Federal Directives, Local Discretion and the Majority Rule\textsuperscript{1}

Antoine LOEPER\textsuperscript{2}

November 2009

\textsuperscript{1}I wish to thank Jacques Crémer, Helmut Cremer, Michel Lebreton and seminar participants at Toulouse University, EEA congress 2008 in MILAN, ASSA meeting 2009 in San Francisco for helpful comments. All remaining errors are mine.

\textsuperscript{2}Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, a-loeper@kellogg.northwestern.edu
Abstract

We consider a federation in which heterogeneous members determine by federal majority rule a discretionary policy space which restricts their sovereignty in order to trade-off the need for flexibility and policy harmonization. Citizens first vote on the size of the discretionary space (the degree of local discretion), then on its location (the federal directive) and finally each state vote on its respective policy within the discretionary space. Because of the conflict of interests between heterogeneous voters at the federal level, in equilibrium, the federal directive varies negatively with the preferences of non-median voters, and the degree of local discretion is too small and insufficiently sensitive to the magnitude of externalities. Hence, the model shows that the inadequacy and excessive rigidity of federal interventions can arise in a perfectly democratic federation without agency costs or institutional inefficiencies.

JEL Classification Numbers: H77, D72

Keywords: Federalism, Voting, Centralization, Decentralization, Directive.
1 Introduction

In its most simple formulation, federalism is about allocating optimally the control of public responsibilities between federal and local jurisdictions in order to exploit their comparative advantage. Economic models typically derive conditions under which a given policy domain is better handled solely at the central or at the local level. However, the functioning of actual federal systems does not necessarily fit this dichotomy. For various policies, decision rights are shared by different layers of governments. In the E.U., most state laws are transpositions of European directives. These federal directives impose binding constraints on member states but leave them some leeway in their transposition and implementation. Likewise, the Stability and Growth Pact boils down to a set of bounds on states’ fiscal policies. In the U.S., the Sentencing Reform Act imposes sentencing ranges to state courts but gives them full flexibility within the range. These non-prescriptive interventions can be viewed as delimiting a space of local discretion to coordinate the policies of the members of the federation while leaving them complete sovereignty within the discretionary space.

In actual federal systems, the formulation of federal directives and the degree of freedom devolved to the local level are often decided by federal bureaucrats and politicians who may not have the correct incentives or information. This paper considers the alternative scenario in which the degree of local discretion is determined through an open and neutral process governed by direct democracy. More precisely, we consider a federation which members have heterogeneous preferences but also care about policy coordination. We characterize the equilibrium of a voting game in which the discretionary
space is determined by federal majority rule and local policies are then chosen within the federal bounds by local majority rule.

The presumed advantage of this type of two-tier decision process is two-fold: first, it allows to combine the comparative advantage of decentralization – responsiveness to local circumstances – and centralization – policy coordination. Second, the information and the degree of verifiability necessary for their implementation is more reasonable than what is required by the kind of contingent transfers that an optimal mechanism design approach would prescribe. However, citizens and local policy makers often complain that federal directives are poorly responsive to local needs and that the discretion devolved to the local level is too limited, even when the gains from policy coordination are negligible. In the sequel, we shall refer to these two biases respectively as the preferences matching and the federal encroachment problems. They can easily be explained by agency costs in a bureaucratic federal administration with centralist preferences. This conclusion has led to the common wisdom that a more democratic political process at the federal level could remedy these biases and bring federal interventions closer to the citizens’ needs.

To investigate this claim, we analyze a model in which voters determine by direct democracy the orientation and flexibility of the federal directives. In each state, a unidimensional policy has to be implemented. Voters have heterogeneous preferences and benefit from policy coordination. The federal intervention consists in an interval $[\Gamma - \Delta, \Gamma + \Delta]$ within which states can choose the policy that best meets the needs of its constituents. Citizens first vote at the federal level on the size of the interval $\Delta$, then on its location
Finally, state policies are chosen by state majority rule within the discretionary interval \([\Gamma - \Delta, \Gamma + \Delta]\). The size of the interval \(\Delta\) can be thought of as the degree of freedom of states and the vote on \(\Delta\) as a constitutional referendum which determines the degree of decentralization. The location of the interval \(\Gamma\) can be interpreted as the federal directive, i.e. a guideline from which states should not depart by more than \(\Delta\). Since all decisions are taken by the voters themselves and no institutional imperfection is introduced, federalism boils down to a simple preferences aggregation problem.

The main results are the following: first, consistently with the aforementioned preferences matching problem, the equilibrium federal directive \(\Gamma\) varies negatively with the preferences of non-median states. The magnitude of this bias depends on the skewness of preferences. Second, consistently with the aforementioned federal encroachment problem, the equilibrium degree of discretion \(\Delta\) is too limited and insufficiently sensitive to the severity of externalities. The magnitude of this bias depends on the on the degree of polarization of preferences. Third, when externalities are sufficiently small, federal directives are socially worse than decentralization and leave a majority of voters worse-off. When externalities are sufficiently large, federal directives are always supported by a majority of voters even when they are socially dominated by decentralization.

