DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

POLITICAL CONFEDERATION

Jacques Crémer
CNRS - GREMAQ and IDEI

Thomas R. Palfrey
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 990

July 1997

Political Confederation*

Jacques Crémer Thomas R. Palfrey

Abstract

Using a spatial model, we compare different rules for aggregating preferences across confederated districts, under the assumption that voters have private information and face uncertainty about the distribution of preferences of other voters.

Our model includes, as special cases, systems of local representation in national assemblies and parliaments and international legislatures with representatives from member states. We show how induced preferences over the degree of centralization and the method of representation (proportional vs. equal representation of districts) vary systematically across voters and districts, depending on such factors as relative size of the districts, the number of districts, and the variance of underlying policy preferences.

We show that each voter has an ideal confederation in which representation consists solely of equal representation. These ideal confederations vary across voters, but are independent district size. Moderate voters prefer a higher degree of centralization than extreme voters. Preference for centralization is increasing in the number of districts and decreasing in the variance of voter ideal points.

With two districts, majority rule equilibria always exists and can have some degree of proportional representation. With three or more districts, majority rule equilibrium often fails to exist. Nonexistence arises due to cycling in the two dimensional—Centralization $\times$ Representation—space of confederations.

Keywords: voting, constitution, federalism

JEL: D72, D78, H11

*This research was supported in part by the National Science Foundation through grants SES-9224787 and SBR-9631627. The work was begun while the second author was visiting IDEI in 1995. Financial support from IDEI, as well as its stimulating research environment, is gratefully acknowledged. That author also thanks CREST-LEI and CERAS for their hospitality and research support during the 1995-96 academic year. We are grateful for comments from seminar participants at London School of Economics, University of Toulouse, the 1996 Gerzensee Worshop on Political Economy, the 1997 SITE summer program on Interregional Competition in Public Economics, and the 1997 annual meeting of the Public Choice Society.

1 Introduction

Episodes of political confederation, or the opposite movement of political dissolution, are widely observed and important historical phenomena, yet the forces underlying the actual process of confederation are not well-understood. There are many different levels of political units, ranging from towns and villages to states or provinces, to nation-states, and even beyond, to supranational bodies such as the United Nations and the European Union. Generally rules exist that govern the interaction between units of different levels, and the allocation of decision-making authority among them. It is the determination of such rules and the development of a theoretical foundation for studying these rules that we address in this paper. We call this the problem of *political confederation*.

The problem of political confederation is one of the most fundamental of all problems in political science. It arises in many contexts, both large and small. Several examples come immediately to mind:

- The ceding of sovereignty to international organizations by member-nations.
- The issue of "states rights" in American politics.
- The nature of representation of provinces in national assemblies.
- The allocation of taxing authority in a federal system.
- The adoption of uniform standards across jurisdictions (pollution, education, roads, etc.).
- The formation of Nation-States
- Constitutional Amendments

A number of specific examples have been studied in great detail by scholars of history and political science, particularly with regard to the evolution and collapse of international organizations and the drafting of constitutions. Most of this work attempts to identify in descriptive terms the important forces that were operative in those isolated incidents, and this has provided some useful insights into the problem. We are attempting here to capture some of the general features of the problem in a formal model, and to derive some theoretical implications.

To this end, we simply model a confederation as a collection of districts, together with rules for the aggregation of preferences of the members of the various districts, which produces a policy outcome that can vary across districts. The districts can be thought of as regions, provinces, counties, villages, or other well-defined decentralized political units. The larger body is referred to as the "Confederation," and it can be thought of as any political unit that contains some collection of smaller political units.

In terms more familiar to economists that political scientists, it is a mechanism design problem, which is a meta-decisionmaking problem in which the decision to be taken is a set of rules for making decisions. It can also be thought of as the problem of constitutional design. What makes the problem of constitutional design especially complicated is that the constitutional designers themselves have preferences[1] over the outcomes and policies that will ultimately be decided according to the rules for preference aggregation that are chosen at the constitution-making stage. Thus, the designers will have *induced preferences* over the rules, that depend upon their preferences over policy outcomes and the specific features of their own district.

The model we use is very simple. It combines the standard spatial model of political competition, with the presence of asymmetric information and multiple districts. It builds on the model of Crémer and Palfrey (1996), in which voters from a district know only their own ideal point in a policy space and some common information about the distribution of preferences in their district and in other districts. With incomplete information, the policy outcomes that emerge from different rules of confederation are uncertain prospects, so the induced preferences over a space of rules reflects not only a voter's ideal point, but also a voter's attitudes about risky outcomes. In the standard Downsian model, preferences necessarily exhibit some degree of concavity over outcomes, which produces risk aversion if the policy outcomes are uncertain.

Our earlier paper considered just two polar extremes of confederation. At one extreme, (the null confederation) there would be no national government at all, and the policy outcomes of different jurisdictions would be determined independently, simply as the median preference of the district (which, it must be remembered, is a random variable in our model). At the polar extreme, the districts cede all sovereignty to a central confederational government, so that all districts have the same policy, which is decided by the median of the nation (composed of all the districts, aggregated together). Because voters do not know the preferences of other voters, the two policy outcomes under the different systems (independent vs. confederated) are random variables which differ in both their conditional mean and variance. Because preferences are quadratic and the uncertainty is Normally distributed, the induced preferences depend only on the mean and variance. This leads to a clean analytical solution.

The basic trade-off for voters is that centralization will result in policy outcomes that are, on average, further from their ideal point, while the risk associated with the centralized government is lower than the risk associated with independent policy making. The latter follows from the fact that the component of risk due to interdistrict variance over policy outcomes is reduced with confederation. We call this risk reduction the *principle of moderation.* Though never formalized before, this principle is not new, and played a central role in the Federalist's justification for the unification of the American states in the late 18^{th} Century. To quote from *The Federalist #10*:

> The influence of faction leaders may kindle a flame within their particular

[1] See the historical analyses of the formation the U.S. constitution, such as Riker (1992).Those studies demonstrate the existence of a clear connection between preferences over policy outcomes and preferences for the institutional structure, and present convincing evidence that this link played a prominent role in the constitutional conventions.

> States, but will be unable to spread a general conflagration through the other States. A religious sect may degenerate into a political faction in a part of the Confederacy; but the variety of sects dispersed over the entire face of it must secure the national councils against any danger from that source. A rage for paper money, for an abolition of debts, for an equal division of property, or for any other improper or wicked project, will be less apt to pervade the whole body of the Union than a particular member of it; in the same proportion as such a malady is more likely to taint a particular county or district, than an entire State.

The central finding of our earlier paper was that moderate voters prefer centralization and extreme voters oppose it. Specifically, there is a critical voter ideal point such that all voters whose ideal point is closer to the expected median prefer centralization and all voters whose ideal point is further away from the expected median prefer independence. The present paper extends this model in two ways.

First we consider a much broader range of constitutions that the two extreme cases of "all or nothing" national government. These intermediate cases are obtained by representing the policy outcome in a district as a weighted average of the local median and the median of the confederation. We call this *the centralization dimension* and show below that this captures a wide range of "federal" systems, in which the relative power of a district in affecting policies in other districts is allowed to vary. For example, the two extreme cases we examined in Crémer and Palfrey (1996) can be identified as special cases, one (all power at the local level) in which no district has any influence over the policies in other districts, and the opposite case, in which the policies in all districts are the same, and are determined by the overall median of the union of all districts.

Second, we introduce another dimension to the confederation design, which we call *the representation dimension.* We consider two contrasting notions of representational formulas, one in which a local unit's representation at the national level is proportional to the population of the unit, and one where each district receives the same absolute representation. We borrow the terminology of *The Federalist #62* and call the first *proportional* representation (with apologies to scholars who study legislative party competition) and the second *equal* representation. An example of these two extremes would be in the U.S. Congress, where, if we think of states as the basic units of the confederation, the House of Representatives approximates proportional representation, while the Senate has equal representation. Hence, we sometimes refer to these two types of representation as "House" and "Senate," respectively. Our representation dimension considers linear combinations of these two schemes, which can loosely be thought of as the allocation of power between two different chambers in a bicameral national legislature (again using the U.S. analogy).

As with the Federalists' arguments about a confederation's ability to moderate the effects of local factions, they also discussed the trade-off between equal and proportional representation. While the latter provides greater retention of local sovereignty for large states, equal representation is a more effective way to moderate factions, especially those from large states. Madison discusses the issue of why one ought to have "equal" (Senate) in addition than "proportional" (House) representation in the legislative body, and he realizes that equal representation of States helps to resolve a trade-off between loss of sovereignty and moderation, in a way that *all* states (large and small) should come to realize it as a good solution to the representation problem. Here's an example of Madison's discussion, taken from *The Federalist #62*[2]:

> ...in a compound republic, partaking both of the national and federal character, the government ought to be founded on a mixture of the principles of proportional and equal representation.... In this spirit, it may be remarked, that the equal vote allowed to each State is at once a constitutional recognition of the portion of sovereignty remaining in the individuals states, and an instrument for preserving that residuary sovereignty.

