EU^C &= -Et^2 \qquad \text{(C.1)} \\ EU^D &= -\frac{1}{1+\sigma_d^2} - \frac{Et^2}{(1+\sigma_d^2)^2} \end{aligned}$$

Some decentralization is preferred if $EU^D > EU^C$, which requires, after some rearrangement of (C.1), that

$$\frac{2+\sigma_d^2}{1+\sigma_d^2} < Et^2 \tag{C.2}$$

Now, by definition, $Et^2 = E(t - m_d)^2 + 2Etm^d - Em_d^2$. Using the facts that $t \sim N(m_d; 1)$; $m_d \sim N(0; \sigma_d^2)$, (so that $(t - m_d)^2 \sim \chi^2(1)$); we see that

$$Et^2 = 1 + \sigma_d^2 > 1$$

So, as the left-hand side of (C.2) is less than 1; (C:2) must always hold. $\square$