These results show that dysfunctional federal institutions or the neglect of local specificities by federal elites may not be the only culprit for the excessive rigidity and inadequacy federal interventions. These inefficiencies can also arise in a neutral democratic process as a result of the conflict of interests between heterogeneous voters at the federal level. Since the fed-
eral directives do not have the same effect on all states, sophisticated voters benefit from imposing an excessively rigid bound on the states that need the most flexibility: by doing so, they can maximize policy coordination at little cost to themselves. However, these biases go beyond the traditional problem of the majority enslaving the minority because in our voting game, the pivotal voters may not be the same at both voting stages. For this reason, a majority of voters can be made worse-off by the federal intervention. To be more concrete, the vote on the directive $\Gamma$ at the second stage generates a conflict between voters on the opposite sides of the preferences spectrum and the voter with median preferences is pivotal. The vote on the degree of local discretion $\Delta$ opposes moderate voters who push for more harmonization versus extreme voters who push for greater flexibility, so the pivotal voter can be on either side of the preferences spectrum. At the first stage, strategic voters consider not only the direct effect of $\Delta$, i.e. the trade-off flexibility/harmonization, but also the indirect, strategic effect on the second stage equilibrium $\Gamma(\Delta)$. Extreme voters on both sides of the preferences spectrum have congruent preferences with respect to the first effect, but their interests collide on the second one. It turns out that the second effect always plays in favor of less discretion because it divides extreme voters while moderate voters always prefer more harmonization. Hence, our model illustrates the simple idea that in a federation composed of a group of “core” states with homogeneous preferences and a more heterogeneous group of “peripheral” states, peripheral states have difficulties forming a cohesive opposition to the centripetal influence of core states.

The paper is organized as follows: section 2 discusses the related litera-
ture. Section 3 lays out the model. Section 4 characterizes the voting equilibrium and section 5 derives the welfare analysis. Section 6 focuses on the case of triadic federations and section 7 concludes. All proofs are relegated to the appendix.

2 Related Literature

Since the seminal work of Oates (1972), a large normative literature has analyzed the costs and benefits of complete centralization and decentralization.¹ A more recent literature has endogenized the cost of centralization by looking at the incentives of voters and politicians at the local and central level.² We depart from this literature in two respects: we consider a mechanism in which decision rights are shared between the local and the central level and we endogenize the degree of decentralization through a popular referendum.

A number of papers have looked at the sharing of fiscal responsibilities in federal systems, in particular voting models of public good provision: unfunded federal mandates (Cremer and Palfrey 2000, 2006), dual provision of pure public goods (Epple and Romano 2003) and local public goods (Hafer and Landa 2007, Lulfesman 2008). Because these papers consider free-riding rather than harmonization problems, the federal intervention is tantamount to a uni-directional constraint on local policies.

Hatfield and Padro-i-Miquel (2008) endogenize the architecture of a federation through a vote on the vertical allocation of public good provision. They

¹See e.g. Oates 1999 or Epple and Nechyba 2004 for a review.
show that an intermediate degree of decentralization allows the capital-poor median voter to commit not to tax capital too heavily. In our model, the intermediate degree of decentralization results from a trade-off between flexibility and coordination.

3 The Model

For all vector $x \in \mathbb{R}^N$, $\overline{x}$ denotes its mean $\frac{1}{N} \sum_n x_n$ and $med(x)$ denotes its median coordinate for the usual order $\leq$ on $\mathbb{R}$. For all $n \in \{1,..,N\}$ and $x_o \in \mathbb{R}$, $(x_o, x_{-n})$ denotes the vector $x$ in which the $n^{th}$ coordinate has been replaced by $x_o$. For all $x, y \in \mathbb{R}^N$, $x \leq y$ means that for all $n$, $x_n \leq y_n$.

3.1 The Federation

We consider a federation composed of a finite and odd number $N$ of jurisdictions that we call states for concreteness. In each state $n = 1, .., N$, a unidimensional policy $x_n$ has to be implemented, which can be thought of as a law or regulation. The welfare of the residents of state $n$ are given by:

$$U_n(\theta_n, x) = -|x_n - \theta_n|^2 - \frac{\beta}{N} \sum_{m \neq n} |x_n - x_m|^2,$$  \hspace{1cm} (1)

The first term in (1) corresponds to the local effect of the state policy. The vector of state types $\theta \in \mathbb{R}^N$ allows for heterogeneity across states. The second term corresponds to the gains from policy coordination. It embodies

\footnote{We could easily introduce heterogeneity of preferences within states as well. All our results carry over if we assume that votes are aggregated at the federal level trough the “one state one vote” rule (i.e. votes are aggregated first at the state level via intra-state majority rule and then at the federal level via inter-state majority rule, see e.g. Crémer}
the uncertainty or unfairness generated by inconsistent laws, the transaction costs and non-tariff barriers to trade caused by a fragmented regulatory system or the mobility costs generated by incompatible school curricula and course systems. The parameter $\beta > 0$ refers to the magnitude of these costs.$^4$

We assume that no majority of voters have the same type, so voters agree that policies should be harmonized to some extent but disagree on the direction of harmonization. The social welfare ordering is derived from the usual utilitarian social welfare function $W = \sum_n U_n$.

Under decentralization, each state has complete sovereignty on its policy. Assuming that voters can secure the policy that maximizes their welfare taking the other policies as given, the equilibrium $x^{dec}$ is given by

$$x_n^{dec} = \frac{\theta_n + \beta \theta}{1 + \beta}, \quad \text{for all } n = 1..N.$$  

(2)

It can easily be shown that $x^{dec}$ is not Pareto optimal whenever states’ types are not uniform: voters do not internalize interstate externalities and choose policies which are too heterogeneous to be socially optimal. For instance, the policy $x^*$ which maximizes $W$ is given by

$$x_n^* = \frac{\theta_n + 2 \beta \theta}{1 + 2 \beta}, \quad \text{for all } n = 1..N.$$  

(3)

which is a mean preserving contraction of $x^{dec}$. Hence, a federal intervention could improve on decentralization by imposing some degree of policy coordination.