Accordingly, in this paper we examine these various trade-offs in a formal theoretical model, by characterizing induced voter preferences over a two-dimensional set of possible constitutions, where the two axes are centralization and representation. We obtain the following results.

1. Voter preferences over representation depend on the degree of centralization, and vice versa. Voters from large states prefer more proportional representation the higher is the level of centralization. Voters from small states prefer less proportional representation as the level of centralization increases.

2. All voters have an induced ideal point which consists entirely of equal representation, and with a most-preferred level of centralization that depends on the voter ideal point in the underlying policy space, but does not depend on the district size. Interestingly, this applies equally to voters from large states (relative population size greater than $1/n$), who sacrifice power to small states under equal representation.

3. Consistent with our earlier findings, more extreme voters want less centralization.

4. A majority rule equilibrium over the two-dimensional confederation issue exists within each district.

5. An equilibrium over all voters in all districts may or may not exist with more than two districts. When it does exist, it can involve some degree of proportional

[2] We hasten to add that the Federalists were also well aware that small states would be more likely to ratify a constitution which provided for absolute as well as proportional representation. This property is also captured in our model and is mentioned several times in *The Federalist #62.* For example:

"The equality of representation in the Senate is another point, which, being evidently the result of compromise between the opposite pretensions of the large and the small States, does not call for much discussion... A government founded on principles more consonant to the wishes of the larger States, is not likely to be obtained from the smaller States."

representation, but this can happen only in knife-edge cases. Generically, if an equilibrium exists, it must involve equal representation. Several examples are given to illustrate the range of possibilities.

The rest of the paper is organized as follows. The model and notation is introduced in Section 2. Induced preferences over the two dimensional space of confederations—representation and centralization—are characterized in Section 3. Section 4 presents our results on existence and properties of majority rule equilibria. Formal proofs are in the Appendix.

2 The Model

A *confederation* is composed of n *districts*. In each district, there is a continuum of individuals. The relative size of district i is denoted $\alpha_i > 0$, where $\sum_{i=1}^{n} \alpha_i = 1$. The underlying policy space is one dimensional and represented by $\Re$, the set of real numbers.

Each individual has an *ideal policy*, $t \in \Re$, which we will also call his *type*. If policy $x \in \Re$ is adopted, the utility of an individual of type t is

$$U(x;t) = -(x-t)^2.$$

When voting over constitutions, the agent will be uncertain about the outcome. In this case, between two distributions over policies he prefers the distribution with the higher expected utility.

An individual is represented by his type t and the district i to which he belongs. Within district i, the types of the individuals are Normally distributed with mean m_i and variance 1 (this normalization is harmless). The district means are independent random variables, normally distributed with mean 0 and variance σ. Each individual only knows his own type, the district to which he belongs, and the statistical distribution of preferences.[3]

We assume that the political process within any district yields a policy that is the ideal point of the median voter of that district.

2.1 Confederations

A confederation is an institutional arrangement in which the policies of different districts are, at least in part, influenced by the preferences of voters from other districts in the confederation. In practice, this is usually accomplished through a complex array of overlapping jurisdictions, representative governments at different levels, and a legal

[3] This can be generalized to allow voters to have additional information about their own district. See Crémer and Palfrey [1996] for a discussion of possible extensions such as this.

system that allocates decisionmaking authority and responsibility across these different levels. The end result is ultimately a vector of policies, one for each district. We denote such a vector of policies by $x = (x_1, \ldots, x_n)$. Given the set of institutions and the legal system governing the overlapping jurisdictions, we can think of x as emerging as a function of the underlying preferences in all the different districts. Since we take "district" to be the smallest political unit, and since we assume that political competition within a district will be driven by the median voter of that district, we write $x = C(m_1, \ldots, m_n) = [C_1(m_1, \ldots, m_n), \ldots, C_n(m_1, \ldots, m_n)]$. That is, the constitution of confederation C is modeled as a function which maps vectors of district medians into vectors of district policies. This function might be different for different districts, and generally will be different in the applications we explore below. To make the analysis tractable, we limit consideration to confederations where $C_i(m_1, \ldots, m_n)$ is linear:

$$C_i(m_1, \ldots, m_n; \lambda) = \sum_{j=1}^{n} \lambda_{ij} m_j,$$

with $\lambda_{ij} \geq 0$ and $\sum_j \lambda_{ij} = 1$. That is, we make the simplifying assumption that a confederation can be represented by a matrix of *influence coefficients*, $\Lambda = [\lambda_{ij}]$, where λ_{ij} is the influence of district j on the policy outcome in district i.

2.2 The Representation/Centralization Axes

In our earlier paper (Crémer and Palfrey [1996]), we considered only two cases, one in which $C_i(m_1, \ldots, m_n) = m_i$ (independent districts) and a second (complete unification) in which policy outcomes correspond to the ideal point of the median voter of the entire union:

$$C_i(m_1, \ldots, m_n) = \sum_{j=1}^{n} \alpha_j m_j, \quad \forall \quad i.$$

That paper analyzed the induced individual preferences between these two extreme confederations, and identified conditions under which a majority of individuals (or a majority of districts) would prefer unification to independence. Here, we study the induced preferences of individuals over a much richer set of possible confederations.

Using the notation of influence coefficients, these two extremes correspond to $\lambda_{ii} = 1$ for all i (independence) and $\lambda_{ij} = \alpha_j$ for all i, j (unification). In fact, we can think of unification somewhat more generally where $\lambda_{ij} = \lambda_j$ for all i, j and λ_j is not necessarily equal to α_j. These correspond to various forms of representation that are possible under unification. The case of $\lambda_{ij} = \alpha_j$, treated in our earlier paper, corresponds to *Proportional Representation*. *Equal Representation* is defined by $\lambda_{ij} = 1/n$ for all i, j. We can represent a continuum of multicameral unified systems using a parameter, $\gamma \in [0, 1]$, that indexes the relative influence of proportional representation. In a unified system this would correspond to $\lambda_{ij} = \gamma\alpha_j + (1-\gamma)/n$. This defines what we call the *representation dimension,* and spans all ranges of mixtures between purely proportional representation ($\gamma = 1$) and purely equal representation ($\gamma = 0$).

Of course, nearly all unified confederations of states also have governments at the district level, which we call *unified systems with federalism* to reflect the fact that they are a cross between the two extremes of pure independence and full unification. We represent a continuum of degrees of centralization using a parameter, $\beta \in [0,1]$, which represents the relative weight of the central government in determining policy outcome in any given district. This defines what we call the *centralization dimension,* and spans all ranges of mixtures between purely independent systems ($\beta = 0$) with no central government, and fully unified system ($\beta = 1$), where districts have no autonomy at all.

Putting these two dimensions together, we have the two main axes—representation and centralization—that determine our space of *confederations.* In the notation of influence coefficients, the policies in each district will be determined according to:

$$\begin{aligned}\lambda_{ii} &= 1-\beta+\beta(\gamma\alpha_i+(1-\gamma)/n),\\ \lambda_{ij} &= \beta(\gamma\alpha_j+(1-\gamma)/n).\end{aligned}$$

There are many other possibilities, including other forms of national representation besides the proportional and equal methods, and also variable devolution for different states. Thus, one can see that, in spite of our complete abstraction from the institutional details of the legal structure and representative organs, this model is able to capture a wide range of realistic possibilities. The actual policy outcome in district i, $C_i(m_1,\ldots,m_n;\alpha,\beta,\gamma)$ is given by:

$$\begin{aligned}C_i(m_1,\ldots,m_n;\alpha,\beta,\gamma) &= \sum_{j\neq i} m_j\beta(\gamma\alpha_j+(1-\gamma)/n)+m_i\\ &\quad [1-\beta+\beta(\gamma\alpha_i+(1-\gamma)/n)]\\ &= (1-\beta)m_i+\beta\left[\gamma\sum_{j=1}^{n}\alpha_j m_j+(1-\gamma)\sum_{j=1}^{n} m_j/n\right]\\ &= (1-\beta)m_i+\beta\left[\gamma M_h+(1-\gamma)M_s\right],\end{aligned}$$

where $M_h = \sum_1^n \alpha_j m_j$ and $M_s = \sum_1^n m_j/n$ can be interpreted as the *House median* and the *Senate median*, respectively.

The rest of the paper will look at induced voter preferences and majority rule equilibria in this two-dimensional space of confederations. However, it should be noted that there are many other possibilities, including other forms of national representation besides the proportional and equal methods, and also variable devolution for different states. Thus, in spite of the complete abstraction from the institutional details of the legal structure and representative organs, this model is able to capture a wide range of realistic possibilities.

3 Induced Preferences of Voters

In order to study the understand political conflict along these two basic dimensions of the confederation question, in this section we characterize the induced preferences of voters over this (β, γ) space, which is represented by the square, $[0,1] \times [0,1]$. This characterization identifies the shape of voter indifference curves in this space, and shows how the ideal points of voters in this space vary as a function of n, i, σ, t, and α.

For an individual with ideal point t in district i, we can compute the interim (i.e., conditional on t) expected squared distance between the policy outcome and the ideal point for any pair of parameters, β and γ. To simplify the formula, we let $\theta = t^2/(1+\sigma)$, and obtain :

$$
\begin{aligned}
W_{i\theta}(\beta, \gamma) \quad = \quad & \frac{\sigma}{1+\sigma}\left[1 - \beta + \beta/n + \beta\gamma\widehat{\alpha}_i\right]^2 \\
& + \sigma \sum_{j \neq i} \left[\beta/n + \beta\gamma\widehat{\alpha}_j\right]^2 \\
& + \frac{\theta}{1+\sigma}\left[1 + \sigma\left(\frac{n-1}{n}\beta - \beta\gamma\widehat{\alpha}_i\right)\right]^2,
\end{aligned}
$$

where

$$\widehat{\alpha}_i = \alpha_i - \frac{1}{n}.$$

The analysis is simplified by the substitution of[4] $\mu = \beta\gamma$, which yields:

$$
\begin{aligned}
W_{i\theta}(\beta, \mu) \quad = \quad & \frac{\sigma}{1+\sigma}\left[(1 - \beta) + \frac{\beta}{n} + \mu\widehat{\alpha}_i\right]^2 \\
& + \sigma \sum_{j \neq i} \left[\frac{\beta}{n} + \mu\widehat{\alpha}_j\right]^2 \\
& + \frac{\theta}{1+\sigma}\left[1 + \sigma\left(\frac{n-1}{n}\beta - \mu\widehat{\alpha}_i\right)\right]^2,
\end{aligned}
$$

with the constraints, $0 \leq \beta \leq 1$ and $0 \leq \mu \leq \beta$.

Level surfaces of $W_{i\theta}$ are the indifference curves of type (i, t) over different (β, μ) confederations. If W is convex, then the first order conditions obtained by differentiating W by β and μ, characterize the minimum value of $W_{i\theta}$, which is the ideal point of voter (i, t), provided this minimum satisfies $\beta \in [0,1]$ and $\mu \in [0,1]$.

3.1 Convexity of W

In this subsection, we establish that W is convex, and therefore the first order conditions can be used to determine voter ideal points. We further show that, except in unusual cases indifference curves are ellipses which are centered at different points of the $\mu = 0$ axis. That is, all voters ideal points correspond to equal, rather than proportional representation, regardless of the size of a voter's state.

[4]That is, μ is the total weight on the house and θ is a "normalized" type.

Theorem 1 *$W_{i\theta}$ is convex, and strictly convex iff $A_i \equiv (n-1)\sum_{j\neq i}\widehat{\alpha}_j^2 - \widehat{\alpha}_i^2 > 0$.*

Proof: See Appendix.

3.2 Indifference maps in (β, μ) space

For most of the remainder of the paper, we will assume that $A_i > 0$ for all districts. The case $A = 0$ is treated in section 5. From the definition of A_i, it is easy to see that A_i will in general be strictly greater than 0 for all districts. There is only one exception to this, if exactly $(n-1)$ districts are the same size. Notice that if $n = 2$, then $\widehat{\alpha}_1 = \widehat{\alpha}_2$, so it is *always* the case that $A_i = 0$. If $n > 2$, it will generically be the case that $A_i > 0$. We next establish the main properties of the indifference maps when $A_i > 0$.

Theorem 2 *If $A_i > 0$ and $\widehat{\alpha}_i \neq 0$, the indifference curves of the agent of type θ in the district i are ellipses, centered at*

$$\mu^0(\theta) = 0$$

$$\beta^0(\theta) = \frac{1-\theta}{1+\theta\sigma+\frac{(1-\theta)\sigma}{n}},$$

whose major axes have slope r defined by

$$|r| - \frac{1}{|r|} = -\frac{n\left((1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2 + \sigma(\theta-1)\widehat{\alpha}_i^2\right)}{\widehat{\alpha}_i[n+\sigma+\theta\sigma(n-1)]} + \frac{(n-1)}{n\widehat{\alpha}_i} \tag{1}$$

and

$$\operatorname{sgn} r = \operatorname{sgn} \widehat{\alpha}_i. \tag{2}$$

The slope of the indifference curves at the points $(\beta, 0)$ is

$$-\frac{\partial W_{i\theta}/\partial\beta}{\partial W_{i\theta}/\partial\mu} = \frac{n-1}{n\widehat{\alpha}_i}. \tag{3}$$

If $\widehat{\alpha}_i = 0$, the ellipses have their main axis parallel to the $\beta = 0$ axis.

Proof: See Appendix.

To summarize, if $A_i > 0$, the slopes of the main axis of their indifference curves of all agents in the district i have the same sign. Furthermore, the slopes of indifference curves of any voter is constant along the β axis. Moreover, this slope is the same for all voters from the district. The slope of the major axis of indifference curves is positive for all voters from large districts ($\widehat{\alpha}_i > \frac{1}{n}$) and negative for all voters with small districts ($\widehat{\alpha}_i > \frac{1}{n}$).

3.3 Voters' Ideal Confederations

Since W is strictly convex, we can characterize the ideal points by first order conditions.[5] Taking partial derivatives of W with respect to β and γ, setting each of these partial derivatives equal to 0, and then simplifying we get:

$$(1-\theta) - \left[\frac{\beta}{n} - \frac{\mu\widehat{\alpha}_i}{n-1}\right][n+\sigma+\theta\sigma(n-1)] = 0 \tag{4}$$

$$(1-\theta)\widehat{\alpha}_i - \beta\frac{\widehat{\alpha}_i}{n}[n+\sigma+\theta\sigma(n-1)] + \mu\left[(1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2 - (1-\theta)\sigma\widehat{\alpha}_i^2\right] = 0 \tag{5}$$

Theorem 3 *If $A_i > 0$ and $\theta < 1$, then the voter's ideal (feasible) confederation is:*

$$\begin{aligned}\mu^*(\theta) &= 0\\ \beta^*(\theta) &= \beta^0(\theta)\end{aligned}$$

and if $A_i > 0$ and $\theta > 1$, then the voter's ideal (feasible) confederation is

$$\mu^*(\theta) = \beta^*(\theta) = 0.$$

Proof: See Appendix.

Figure 1 illustrates the feasible set and some representative indifference curves for voters from three hypothetical districts. The horizontal axis is the centralization (β) dimension and the vertical axis is the representation (μ) dimension. The shaded area marks the set of feasible confederations. Voter 1 is an extreme voter from a large district. The figure includes the indifference curve of voter 1 that passes through the origin, to illustrate the fact that every extreme voter's ideal confederation is $(0, 0)$. Voters 2 and 3 are moderate voters from large and small districts, respectively.

FIGURE 1 ABOUT HERE

3.4 Discussion and Comparative Statics

The first part of this theorem says that moderate voters ($\theta < 1$), regardless of the size of their district, are unanimously opposed to population-based representation. They prefer the national policy to be decided by representative institutions with the same number of delegates from each district, and prefer power to be transferred to districts via decentralization rather than via population-based representation. For any level of centralization, β, the variance of the centralized component is minimized by setting $\mu = 0$. Since the only advantage of centralization is risk reduction, it follows that any voter's ideal confederation must have $\mu = 0$.

[5]For the moment, we are ignoring the constraints $\beta \in [0,1]$ and $\mu \in [0, \beta]$. We take account for these constraints later.

The second half of the theorem is even simpler to understand, in spite of the fact that the expression obtained for β^* from the first order conditions is negative, so the feasibility constraints bind. In this case, sufficiently extreme voters ($\theta > 1$) want no centralization at all, which corresponds to $\beta = 0$. At $\beta = 0$, the only feasible value of μ is $\mu = 0$.

Thus, we have established that the optimal confederation for all voters from all districts calls for *no proportional representation* at the national level. However, we should keep in mind that this is not quite the same as saying that all voters *unconditionally* prefer equal representation.

The comparative statics of how the ideal points change with respect to the exogenous parameters of the model are straightforward. For voters with $\theta < 1$ it is easy to see that $\beta^*(\theta)$ is increasing in n. The intuition behind this is simple. The moderation, or risk-reduction, benefits of centralization are greater as n is greater, independently of district size, since having more districts reduces the probability that the centralized policy will be dictated by an extremist majority of one wing or the other. For similar reasons, $\beta^*(\theta)$ is decreasing in σ. The higher is σ, the more dispersed is the distribution of voter ideal points within a district, so there will be a higher probability that the centralized policy will be on one extreme or the other.