$^4$For the case $\beta = 0$, see Cremer and Palfrey 1996, 1999 and 2000.
Typically, decentralization is compared to a centralized regime in which a uniform policy vector is chosen by federation-wide majority rule. Since induced preferences on uniform policies are single-peaked, the centralized voting equilibrium is:

\[ x^c = (\theta_\mu, ..., \theta_\mu). \]  

(4)

States are indexed so that \( \theta_1 \leq ... \leq \theta_N \) and \( \mu \equiv \frac{N+1}{2} \) refers to the state with median preferences. By convention, the states \( n \) such that \( \theta_n < \theta_\mu \) (resp. \( \theta_n > \theta_\mu \)) are called leftist (resp. rightist) states, although the type space should not necessarily be interpreted as an ideological spectrum. This admittedly stylized one dimensional setup captures nevertheless an important aspect of heterogeneous federations which plays a central role in our argument: states which are better-off under centralization than under decentralization form a relatively homogeneous coalition. On the contrary, states which oppose centralization form a polarized group of right and left extreme type. For concreteness, we shall refer to the former states (loosely defined as states which type is relatively close to \( \theta_\mu \)) as core states and the latter (states which preferences are relatively distant from \( \theta_\mu \)) as peripheral states.

### 3.2 The Federal Mechanism

As stated in Oates Decentralization Theorem, the gains from policy coordination under unitarian centralization may not outweigh the cost of a

---


6See Oates 1972, or Loeper 2008 for a version of the decentralization theorem in a setup similar to this one.
one-size fits-all policy. This paper analyzes a more flexible form of federal intervention in which states’ sovereignty is limited to a discretionary interval $[\Gamma - \Delta, \Gamma + \Delta]$ within which states can choose the policy that best meets their specific needs. Citizens first vote at the federal level on the discretionary interval and then vote at the state level on their respective policy within the federal bounds $[\Gamma - \Delta, \Gamma + \Delta]$. This class of federal mechanism can accommodate different degrees of decentralization (via the range of the interval) and any policy orientation (via the location of the interval on the preferences spectrum). In particular, it encompasses both decentralization (as $\Delta \to +\infty$) and centralization (for $\Gamma = \theta, \Delta = 0$). By submitting the discretionary interval to a popular vote, we let voters decide on the degree of decentralization.

As argued earlier, this kind of mechanism is reminiscent of the way actual federal systems coordinate the policies of their members. European directives leave member states some leeway in their transposition and implementation and often specify an interval of time for implementation. In the U.S., the Sentencing Reform Act provides a grid of sentencing ranges for fine and jail time for each offence category. Its goal is to “provide certainty and fairness” while “avoiding unwarranted sentencing disparities” and “maintaining sufficient flexibility”. Likewise, several nations have a core curriculum which specifies a set of goals for student achievement but leaves some degree of dis-

---

7 Article 249 of the Treaty establishing the European Community: “A directive shall be binding, as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods.”

cretion to localities and schools as to the course organization, readings and teaching methods.

At a more speculative level, this type of coordination mechanism can be considered in any policy domain for which some degree of harmonization is beneficial but complete centralization is undesirable or politically infeasible. For instance, it could be used by central banks to coordinate their policy so as to mitigate currency swings without adopting to a monetary union. Likewise, it could help the members of a monetary union to coordinate their fiscal policy in times of crisis without resorting to complete political integration.\footnote{In that case, the upper bound would prevent states which need large fiscal stimulus from exerting too much inflationary pressures on the common currency while the lower bound would force the other states not to free-ride on the fiscal expansion of the most distressed states.}

Intuitively, the location $\Gamma$ of a socially optimal discretionary interval should vary positively with the voters’ types $\theta$ and the degree of local discretion $\Delta$ should decrease with the need for coordination $\beta$. The central question of the paper is which directive and what degree of discretion would emerge under direct democracy. To answer this question, we consider the following voting game:

1. The federation vote on the size of the discretionary interval $\Delta$.

2. The federation vote on the location of the discretionary interval $\Gamma$.

3. Each state vote on its respective policy within $[\Gamma - \Delta, \Gamma + \Delta]$.

The first stage can be interpreted as a constitutional referendum on the degree of decentralization. The second stage can be viewed as voting on a
federal directive, a guideline from which states should not depart by more than $\Delta$. We define formally the solution concept in section 4.

### 3.3 The State Equilibrium

We first derive the equilibrium at the third stage of the above sequential voting mechanism for any subgame $(\Gamma, \Delta)$. Since citizens in each state share the same preferences, they all vote for their most preferred policy and we assume that they take the policies of the other states as given.