3.5 Conditional Ideal Points

The final question we address is a voter's optimal value of μ as a function of β. This will be useful later in the paper, where we study the existence of a majority rule equilibrium. For voters with $\widehat{\alpha}_i \neq 0$, the optimal value of μ generally will depend on the choice of β since the axes of the indifference curves are not rotated relative to the coordinate axes. Therefore, this relationship, $\mu^*(\beta)$, will depend on the orientation of the (ellipsoidal) indifference curves of the agents in (β, μ)-space. To see what the solution is, consider again the first order condition from differentiation of W with respect to μ, holding β fixed:

$$(1-\theta)\widehat{\alpha}_i - \beta\frac{\widehat{\alpha}_i}{n}\left[n+\sigma+\theta\sigma(n-1)\right] + \mu\left[(1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2 - (1-\theta)\sigma\widehat{\alpha}_i^2\right] = 0$$

Solving for μ, we get:

Theorem 4 *Fix β. The optimal ideal confederation for a voter θ from district i is*

$$\mu^*(\beta,\theta) = -\frac{(1-\theta)\widehat{\alpha}_i - \beta\frac{\widehat{\alpha}_i}{n}\left[n+\sigma+\theta\sigma(n-1)\right]}{(1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2 - (1-\theta)\sigma\widehat{\alpha}_i^2}$$

if

$$0 \leq -\frac{(1-\theta)\widehat{\alpha}_i - \beta\frac{\widehat{\alpha}_i}{n}\left[n+\sigma+\theta\sigma(n-1)\right]}{(1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2 - (1-\theta)\sigma\widehat{\alpha}_i^2} \leq \beta.$$

Otherwise, $\mu^*(\beta,\theta)$ equals either 0 or β depending on whether the right hand side is less than 0 or greater than β, respectively.

Proof: See Appendix.

Whether or not $\mu^*(\beta,\theta)$ is greater than 0 depends on whether $\widehat{\alpha}_i$ is greater than 0 (large districts) or less than 0 (small districts), and whether β is greater than or less than $\beta^*(\theta)$.

Consider first the case of large districts ($\widehat{\alpha}_i > 0$). Since $\widehat{\alpha}_i$ is greater than 0, the sign of $\mu^*(\beta)$ is the opposite of the sign of

$$\frac{(1-\theta) - \beta\left[1 + \theta\sigma + \frac{(1-\theta)\sigma}{n}\right]}{(1+\sigma)\sum_{j\neq i}\widehat{\alpha}_j^2 + (1+\theta\sigma)\widehat{\alpha}_i^2}.$$

Since the denominator is positive, the whole expression is negative for all voters such that $\theta > 1$. Hence $\mu^*(\beta,\theta) > 0$ for these voters. This establishes that all extremists from larger states prefer at least some degree of proportional representation, for every value of $\beta > 0$.

More generally, in large districts $\mu^*(\beta,\theta) > 0$ if and only if $\beta > \beta^*(\theta)$. That is, voters from large districts are in favor of some degree of a proportional system if and only if there is too much centralization relative to their ideally preferred level of centralization. The intuition is that by putting some weight on the proportional system, the large district can effectively take back some of the sovereignty that was lost because of a high β. Voters from large districts who are sufficiently extreme will want to do this. This can be also be interpreted in the following way. Let $\theta^*(\beta)$ be the solution of $\beta^*(\theta) = \beta$. That is, $\theta^*(\beta)$ is the voter for whom β is the ideally preferred level of centralization. The critical value $\theta^*(\beta)$ is decreasing in β, so voters from large districts prefer some degree of a proportional system if and only if $\theta < \theta^*(\beta)$. That is, voters from large districts who are more extreme than $\theta^*(\beta)$ would like some proportional representation. The critical value $\theta^*(\beta)$ is decreasing in β.

Not surprisingly, preferences over β in small districts go in exactly the opposite direction from the case of large districts. These voters are in favor of some degree of a proportional system if and only if $\beta < \beta^*(\theta)$. The intuition here is exactly the reverse of the intuition for why relatively extreme voters from large districts prefer some degree of proportional representation. In this case, relatively moderate voters (i.e., voters for whom $\beta^*(\theta) > \beta$) will want to cede even more sovereignty. This is done by increasing the weight on proportional representation, since doing so reduces the influence of a small district on the centralized component of the policy. For any fixed value of β there is a critical voter $\theta^*(\beta)$ such that voters from small districts prefer some degree of a proportional system if and only if $\theta > \theta^*(\beta)$.

4 Majority Rule Equilibrium

4.1 Majority Rule Equilibrium Within a District

A *majority rule equilibrium in district* i is a centralization-representation pair (β, μ) with the property that there does not exist another pair (β', μ') such that a majority of members in district i prefer (β', μ') to (β, μ). One can show that every district has a majority rule equilibrium, $(\beta_i^*, 0)$, where $\beta_i^* = \beta^*(\theta_{med}^i)$ and θ_{med}^i is the median value of θ in district i. Moreover, if $A_i > 0$ then this is the unique majority rule equilibrium in that district. This is stated formally and proved below, with an informal proof and explanation in Figure 2.

Theorem 5 *If $A_i > 0$, then within district i, there is a unique majority rule equilibrium confederation at $(\beta^*(\theta_{med}^i), 0)$.*

Proof: See Appendix.

FIGURE 2 ABOUT HERE

It is also true that the within-district majority rule equilibria, $\beta^*(\theta_{med}^i)$, is independent of the relative size of the district, α_i, but varies systematically with the underlying parameters of the model, n and σ, and to a limited extent, on the profile of sizes of other districts, α_{-i}. specifically, $\beta^*(\theta_{med}^i)$ is increasing in n, decreasing in σ, is higher the more moderate is the district's median voter ideal point.

4.2 Equilibrium Confederations

The analysis in the previous section does not address the question of how the induced preferences in the various districts are aggregated to reach an inter-district decision about the degree of centralization and the extent to which representation is proportional. We consider one possibility here. A referendum is conducted in which every individual in every district votes, and outcomes are determined by aggregating votes at the confederation level[6]. Thus, an *equilibrium confederation* is a centralization-representation pair (β, μ) with the property that there does not exist another pair (β', μ') such that a majority of voters in the confederation prefer (β', μ') to (β, μ).

The argument for existence of majority rule equilibria within a district, relied heavily on the assumption that the induced (β, μ)– indifference maps of voters from the same district are elliptical and have the same slopes along the β-axis. In fact, this is true for voters whose districts are the same size. Hence the previous analysis holds up if all districts are the same size, i.e., $\alpha_i = \alpha_j = 1/n, \forall i, j$. This case is easy to analyze since it

[6]There are other possible ways to aggregate votes at the confederation level, including, for example, "conventions," in which each district sends a delegate, who is assumed to vote at the conventions based on the majority preference within his district.

is degenerate, in the sense that the centralized policy is neutral with respect to μ, since the proportional and equal representation models are the same when districts are of equal size. For this special case, majority rule equilibria exist and are given by $(\beta^*(\overline{\theta}), \mu)$ where $\overline{\theta}$ is the median of the medians and μ is any value between 0 and 1.

The more interesting case, at least for addressing issues of representation, is that case where districts differ in size. In this case, the indifference curves for voters can vary in important ways across districts. In particular, the major axes of the elliptical indifference curves are sloped upward for large districts, and downward for small districts. This leads to situations where majority rule equilibria may not exist, in the sense that there may not exist (β, μ) pairs that are undefeated by majority rule.

The following result fully characterizes majority rule equilibria when they exist. Specifically, $\mu = 0$ is a necessary condition for (β, μ) to be a majority rule equilibrium. Moreover, $\beta = \beta^*(\overline{\theta})$ is also a necessary condition. That is:

Theorem 6 *If $N > 2$ and there does not exist i such that $\alpha_j = \alpha_k$, $\forall j, k \neq i$, then $(\tilde{\beta}, \tilde{\mu})$ is a majority rule equilibrium only if $\tilde{\mu} = 0$ and $\tilde{\beta} = \beta^*(\overline{\theta})$.*

Proof: See appendix. An informal proof is given in Figure 3.

FIGURE 3 ABOUT HERE

The following example, which is robust, shows that majority rule equilibria may not exist. Suppose there are three districts, whose median values of θ are given by $\theta^1_{med} \geq 0$, $\theta^2_{med} > 0$, and $\theta^3_{med} > 0$. Further suppose that $\alpha_1 > 1/3 > \alpha_2 > \alpha_3$ and $\beta^*(\theta^1_{med}) < \beta(\overline{\theta}) < \beta^*(\theta^2_{med}) < \beta^*(\theta^3_{med})$, where $\overline{\theta} = \alpha_1\theta^1_{med} + \alpha_2\theta^2_{med} + \alpha_3\theta^3_{med}$. This is shown on Figure 4.