**Definition 1** A state equilibrium is a policy vector $x$ such that

$$
\text{for all } n, \ x_n \in \arg \max_{y_n \in [\Gamma-\Delta, \Gamma+\Delta]} U_n (\theta_n, (y_n, x_{-n})).
$$

(5)

**Proposition 1** For all $\theta \in \mathbb{R}^N$, $\Gamma$ and $\Delta \geq 0$, there is a unique state equilibrium which we denote $x(\theta, \Gamma, \Delta)$. It is characterized by the following fixed point:

$$
\text{for all } n, \ x_n (\theta, \Gamma, \Delta) = \max (\Gamma - \Delta, \min (\Gamma + \Delta, \frac{\theta_n + \beta \bar{x}(\theta, \Gamma, \Delta)}{1+\beta})).
$$

We denote respectively $l(\theta, \Gamma, \Delta)$ and $r(\theta, \Gamma, \Delta)$ the number of states constrained by the left and right bound, i.e. the states such that respectively $\frac{\theta_n + \beta \bar{x}}{1+\beta} < \Gamma - \Delta$ and $\frac{\theta_n + \beta \bar{x}}{1+\beta} > \Gamma + \Delta$. The unconstrained states $\{l+1, \ldots, N-r\}$ are not directly affected by the federal bounds although their policy is indirectly affected since $\bar{x}(\theta, \Gamma, \Delta)$ is typically not constant in $(\Gamma, \Delta)$.

### 4 The Federal Equilibrium

A federal equilibrium $(\Gamma^e, \Delta^e)$ is a subgame perfect voting equilibrium of the game defined in subsection 3.2: at the third stage, for all $(\Gamma, \Delta)$, the
subgame equilibrium is the state equilibrium \( x(\theta, \Gamma, \Delta) \) (c.f. definition 1). For any \( \Delta \geq 0 \), the set of second stage subgame equilibria \( G(\theta, \Delta) \) is the set of \( \Gamma \) such that for all \( \Gamma' \), \( x(\theta, \Gamma', \Delta) \) is not preferred by simple majority rule to \( x(\theta, \Gamma, \Delta) \). Since the equilibria at the second stage may not be unique, an equilibrium at the first stage is a pair \((\Gamma^e, \Delta^e)\) such that \( \Gamma^e \in G(\theta, \Delta^e) \) and for all \( \Delta \geq 0 \) and \( \Gamma \in G(\theta, \Delta) \), \( x(\theta, \Gamma, \Delta) \) is not preferred by simple majority rule to \( x(\theta, \Gamma^e, \Delta^e) \).

### 4.1 Federal Directive

The state equilibrium defines an induced utility function \( V_n(\Gamma, \Delta) \) for each state \( n \). For a given \( \Delta \), \( l(\theta, \Gamma, \Delta) \) and \( r(\theta, \Gamma, \Delta) \) have discontinuous jumps in \( \Gamma \), which creates a kink in the induced utility function of all states. These kinks can be concave or convex, so \( V_n(\Gamma, \Delta) \) may have multiple peaks in \( \Gamma \). Nevertheless, a Condorcet winner always exists at the second stage.

**Proposition 2** For all \( \Delta \geq 0 \), the second stage equilibria \( G(\theta, \Delta) \) are the most preferred \( \Gamma \) of the voters of the median state. For all \( \Gamma \in G(\theta, \Delta) \), the median state is unconstrained: \( l(\theta, \Gamma, \Delta) > \mu < r(\theta, \Gamma, \Delta) \), and

\[
\Gamma = \theta_\mu + \frac{l(\theta, \Gamma, \Delta) - r(\theta, \Gamma, \Delta)}{l(\theta, \Gamma, \Delta) + r(\theta, \Gamma, \Delta)} \Delta. \tag{6}
\]

Proposition 2 suggests that, in equilibrium, the federal directive is not positively sensitive to the preferences of peripheral states. Indeed, the farther from \( \theta_\mu \) the types of the rightist states are relative to the type of the leftist states, the greater \( r \) relative to \( l \), so from (6), the more leftist the federal directive. The intuition is the following: when choosing the federal directive, the voters of the median state determine which states can implement the
policy they prefer and which states will be constrained by the federal bounds. This has two consequences: first, they will vote for a directive which lets them choose the policy they prefer. Roughly speaking, this corresponds to the $\theta_\mu$ term in (6). Second, conditional on being unconstrained by the federal bounds, their most preferred federal directive is the one that minimizes the heterogeneity of policies across the federation. For this reason, they will choose a federal directive which constrains the most extreme states. This explains the $\frac{\mathcal{I}_+ \Delta}{1 + \mathcal{I}_+}$ term in (6). By imposing a stricter discipline on the states that need the most flexibility, the voters of the median state maximize policy harmonization at a minimal cost to themselves, more so than under the same nominal degree of local discretion $\Delta$ but with a socially optimal federal guideline $\Gamma^* (\theta)$. Indeed, $\Gamma^* (\theta)$ is increasing in the type of constrained states and thus leans towards the most extreme side of the preferences distribution. As the next proposition shows, the equilibrium directive does exactly the opposite.

**Definition 2** A distribution of type $\theta$ is skewed to the right (resp. to the left) if for all $n \in \{1, \ldots, \mu - 1\}$, $|\theta_n - \theta_\mu| < |\theta_{2\mu - n} - \theta_\mu|$ (resp. $>-$).

**Proposition 3** $G(\theta, \Delta)$ is weakly decreasing in non median types for the strong set order: for all $\theta, \theta'$ such that $\text{med}(\theta) = \text{med}(\theta')$ and $\theta \leq \theta'$, if

$^{10}$For instance for any $n \leq l(\theta, \Gamma^*, \Delta)$, $x(\theta, \Gamma^*, \Delta)$ is locally constant in $\theta_n$ so

$\frac{\partial^2 W(U(x(\theta, \Gamma^*, \Delta)))}{\partial \theta_n \partial \Gamma^*} = \frac{\partial}{\partial \Gamma^*} \left( \frac{\partial U}{\partial \theta_n} \right)$. We show in the appendix (lemma 2) that $\frac{\partial U}{\partial \theta_n}$ is increasing in $\Gamma$. From Topki’s theorem (see e.g. Milgrom and Shannon 1994), $\Gamma^*$ must be increasing in $\theta_n$.

$^{11}$Let $X \subset \mathbb{R}$ be weakly greater than $Y \subset \mathbb{R}$ for the strong set order if for all $x \in X$ and $y \in Y$, $\max (x, y) \in X$ and $\min (x, y) \in Y$.  

$^{13}$
\( \Gamma \in G(\theta, \Delta), \Gamma' \in G(\theta', \Delta) \) and \( \Gamma \leq \Gamma' \), then \( \Gamma \in G(\theta', \Delta) \) and \( \Gamma' \in G(\theta, \Delta) \). Moreover, if \( \theta \) is skewed to the right (resp. left), then for all \( \Gamma \in G(\theta, \Delta) \), \( \Gamma \leq \theta_\mu \) (resp. \( \Gamma \geq \theta_\mu \)).

In words, a distribution of preferences is skewed to the right if rightist states are farther from the median state than leftist states. In this case, proposition 3 shows that the voters of the median state bias the directive towards the moderate (i.e. left) side of the preferences spectrum so as to reduce the leeway of the most extreme (i.e. rightist) states and thus force their policies to be more aligned with their own preferences.

This bias is reminiscent of the preferences matching problem mentioned in the introduction. Moreover, it contrasts with the normative approach underlying Oates Decentralization Theorem which relates the welfare performance of a federal intervention to the heterogeneity of preferences. Our model shows that the skewness of the distribution of preferences matters in that it determines in which direction the pivotal voters will tilt the federal directive so as to avoid policies which are most incompatible with their ideal policy.

### 4.2 State Discretion

The set of states which are constrained is not monotonic (in the inclusion sense) in \( \Delta \): as \( \Delta \) decreases, more states become constrained and (6) implies that \( \Gamma \) jumps discontinuously and not necessarily monotonically. As a consequence, induced preferences on \( \Delta \) are neither single-peaked nor order restricted nor continuous. For this reason, a global voting equilibrium may fail to exist at the first stage. In section 6, we show that a global federal
equilibrium always exists in the case of a triadic federation. In the general case, we restrict our attention to local majority rule equilibria (see Kramer and Klevorick 1973) and we call the corresponding pair \((\Gamma, \Delta)\) a local federal equilibrium (henceforth LFE).

**Definition 3** \((\Gamma^e, \Delta^e)\) is LFE if \(\Gamma^e \in G(\theta, \Delta^e)\) and if there exists \(\varepsilon > 0\) such that for all \(\Delta \in [\Delta^e - \varepsilon, \Delta^e + \varepsilon]\) and for all \(\Gamma \in G(\theta, \Delta)\), \(x(\theta, \Gamma, \Delta)\) is not preferred by simple majority rule to \(x(\theta, \Gamma^e, \Delta^e)\).

Since the local majority rule requirement is more permissive, there may be multiple LFE. However, we will see that this equilibrium concept is sufficiently discriminative to derive interesting welfare conclusions.

**Proposition 4** A LFE exists.\(^{12}\) At any LFE \((\Gamma^e, \Delta^e)\), \(\Delta^e > 0\) and at least a majority of states is constrained: \(l(\theta, \Gamma^e, \Delta^e) + r(\theta, \Gamma^e, \Delta^e) \geq \frac{N+1}{2}\).

To see why a majority of state must be constrained at any LFE, it is helpful to notice the following: If \(D\) and \(D'\) are two consecutive points of discontinuity of \(G(\theta, \Delta)\) and if \(G(\theta, \Delta)\) is single-valued on \([D, D']\), we show in the appendix (c.f. lemma 4) that for all \(\Delta \in [D, D']\),

\[
\begin{align*}
\text{for } n & \leq l, \quad \frac{\partial [x_n(\theta, G(\theta, \Delta), \Delta)]}{\partial \Delta} < 0, \quad (7) \\
\text{for } n > N - r, \quad \frac{\partial [x_n(\theta, G(\theta, \Delta), \Delta)]}{\partial \Delta} > 0, \\
\text{for } l < n \leq N - r, \quad \frac{\partial [x_n(\theta, G(\theta, \Delta), \Delta)]}{\partial \Delta} = 0.
\end{align*}
\]

\(^{12}\)Observe that since \(G(\theta, \Delta)\) is neither single-valued nor continuous, we cannot resort to the existence theorem of Kramer and Klevorick (1973).
As the degree of local discretion decreases, the effort of policy harmonization is borne entirely by the constrained states, since the policies of the unconstrained states are unaffected by $\Delta$. For this reason, the unconstrained states unanimously prefer less local discretion. As long as they are not in minority, they will form a “free-riding coalition” pushing for more harmonization.