FIGURE 4 ABOUT HERE

We next show that $(\beta^*(\overline{\theta}), 0)$ is not an equilibrium. Since district 1 is a large district, all voters in district 41 with ideal points, $\tilde{\theta}$ such that $\tilde{\theta} > \overline{\theta}$ have $\mu^*(\beta^*(\overline{\theta}), \tilde{\theta}) > 0$. Since $\beta^*(\theta^1_{med}) < \beta^*(\overline{\theta})$, this consists of a strict majority of members of district 1. Similarly, since districts 2 and 3 are small districts, all voters in districts 2 and 3 with ideal points $\tilde{\theta}$ such that 4 $\tilde{\theta} < \overline{\theta}$ have $\mu^*(\beta^*(\overline{\theta}), \tilde{\theta}) > 0$. Since $\beta^*(\theta^2_{med}) > \beta^*(\overline{\theta})$ and $\beta^*(\theta^2_{med}) > \beta^*(\overline{\theta})$, this consists of a strict majority of members of each of these small districts. Thus, for sufficiently small values of $\mu > 0$, a strict majority prefers $(\beta^*(\overline{\theta}), \mu)$ to $(\beta^*(\overline{\theta}), 0)$. Since $(\beta^*(\overline{\theta}), 0)$ is the only possible majority rule equilibrium, this implies that there is no majority rule equilibrium in this case.

4.3 The Case of $A_i = 0$

We treat this special case separately, because the results are somewhat different, and because it corresponds to several interesting problems similar to confederation, which commonly arise when there are only two groups (districts). One such situation is civil war or separatist movements, which are the reverse of the confederation problem, and presumably arise because a large majority of one of the districts in the confederation have preferences that are sufficiently distant from the median of the confederation. Frequently these are dyadic in character, pitting one member of the confederation against its complement.[7] We hasten to add that there are also many examples of unification of two districts, including Czechoslovakia, the Belgium-Luxembourg monetary agreement, and the common U.S. phenomenon of municipality annexation and consolidation.

Theorem 7 *If $A_i = 0$, the indifference curves of the agents of district i are pairs of parallel lines given by:*

$$\frac{\beta}{n} - \frac{\mu\widehat{\alpha}_i}{(n-1)} = K$$

where K is a constant. Each voter in such a district has a continuum of ideal points, consisting of the set of all confederations lying on the indifference line passing through the point $\beta^*(\theta)$, defined in Theorem 1.

Proof: See Appendix.

When $n = 2$, we can show that there is always a unique equilibrium, and it may lie anywhere between the β axis and the $\beta = \mu$ diagonal. In the special case when $\alpha_1 = \alpha_2$, all confederations with the same β produce equivalent policies in every district, so uniqueness is defined relative to such equivalence classes of confederations.

Theorem 8 *If $n = 2$, then there exists a unique majority rule equilibrium.*

Proof: See Appendix.

5 Conclusions

This paper investigated questions of sovereignty and representation in confederations. We model sovereignty formally by allowing policy outcomes in one district to be affected by the distribution of voters in other districts and by the relative sizes of districts. We model the representation dimension as a choice of the degree to which each district is represented proportionally as opposed to equally. The balance of power between the two houses of the U.S. Congress is a classic example of how such trade-offs may be achieved in practice. The questions are posed, first at the individual level, and then at the aggregate level

At the individual level the questions reduce to:

[7]Examples would include the War Between the States, contemporary separatist movements such as Quebec and the Basques, and historical colonial independence struggles such as Algeria and India.

(I1) To what extent would voters prefer to sacrifice sovereignty of their own district in order to ensure more moderate policy outcomes?

(I2) How will voters from large states differ from voters from small states in their induced preferences over proportional vs. equal representation?.

The answer to (I1) generalizes the earlier findings in Crémer and Palfrey (1996). Induced preferences over sovereignty cut across traditional left-right political cleavages. Instead, sovereignty issues will find extremists from both sides of the political spectrum agreeing with each other and in opposition to relatively moderate voters, who prefer more centralization. Sufficiently extreme voters will be "separatists," who ideally prefer no confederation at all. We also show that preference for more centralization (i.e., willingness to cede sovereignty) increases with n, decreases with σ, and is independent of the size distribution of districts.

The answer to (I2) is more complicated. All voters' ideal confederations would consist entirely of equal representation. However, preferences over representation cannot be disentangled from the sovereignty issue. That is, voters induced preferences over proportional vs. equal representation change depending on the degree of centralization. For voters from large states, if there is too much centralization, they would like some degree of proportional representation, which restores some of their ceded sovereignty. The opposite is true for voters from small states. They would like some degree of proportional representation only if there is too little centralization.

At the aggregate level, we the central questions are:

(A1) What are the properties of majority rule equilibrium in a district?

(A2) What are the properties of majority rule equilibrium across the whole federation?

The answer to (A1) is that within each district a majority rule equilibrium exists and coincides with the ideal confederation of the median value of θ in that district. The answer to (A2) is that a majority rule equilibrium may not exist. If there are three or more districts, special conditions are required in order for an equilibrium to exist. When equilibrium confederations do exist, the equilibrium form of representation must be entirely equal representation. The case of two districts is special. With two districts, a unique equilibrium confederation always exists and, depending on the relative extremism of the larger vs. the smaller state, the equilibrium degree of proportional representation can range from entirely equal representation to entirely proportional representation.

The nonexistence of equilibrium in the general case of more than two districts is problematic, and is indicative of the importance of agenda manipulation and procedural control at constitutional conventions. Riker (1986) recounts numerous episodes of agenda manipulation at the U.S. Constitutional Convention of 1787, consistent with the sort of preference aggregation problem identified in this paper.

6 APPENDIX

6.1 Proof of Theorem 1

The second order condition for convexity of W requires that the determinant

$$D_{it} = \begin{vmatrix} \partial^2 W_{i\theta}/\partial\beta^2 & \partial^2 W_{i\theta}/\partial\beta\partial\mu \\ \partial^2 W_{i\theta}/\partial\beta\partial\mu & \partial^2 W_{i\theta}/\partial\mu^2 \end{vmatrix}$$

is nonnegative. W is strictly convex if and only if $D_{it} < 0$. We first establish that D_{it} has the same sign as $A_i = (n-1)\sum_{j\neq i}\widehat{\alpha}_j^2 - \widehat{\alpha}_i^2$. Expanding $W_{i\theta}$ gives:

$$\begin{aligned} W_{i\theta}(\beta,\mu) &= \frac{\sigma}{1+\sigma}[1+\beta^2(1-\frac{1}{n})^2+\mu^2\widehat{\alpha}_i^2-2\beta(1-\frac{1}{n})+ \\ &\quad 2\mu\widehat{\alpha}_i - 2\beta\mu(1-\frac{1}{n})\widehat{\alpha}_i]+ \\ &\quad \sigma\sum_{j\neq i}[\frac{\beta^2}{n^2}+\mu^2\widehat{\alpha}_j^2+2\frac{\beta}{n}\mu\widehat{\alpha}_j]+ \\ &\quad \frac{\theta}{1+\sigma}[1+\sigma^2(1-\frac{1}{n})^2\beta^2+\sigma^2\mu^2\widehat{\alpha}_i^2+ \\ &\quad 2\beta\sigma(1-\frac{1}{n})-2\mu\sigma\widehat{\alpha}_i-2\sigma^2\beta\mu(1-\frac{1}{n})\widehat{\alpha}_i] \\ &= \frac{\beta^2}{n^2}\frac{\sigma(n-1)}{1+\sigma}[n+\sigma+\sigma\theta(n-1)] \\ &\quad +\mu^2\frac{\sigma}{1+\sigma}[(1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2+\sigma(\theta-1)\widehat{\alpha}_i^2] \\ &\quad -2\beta\mu\frac{\sigma}{1+\sigma}\frac{\widehat{\alpha}_i}{n}[n+\sigma+\theta\sigma(n-1)] \\ &\quad +2\mu\frac{\sigma}{1+\sigma}\widehat{\alpha}_i[1-\theta]-2\beta\frac{\sigma}{1+\sigma}\frac{(n-1)}{n}[1-\theta]+\frac{\sigma+\theta}{1+\sigma}. \end{aligned} \tag{6}$$

Taking partial derivatives of W gives

$$\begin{aligned} \frac{(1+\sigma)^2}{\sigma^2}D_{it} &= \begin{vmatrix} \frac{n-1}{n^2}(n+\sigma+\sigma\theta(n-1)) & -\frac{\widehat{\alpha}_i}{n}[n+\sigma+\theta\sigma(n-1)] \\ -\frac{\widehat{\alpha}_i}{n}[n+\sigma+\theta\sigma(n-1)] & (1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2+\sigma(\theta-1)\widehat{\alpha}_i^2 \end{vmatrix} \\ &= \frac{n+\sigma+\sigma\theta(n-1)}{n^2}\begin{vmatrix} n-1 & -\widehat{\alpha}_i[n+\sigma+\theta\sigma(n-1)] \\ -\widehat{\alpha}_i & (1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2+\sigma(\theta-1)\widehat{\alpha}_i^2 \end{vmatrix} \end{aligned}$$

Hence D_{it} has the same sign as:

$$(n-1)\left((1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2+\sigma(\theta-1)\widehat{\alpha}_i^2\right)-\widehat{\alpha}_i^2[n+\sigma+\theta\sigma(n-1)] = (1+\sigma)A_i.$$

The final step of the proof is to show that $D_{it} \geq 0 \; \forall \; i, t.$ and that $D_{it} = 0$ if and only if $\widehat{\alpha}_j = \widehat{\alpha}_k \; \forall \; j, k$ such that $j \neq i$ and $k \neq i$. Fixing $\widehat{\alpha}_i$, we must have $\sum_{j=1}^{n} \widehat{\alpha}_j = -\widehat{\alpha}_i$. Under this constraint $\sum_{j=1}^{n} \widehat{\alpha}_j^2$ has a unique minimum where all the $\widehat{\alpha}_j$'s are equal to each other, which implies that they are all equal to $\widehat{\alpha}_i/(n-1)$. In this case

$$A_i = (n-1) \sum_{j \neq i} \left(\frac{\widehat{\alpha}_i}{n-1} \right)^2 - \widehat{\alpha}_i^2 = 0,$$

which proves the result. ∎

6.2 Proof of Theorem 2

To show that a quadratic function defines an ellipse, it is sufficient to show that the function is strictly convex, which is guaranteed by Theorem 1.

To see that the absolute value of r satisfies equation (1), we apply the standard formula for the slope of an ellipse, which is

$$r - \frac{1}{r} = \frac{A - B}{C},$$

where A is the coefficient on β^2, B is the coefficient on μ^2, and C is the coefficient on $\beta\mu$ in the quadratic expression for $W_{i\theta}(\beta, \mu)$. Note that $r - 1/r$ is increasing on $(-\infty, 0)$ and $(0, +\infty)$, and takes values on both these intervals on $(-\infty, +\infty)$. Therefore the formula only determines r up to its sign. The formula gives

$$\begin{aligned} r - \frac{1}{r} &= \frac{(1+\sigma)\sum_{j=1}^{n} \widehat{\alpha}_j^2 + \sigma(\theta - 1)\widehat{\alpha}_i^2 - \frac{n-1}{n^2}[n + \sigma + \sigma\theta(n-1)\sigma\theta]}{-\frac{\widehat{\alpha}_i}{n}[n + \sigma + \theta\sigma(n-1)]} \\ &= -\frac{n\left((1+\sigma)\sum_{j=1}^{n} \widehat{\alpha}_j^2 + \sigma(\theta-1)\widehat{\alpha}_i^2\right)}{\widehat{\alpha}_i[n + \sigma + \theta\sigma(n-1)]} + \frac{n-1}{n\widehat{\alpha}_i} \\ &= \frac{1}{\widehat{\alpha}_i}\left(\frac{n-1}{n} - \frac{n\left((1+\sigma)\sum_{j\neq i} \widehat{\alpha}_j^2 + (1+\sigma\theta)\widehat{\alpha}_i^2\right)}{n + \sigma + \theta\sigma(n-1)} \right) \end{aligned}$$

The sign of r follows direction. To prove (3), observe that

$$\begin{aligned} -\left.\frac{\partial W_{i\theta}/\partial\beta}{\partial W_{i\theta}/\partial\mu}\right|_{\mu=0} &= -\frac{\frac{(n-1)\sigma}{n(1+\sigma)}\left[-1 + \theta + \frac{\beta}{n}(n + \sigma + \sigma\theta(n-1))\right]}{\frac{\sigma}{1+\sigma}\left[\widehat{\alpha}_i(1-\theta) - \frac{\beta}{n}\widehat{\alpha}_i\,(n + \sigma + \theta\sigma(n-1))\right]} \\ &= -\frac{n-1}{n} \times \frac{1}{-\widehat{\alpha}_i}. \end{aligned}$$

∎

6.3 Proof of Theorem 3

Proof: From Lemma 3, these voters have elliptical indifference curves. Multiplying (4) by $-\widehat{\alpha}_i$ and adding it to (5), we obtain:

$$\mu\left[(1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2-\sigma\widehat{\alpha}_i^2-\frac{\widehat{\alpha}_i^2}{n-1}(n+\sigma)\right]=0.$$

Since $A_i>0$, the term in brackets is not equal to 0, so $\mu^*=0$. Plugging $\mu^*=0$ into 4, yields

$$\beta^*=\frac{n(1-\theta)}{n+\theta+(n-1)\theta\sigma}.$$

If $\theta<1$, then $\beta^*\in[0,1]$ so the feasibility constraints ($\beta\in[0,1]$ and $\mu\in[0,\beta]$) do not bind, so the solution given by the first order conditions characterizes the minimum value of W over the set of feasible confederations. Hence

$$\mu^*=0$$

$$\beta^*=\frac{n(1-\theta)}{n+\sigma+(n-1)\theta\sigma}.$$

for these voters.

If $\theta>1$, then the unconstrained value of β^* is negative, so at least one constraint ($\beta\in[0,1]$ and $\mu\in[0,\beta]$) must bind. Since W is convex, the constraint, $\mu\leq\beta$ must bind. This is illustrated in Figure 1. Thus, $\theta>1$ implies $\mu^*=\beta^*$. For this case, the constrained optimum can be obtained by replacing μ by β in W, differentiating and solving. To verify that the solution is $\mu^*=\beta^*=0$, we rewrite the objective function as:

$$\begin{aligned} W_{i\theta}(\beta,\beta) &= \frac{\sigma}{1-\sigma}\left[(1-\beta+\beta\alpha_i)\right]^2 \\ &\quad+\sigma\beta^2\sum_{j\neq i}\alpha_j^2 \\ &\quad+\frac{\theta}{1+\sigma}\left[1+\sigma\beta\left(1-\alpha_i\right)\right]^2 \end{aligned}$$

Setting the derivative of this with respect to β equal to 0 and simplifying gives:

$$\beta=\frac{1-\theta}{\frac{\sum_{j\neq i}\alpha_j^2}{1-\alpha_i}(1+\sigma)+(1-\alpha_i)(1+\theta\sigma)}.$$

Since $A_i>0$ and $\theta>1$, this is always negative. Hence the optimal nonnegative value of β is 0. ∎

6.4 Proof of Theorem 4

Recall the first order condition from differentiation of W with respect to μ, holding β fixed:

$$(1-\theta)\widehat{\alpha}_i - \beta\frac{\widehat{\alpha}_i}{n}\left[n+\sigma+\theta\sigma(n-1)\right] + \mu\left[(1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2 - (1-\theta)\sigma\widehat{\alpha}_i^2\right] = 0$$

Solving for μ, we get:

$$\mu = -\frac{(1-\theta)\widehat{\alpha}_i - \beta\frac{\widehat{\alpha}_i}{n}\left[n+\sigma+\theta\sigma(n-1)\right]}{(1+\sigma)\sum_{j=1}^{n}\widehat{\alpha}_j^2 - (1-\theta)\sigma\widehat{\alpha}_i^2}$$

If this satisfies the feasibility constraints, then this determines μ^*. Otherwise, by convexity of W, the solution lies on the boundary, as described in the theorem.

6.5 Proof of Theorem 5

First notice that if $\beta^*(\theta^i_{med}) \leq 0$, then the result is trivial. Then $(0,0)$ is the preferred point in the feasible set for all since then the median is the ideal point for a majority of voters (i.e., those voters with $\beta^*_i(\theta) = 0$).

Thus, suppose $\beta^*(\theta^i_{med}) > 0$. Consider the two halfplanes defined by the line, $L(\theta^i_{med})$, following the major axis of the indifference map of voter θ^i_{med} and passing through $(\beta^*(\theta^i_{med}), 0)$. This is illustrated in Figure 2. Consider any point in the halfplane consisting of all points to the left of this line, such as the dot in the figure. Since the major axes of the voters are parallel to each other along the $\beta - axis$, it follows immediately that all voters with θ such that $\beta^*_i(\theta) > \beta^*(\theta^i_{med})$ prefer $\beta^*(\theta^i_{med})$ to y. This is, by definition of $\beta^*(\theta^i_{med})$, exactly half the voters, so that y does not defeat $\beta^*(\theta^i_{med})$ by a majority. A similar argument applies to any point in right halfplane, such as the point marked with a cross in the figure. In this case, all voters with θ such that $\beta^*_i(\theta) < \beta^*(\theta^i_{med})$ prefer $\beta^*(\theta^i_{med})$ to y, so y does not defeat $\beta^*(\theta^i_{med})$ by a majority.

To show uniqueness of the majority rule equilibrium, notice that for any point (β', μ') not equal to $(\beta^*(\theta^i_{med}), 0)$, continuity implies that all voters with ideal points in a neighborhood of θ^i_{med} prefer $(\beta^*(\theta^i_{med}), 0)$ to (β', μ'). By the argument above, it is also true that either all the voters to the left of θ^i_{med} or all the voters to the right of θ^i_{med} prefer $(\beta^*(\theta^i_{med}), 0)$ to (β', μ'). Thus, a strict majority prefers $(\beta^*(\theta^i_{med}), 0)$ to (β', μ'), hence none of these other points can be majority rule equilibria. ∎

6.6 Proof of Theorem 6

If $N > 2$ and there does not exist i such that $\alpha_j = \alpha_k$, $\forall j,k \neq i$, then $(n-1)\sum_{j\neq i}\widehat{\alpha}_j^2 - \widehat{\alpha}_i^2 > 0$ for all i, so all indifference curves are ellipses. We first show that $\widetilde{\mu} = 0$.