5 Welfare Analysis

Proposition 4 implies that in equilibrium, the pivotal voters are necessarily from a state constrained by the federal bounds. When voting on $\Delta$, they do not internalize the benefits of harmonization for the unconstrained states and of flexibility for the other constrained states. For small coordination costs, the former is negligible while the latter is not, so the pivotal voters have an incentive to excessively restrict the sovereignty of peripheral states to secure small coordination gains. As a result there will be too little local discretion, which is reminiscent of the federal encroachment problem mentioned in the introduction.

Proposition 5 For $\beta$ sufficiently small, for any LFE $(\Gamma^e, \Delta^e)$, decentralization socially dominates and is majority preferred to $(\Gamma^e, \Delta^e)$. Fixing $\Gamma = \Gamma^e$, the welfare of a majority of voters is strictly increasing in $\Delta$ around $\Delta^e$.

Proposition 5 shows that the inefficiency of the federal intervention goes beyond the usual problem of a majority enslaving a minority of voters. Indeed, the latter argument alone cannot explain why the federal equilibrium is not majority preferred to decentralization (contrary to Cremer and Palfrey 2000, Hafer and Landa 2007 or Lulfesman 2008). That result is somewhat
surprising since the class of intervention we consider encompasses decentralization and all decisions are taken by majority rule. The reason is that at the first stage, the vote on the degree of local discretion opposes moderates versus extreme voters so the pivotal state is not anymore the median state. Since the pivotal voters at each voting stage have distinct incentives, their combined effect may leave a majority of voters worse-off.

The intuition behind the second point of proposition 5 is the following: The pivotal voters at the first stage take into account not only the direct effect of $\Delta$, i.e. flexibility versus coordination, but also the strategic effect on the equilibrium at the second stage, i.e. the effect of $\Delta$ on the incentives of the median state when voting on $\Gamma$. With respect to the first effect, peripheral states push unanimously in favor of more discretion. With respect to the second effect, their preferences collide since at the second stage, rightist and leftist states have conflicting interests. It turns out that the second effect always plays in favor of less discretion: because the median voters always prefer less local discretion, the side of the preferences spectrum towards which the second stage equilibrium $G(\Delta)$ is leaning as $\Delta$ decreases will always command a majority. In words, the heterogeneity of peripheral states makes their incentives at the first voting stage not perfectly aligned. Hence, the excessive rigidity of federal directives is linked to the difficulty of the peripheral states to form a unified opposition against the centripetal influence of core states.

As argued earlier, the pivotal voters do not internalize the benefits of flexibility for peripheral states and policy harmonization for unconstrained states. As the magnitude of externalities increases, both effects are of simi-
lar order and the bias of the pivotal voters becomes ambiguous. Hence, the welfare properties of the federal intervention depend on the details of the distribution of preferences. For simplicity, we restrict attention to symmetric profiles of type, i.e. \( \theta \) such that for all \( n, \theta_{2\mu-n} = 2\theta_{\mu} - \theta_n \). In the symmetric case, we show in the appendix that at any LFE, \( G(\Delta^c) = \{ \theta_{\mu} \} \). Since the second stage equilibrium is socially optimal, we can focus on the first stage incentives in isolation. Hence, if anything, the symmetry assumption should lessen the bias towards excessive rigidity. As we will see, the main determinant of the equilibrium degree of decentralization is then the polarization of preferences:

**Definition 4** The degree of polarization \( \pi(\theta) \) of a profile of type \( \theta \) is given by

\[
\pi(\theta) = \frac{\text{med}_n |\theta_n - \theta_{\mu}|}{\frac{1}{N} \sum_n |\theta_n - \theta_{\mu}|}.
\]

A profile of preferences is maximally polarized if there exists \( \delta > 0 \) such that for \( n \in \{1, \ldots, \lceil \frac{N+1}{4} \rceil \} \), \( \theta_n = \theta_{\mu} - \delta \), for \( n \in \{n - \lceil \frac{N+1}{4} \rceil + 1, \ldots, n\} \), \( \theta_n = \theta_{\mu} + \delta \) and \( \theta_n = \theta_{\mu} \) otherwise.\(^{13}\)

In words, a profile of type is less polarized the more homogeneous the core states are (i.e. states \( p \) such that \( |\theta_p - \theta_{\mu}| < \text{med}_n |\theta_n - \theta_{\mu}| \)) and more extreme the peripheral states are (i.e. states \( c \) such that \( |\theta_c - \theta_{\mu}| > \text{med}_n |\theta_n - \theta_{\mu}| \)). One can show that a maximally polarized preferences profile indeed maximizes \( \pi(\theta) \) among symmetric profiles of type, and it is a global maximizer when \( \frac{N+1}{4} \) is an integer. Notice that polarization is orthogonal to

\(^{13}\lceil \frac{N+1}{4} \rceil \) denotes the smallest integer weakly greater than \( \frac{N+1}{4} \).
the notion of heterogeneity since any linear transformation of \( \theta \) leaves \( \pi(\theta) \) unchanged.\(^{14}\)

**Proposition 6** Suppose \( \theta \) is symmetric and let \((\Gamma^e, \Delta^e)\) be a LFE,

(i) if \( \pi(\theta) \leq \frac{1+\frac{3}{4} \beta}{1+2\beta} \), then for all \( \beta \), \( \Delta^e < \Delta^* \) where \( \Delta^* \) is the socially optimal degree of local discretion

(ii) if \( \theta \) is maximally polarized and \( N > 3 \), then for \( \beta \) sufficiently large, \( \Delta^e > \Delta^* \).