Suppose not, and consider an equilibrium $e = (\widetilde{\beta}, \widetilde{\mu})$ with $\widetilde{\mu} \in (0, \widetilde{\beta}]$, and $\widetilde{\beta} > 0$. Denote by $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i)$ the optimal β for a type θ voter from district with relative size $\widehat{\alpha}_i$, given $\widetilde{\mu}$. Formally, this is obtained by rearranging the first order condition, equation (4), to get:

$$\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i) = \frac{n(1-\theta)}{n + \sigma + (n-1)\theta\sigma} + \frac{n\widetilde{\mu}\widehat{\alpha}_i}{n-1}$$

Note that this may fall outside the feasible bounds given by the restriction that $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i) \in [\widetilde{\mu}, 1]$. Since we have supposed that $(\widetilde{\beta}, \widetilde{\mu})$ is a majority rule equilibrium, it must be the case that $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i) > \widetilde{\beta}$ for 50% of the voters and $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i) < \widetilde{\beta}$ for 50% of the voters. If not, one could either increase or decrease β slightly, keeping μ fixed at $\widetilde{\mu}$, and a majority would vote for the new proposal. Now consider all voters for whom $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i) < \widetilde{\beta}$. A typical voter like this is shown in Figure 3 where we denote $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i)$ by v^2.

Now consider rotating the horizontal line, defined by $P = \{(\beta, \mu) \mid \mu = \widetilde{\mu}\}$ clockwise by a very small angle of rotation, ϕ. This new, downward sloping line, denoted PQ, is shown in the figure. For each voter type $(\theta, \widehat{\alpha}_i)$, define $p^*_{P_\phi}(\theta, \widehat{\alpha}_i)$ as this voter's most preferred feasible confederation, restricted to PQ. The point, $p^*_{P_\phi}(\theta, \widehat{\alpha}_i)$, for the typical voter described above is marked v^3 in the figure. Because the preference gradient for this voter points directly downward at v^2, this implies that v^3 lies to the left of e. This is also true for any voter for whom $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i) = \widetilde{\beta}$, as illustrated by the point labelled v^1 in the figure, and by continuity will also be true for a positive measure of voters for whom $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i) < \widetilde{\beta}$, such that $\beta^*_{\widetilde{\mu}}(\theta, \widehat{\alpha}_i)$ is sufficiently close to $\widetilde{\beta}$. Thus there must exist a proposal lying to the left of e on P_ϕ which a majority of voters prefer to $(\widetilde{\beta}, \widetilde{\mu})$. Therefore, $(\widetilde{\beta}, \widetilde{\mu})$ cannot be a majority rule equilibrium if $\widetilde{\mu} > 0$.

Finally, suppose that $(\widetilde{\beta}, 0)$ is a majority rule equilibrium where $\widetilde{\beta} \neq \beta^*(\overline{\theta})$. Then there exists a majority of voters such that either $\beta^*(\theta) < \widetilde{\beta}$ or $\beta^*(\theta) > \widetilde{\beta}$. Suppose w.l.o.g. there is a majority of voters such that $\beta^*(\theta) > \beta$. Then for small enough ϵ, for all $\beta' \in (\widetilde{\beta} + \epsilon, \widetilde{\beta})$, a majority would prefer $(\beta', 0)$ to $(\widetilde{\beta}, 0)$. Thus a necessary condition for $(\widetilde{\beta}, \widetilde{\mu})$ to be an equilibrium is that $\widetilde{\beta} = \beta^*(\overline{\theta})$ and $\widetilde{\mu} = 0$. ∎

6.7 Proof of Theorem 7

The coefficient of μ^2 in (6) is equal to

$$\begin{aligned} & \frac{\sigma}{1+\sigma}[(1+\sigma)\sum_{j\neq i}\widehat{\alpha}_j^2 + (1+\sigma+\sigma(\theta-1))\,\widehat{\alpha}_i^2] \\ = \;& \frac{\sigma}{1+\sigma}[(1+\sigma)\sum_{j\neq i}\widehat{\alpha}_j^2 + (1+\sigma\theta)\,\widehat{\alpha}_i^2]. \end{aligned}$$

If $A_i = 0$, this is equal to

$$\frac{\sigma}{1+\sigma}\widehat{\alpha}_i^2[\frac{1+\sigma}{n-1} + 1 + \sigma\theta] = \frac{\sigma}{(n-1)(1+\sigma)}\widehat{\alpha}_i^2[\sigma + n + \sigma\theta(n-1)].$$

This enables us to rewrite (6) as follows

$$\begin{aligned}
W_{i\theta}(\beta,\mu) &= \frac{\sigma(n-1)}{1+\sigma}[n+\sigma+\sigma\theta(n-1)]\left(\frac{\beta}{n}-\frac{\mu\widehat{\alpha}_i}{(n-1)}\right)^2 - \\
& \quad 2\frac{\sigma}{1+\sigma}(n-1)[1-\theta]\left(\frac{\beta}{n}-\frac{\mu\widehat{\alpha}_i}{(n-1)}\right)+\frac{\sigma+\theta}{1+\sigma}
\end{aligned}$$

which proves the result. ∎

6.8 Proof of Theorem 8

If $\alpha_1 = \alpha_2$ then all voters have vertical indifference lines. Consider the ideal line of voter $\overline{\theta} = \frac{1}{2}\theta_1 + \frac{1}{2}\theta_2$, where θ_1 and θ_2 are the medians of district 1 and district 2, respectively. Any point $(\beta^*(\overline{\theta}), \mu)$ such that $\mu \in [0, \beta]$ is an equilibrium.

Next suppose that $\alpha_1 > \alpha_2$. Let $P^*(\theta_1)$ and $P^*(\theta_2)$ denote the ideal lines of θ_1 and θ_2, respectively. It is easily verified that these two lines have slopes that are equal in magnitude (greater than 1), with opposite signs; $P^*(\theta_1)$ is upward sloping and $P^*(\theta_2)$ is downward sloping.

There are three cases to consider. The intersection of $P^*(\theta_1)$ and $P^*(\theta_2)$ lies either above the diagonal (Case 1), below the horizontal axis (Case 2), or at some point $(\widetilde{\beta}, \widetilde{\mu})$ with $\mu \in (0, \beta)$ (Case 3).

Case 3 is easiest since for any $(\beta, \mu) \neq (\widetilde{\beta}, \widetilde{\mu})$, less than 50% of the voters prefer (β, μ) to $(\widetilde{\beta}, \widetilde{\mu})$. Thus $(\widetilde{\beta}, \widetilde{\mu})$ is a majority rule equilibrium. Obviously no other point can be a majority rule equilibrium for this case.

In case 2, consider the indifference line of the voter with the median preference, $\overline{\theta} = \alpha_1\theta_1 + \alpha_2\theta_2$. Voters with this ideal point in the underlying policy space will have the median ideal confederation, $\beta = \beta^*(\overline{\theta})$, on the horizontal axis ($\mu = 0$), regardless of whether they come from the large district or the small district. In Figure 3, we illustrate the intersection of the ideal lines of two representative $\overline{\theta}$ voters, one from a large district, and one from a small district. We also show in that figure the ideal lines of θ_1 and θ_2 voters, the median voters of district 1 and district 2, respectively. Notice, that since the intersection of $P^*(\theta_1)$ and $P^*(\theta_2)$ lies below the horizontal axis (Case 2), a majority of district 1 voters have ideal lines to the right of the $\overline{\theta}$ ideal line, and a majority of district 2 voters have ideal lines to the left of the $\overline{\theta}$ ideal line. Therefore, $(\beta^*(\overline{\theta}), 0)$ is a majority rule equilibrium. It is easy to verify than there is no other (β, μ) that can be an equilibrium.

Finally, for case 1, consider the preferences of voters restricted to (β, μ) on the diagonal. The ideal policy on this diagonal for a θ_1 voter, $\beta^*_{\beta=\mu}(\theta_1)$, is to the left of the ideal policy on this diagonal for a θ_2 voter, $\beta^*_{\beta=\mu}(\theta_2)$. Hence there is a point between

$\beta^*_{\beta=\mu}(\theta_1)$ and $\beta^*_{\beta=\mu}(\theta_2)$ which represents the median ideal point restricted to the diagonal, which is denoted M^*. Notice, however, that it does not correspond to $\beta^*_{\beta=\mu}(\overline{\theta})$, since the ideal point on the diagonal for a $\overline{\theta}$ type voter will differ depending on whether that voter is from a small or a large district. The indifference line for the district 1 voter that passes through M^* corresponds to a voter with a lower θ than the indifference line for the district 2 voter that passes through M^*. This is illustrated in the figure. Now, by an argument similar to Case 2, it is clear than a majority of district 1 voters have ideal lines to the left of M^* and a majority of district 2 voters have ideal lines to the right of M^*, but exactly half the voters overall have ideal lines to the right of M^*, and half to the left. Because the slope of the indifference lines for voters from district 1 is greater than 1, those indifference lines cut the diagonal from below. Therefore, any move to the right of M^* is opposed by a majority, as is any move down. Therefore, M^* is a majority rule equilibrium. It is easily verified that no other point can be a majority rule equilibrium.

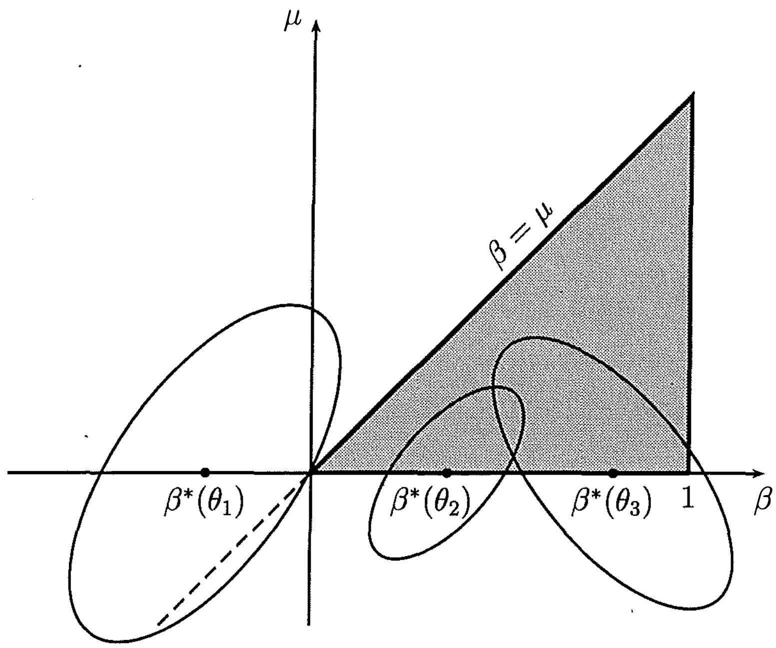

Figure 1: On this figure, we have represented the indifference curves of three agents. Because the ideal points of the voters depend only on their θs, we see that we have $\theta_3 < \theta_2 < \theta_1$. Voter 3 is a moderate (his preferred confederation has positive β) from a small district (the slopes of his indifference curves for $\mu = 0$ is negative). Voters 2 and 1 are respectively a moderate and an extremist from large districts. Voter 1 enables us to illustrate another point. She would like a negative β, but because the slope of her indifference curve at confederation $(0, 0)$ is strictly greater than 1, as it must be for all agents from large districts, her preferred feasible confederation is $(0, 0)$.

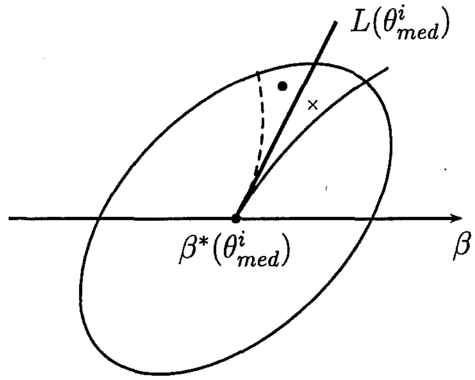

Figure 2: The point marked $\beta^*(\theta^i_{med})$ is the median ideal confederation in district i. It is also the majority voting equilibrium on the whole feasible set. To show this, we have represented a typical indifference curve of an agent with a small θ (the solid curve) and a typical indifference curve of an agent with a large θ (the dashed curve). Any move to a constitution such as the point marked with a cross will be voted down by all the agents with a large θ and some of those with a small θ. A move to a point such as the one marked with a dot will be voted down by all the agents with a small θ and some of those with a large θ.

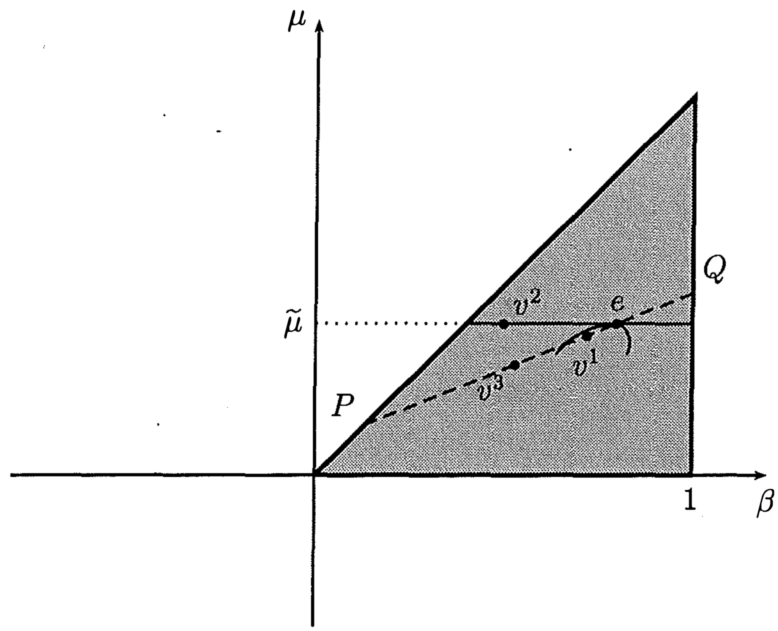

Figure 3: Assume that e is the putative equilibrium. Therefore, it must also be the majority voting equilibrium when choices are restricted to $\mu = \widetilde{\mu}$. This implies, however, that e cannot be a majority voting equilibrium restricted to PQ. To see this, consider first an agent whose preferred point on $\mu = \widetilde{\mu}$ is v^2. By strict concavity and because his ideal confederation is on the horizontal axis, he prefers a point such as v^3 to v^2, and therefore to e. Also by concavity of preferences, his ideal confederation on PQ will be to the left of e. Now, consider the median on $\mu = \widetilde{\mu}$, whose indifference curve is represented on the figure. She prefers v^1 to e, and her ideal confederation on PQ will be to the left of e. By continuity, this will also be true for some voters whose ideal points on $\mu = \widetilde{\mu}$ are "just" to the right of e. Hence, more than half of the voters will have an ideal confederation on PQ to the left of e, and the contradiction is established. The proof in the appendix formalizes the argument, and takes care of potential equilibria on the boundary of the feasible set.

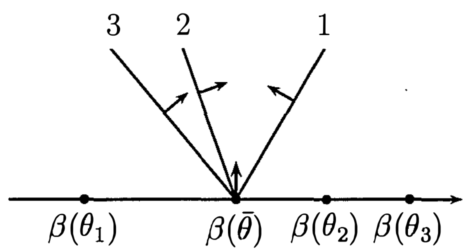

Figure 4: This figure illustrates the fact that there can be no equilibrium in votes aggregated at the confederation level. There are three districts, $i = 1, 2, 3$, with $\alpha_3 < \alpha_2 < 1/3 < \alpha_1$, whose medians are denoted by θ_i. Let $\bar{\theta} = \sum_{i=1}^{3} \alpha_i \theta_i$. If there were a majority rule equilibrium, it would have to be the confederation $(\beta = \beta^*(\bar{\theta}), \mu = 0)$. This is not possible. We have represented the tangents to the indifference curves of the median voters of the three districts going through $\bar{\theta}$. The light arrows indicate the direction of the gradient of preferences at the presumed equilibrium. A small movement away from that equilibrium, such as that indicated by the thick arrow, would increase the welfare of more than half of the voters in each district.

References

[1] Crémer, J., A. Eustache, P. Seabright, 1993, "Lessons of the Theory of the Firm for the Decentralization of Public Services," typescript, IDEI, University of Toulouse.

[2] Crémer, J. and T. Palfrey, 1996, "In or Out?: Centralization by Majority Vote," *European Economic Review*, 40:43–60.

[3] de Groot, M., 1970, *Optimal Statistical Decisions*, New York: McGraw-Hill.

[4] Hamilton, A., J. Jay, and J. Madison, 1788, *The Federalist: A Commentary on the Constitution of the United States*, New York:The Modern Library

[5] Oates, W. E., 1972, *Fiscal Federalism*, New York: Harcourt, Brace, Jovanovich.

[6] Riker, W. H., 1986, *The Art of Political Manipulation*, New Haven: Yale University Press.

[7] Sinn, H.-W., 1993, "How Much Europe? Subsidiary, Centralization, and Fiscal Competition," *CEPR Discussion Paper #834*, Centre for Economic Policy Research, London.