(iii) for any \( \beta \), if \( \theta \) is maximally polarized, any LFE is socially better and majority preferred to decentralization, while if \( \pi(\theta) \) is sufficiently small, any LFE is socially worse but still majority preferred to decentralization.

Point (i) implies that if \( \pi(\theta) < \frac{3}{4} \), then for all \( \beta \), the equilibrium degree of local discretion is too small: a low degree of polarization exacerbates the federal encroachment problem.\(^{15}\) Conversely, from point (ii), when preferences are very polarized and the need for coordination is greater, the equilibrium entails an inefficiently low degree of policy harmonization. The latter result, together with proposition 5, suggests that in this case, the degree of decentralization \( \Delta \) is insufficiently sensitive to the magnitude of externalities. Point (iii) is reminiscent of the “delimitation problem” highlighted by Cremer and Palfrey (2000, 2006): a majority of voters can support a federal intervention even when it is not socially beneficial. Observe that the federal

\(^{14}\)See Esteban and Ray 1994 for a related notion of polarization.

\(^{15}\)For instance, as \( N \to \infty \), if each \( \theta_n \) is normally distributed \( \pi(\theta) \to 0.84 > 3/4 \), and if it is chi square distributed on each side of \( \theta_{\mu} \), \( \pi(\theta) \to 0.46 < 3/4 \). If \( \theta \) is maximally polarized, \( \pi(\theta) = \frac{N}{2(1+\mu)} \to 2 \).
intervention can be socially detrimental even if externalities are arbitrarily large.

Finally, proposition 3 and 6 show that the welfare effect of federal directives under direct democracy depends on the skewness and the degree of polarization of preferences. Symmetric preferences minimize the perverse incentives of the median voter at the second stage while polarized preferences counter the federal encroachment problem at the first stage.

6 The Case of Triadic Federations

In this section, we restrict attention to triadic profiles of types, i.e. $\theta$ such that $\theta_1 = \ldots = \theta_\lambda$, $\theta_{N-\rho+1} = \ldots = \theta_N$, and $\theta_m = \theta_\mu$ for $\lambda < m \leq N - \rho$ with $\theta_1 < \theta_\mu < \theta_N$ and $\rho + \lambda > \frac{N+1}{2}$.

**Proposition 7** If $\theta$ is triadic, the unique LFE $(\Gamma^e, \Delta^e)$ is a global federal equilibrium. It is given by

$$
\Delta^e = \min \{ D_\lambda^o, D_\rho^o, D_\lambda, D_\rho \}, \quad \Gamma^e = \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta^e.
$$

where

$$
D_\lambda^o = \sqrt{\frac{(\lambda + \rho)(N + \beta \lambda + \beta \rho)}{(\beta + 1)^2 \rho (N + \beta \rho)} \frac{\theta_\mu - \theta_1}{2}}, \quad D_\rho^o = \sqrt{\frac{(\lambda + \rho)(N + \beta \lambda + \beta \rho)}{(\beta + 1)^2 \lambda (N + \beta \lambda)} \frac{\theta_N - \theta_\mu}{2}},
$$

$$
D_\lambda \equiv \frac{\lambda + \rho}{\rho + \frac{\rho}{N} (\lambda^2 + \rho \lambda + N \rho)} \frac{\theta_\mu - \theta_1}{2}, \quad D_\rho \equiv \frac{\lambda + \rho}{\lambda + \frac{\rho}{N} (\rho^2 + \lambda \rho + N \lambda)} \frac{\theta_N - \theta_\mu}{2}.
$$

The threshold $\min \{ D_\lambda^o, D_\rho^o \}$ is the degree of local discretion below which both rightist and leftist states are constrained. From (7), $\Delta > \min \{ D_\lambda^o, D_\rho^o \}$ cannot be a LFE because a majority of unconstrained states push for less
discretion. $D_\lambda$ and $D_\rho$ are the ideal degree of discretion of respectively leftist and rightist states conditional on being both constrained by the federal bounds.

When $\beta$ is sufficiently large, $D^o_\lambda > D_\lambda$ and $D^o_\rho > D_\rho$ so $\Delta^e = \min \{D_\lambda, D_\rho\}$. In this case, consider a perfectly symmetric situation in which $|\theta_\mu - \theta_1| = |\theta_N - \theta_\mu|$ and $\lambda = \rho$ (so $D_\lambda = D_\rho$). Keeping everything else constant, if rightist states become more numerous (i.e. $\rho$ increases), then simple calculus shows that $D_\lambda$ decreases: leftist states have greater incentive to reduce the leeway of rightist states. Likewise, if rightist states become more extreme (i.e. $|\theta_N - \theta_\mu|$ increases), then $D_\rho$ increases but $D_\lambda$ stays constant and so does $\Delta^e = \min \{D_\lambda, D_\rho\}$: the pivotal voters (leftist states) do not internalize the benefits of more flexibility for the rightist states. In line with proposition 3, these comparative statics suggest that the excessive rigidity of federal directives is exacerbated by the skewness of the distribution of preferences. The next proposition confirms this intuition but shows that quantitatively, these two forms of skewness do not have the same welfare consequences.

**Proposition 8** For $\beta$ sufficiently large,

(i) if $|\theta_\mu - \theta_1| = |\theta_N - \theta_\mu|$ and $\lambda = \rho$, the federal equilibrium Pareto dominates decentralization while some state strictly prefer decentralization whenever $\lambda |\theta_\mu - \theta_1| \neq \rho |\theta_N - \theta_\mu|$

(ii) if $|\theta_\mu - \theta_1| = |\theta_N - \theta_\mu|$ and $\lambda \neq \rho$ the federal equilibrium is socially better than decentralization,

(iii) if $\lambda = \rho$ and $|\theta_\mu - \theta_1| \neq |\theta_N - \theta_\mu|$, decentralization is socially better than decentralization,

(iv) for all triadic preferences profile, the federal equilibrium is majority pre-
ferred to decentralization.

The intuition behind points (i) and (ii) is the following: as externalities grow larger, $\Delta^c$ decreases and from (6), both federal bounds converge towards the ideal point of the median state $\theta_\mu$. If $|\theta_\mu - \theta_1| = |\theta_N - \theta_\mu|$, voters from the rightist and leftist states trade-off policy coordination and flexibility in a similar manner. Therefore, when voting on $\Delta$, the gap between the incentives of leftist and rightists states has limited welfare consequences. When furthermore $\lambda = \rho$, voters from both rightist and leftist states are pivotal at the first stage and can secure a jointly profitable equilibrium. Points (iii) and (iv) extends proposition 6 to triadic preferences profiles: federal directives can be socially worse than decentralization even when externalities are arbitrarily large, but they will always receive the support of a majority of voters.

7 Conclusion

This paper analyzes an intermediate regime between complete centralization and decentralization in which citizens vote at the federal level on a common discretionary space within which each state can chose its policy. This non-prescriptive federal mechanism imposes some degree of coordination while leaving state sovereignty within the federal bounds. By letting voters decide directly the degree of decentralization and the direction of harmonization, the contention is that direct democracy allows the voters to trade-off the need for policy coordination and flexibility according to their preferences and thus avoid the excessive rigidity and inadequacy of a top-down bureaucratic
organization. However, we show that in equilibrium, the discretionary space is negatively sensitive to voters preferences and too restrictive even when the gains from coordination are negligible.

The first bias is due to the incentives of the pivotal voters at the second stage to restrict the discretion of the states whose needs are most incompatible with their own ideal policy. The second bias is driven by the incentives of the pivotal voters at the first stage to reduce the degree of discretion in order to tilt the incentives of the pivotal voter at the second stage towards their own preferences. Since the pivotal voters at the two voting stages are different, the federal intervention may leave a majority of voters worse-off even when externalities are arbitrarily large. Since the model rules out any form of institutional inefficiencies, these biases are due solely to the conflict of interests between heterogeneous voters at the federal level.

This paper does not solve for the optimal voting mechanism for choosing the discretionary interval. An interesting direction for further research would be to characterize simple, robust democratic mechanisms to mitigate the biases exhibited in this paper. A quick look at natural alternatives suggests that there is no obvious way to fix these inefficiencies. If \( \Gamma \) and \( \Delta \) were voted upon simultaneously instead of sequentially, the equilibria would have qualitatively similar properties.\(^{16}\) Moreover, having distinct pivotal voters at different stages is necessary to avoid the tyranny of the median state. For instance, if the left bound \( L \) and right bound \( R \) of the discretionary interval

\(^{16}\)Observe first that if \( \Gamma \) and \( \Delta \) where voted upon jointly, there would be no Condorcet winner. If they were voted upon simultaneously and separately, one can easily see that propositions 2, 3 and 5 would still hold.
were voted upon sequentially in any order, one can show that the unique equilibrium is equivalent to unitarian centralization: \( L = R = \theta_{\mu} \). Opposing rightist to leftist at both voting stages allows the median voter to secure its most preferred policy vector.

8 Appendix

To simplify notations, in the following two lemmas, \( L \) (resp. \( R \)) refers to the left (resp. right) bound of the federal interval, i.e. \( L = \Gamma - \Delta \) and \( R = \Gamma + \Delta \). When no ambiguity arises, we shall use the notation \( x(\theta, L, R) \) for \( x(\theta, \Gamma = \frac{L+R}{2}, \Delta = \frac{R-L}{2}) \), and \( V_n(\theta, L, R) \) for \( V_n(x(\theta, \Gamma = \frac{L+R}{2}, \Delta = \frac{R-L}{2})) \).

The next lemma proves proposition 1.

Lemma 1 For all \( \Gamma, \Delta \), there is a unique state equilibrium characterized by

\[
\begin{align*}
n \in \{1, \ldots, l\} & \Rightarrow x_n = \Gamma - \Delta, \\
n \in \{l+1, \ldots, N-r\} & \Rightarrow x_n = \frac{\theta_n + \beta \bar{x}}{1 + \beta}, \\
n \in \{N-r+1, \ldots, N\} & \Rightarrow x_n = \Gamma + \Delta.
\end{align*}
\]

Each equilibrium policy \( x_n(\theta, L, R) \) weakly increasing in \( \theta, L, R \) and \( n \) and

\[
\begin{align*}
\frac{\partial \bar{x}(\theta, L, R)}{\partial L} &= \frac{(1 + \beta) l}{N + \beta (l + r)}, \\
\frac{\partial \bar{x}(\theta, L, R)}{\partial R} &= \frac{(1 + \beta) r}{N + \beta (l + r)}.
\end{align*}
\]

Proof. The unconstrained first-order condition for state \( n \) is given by \( x_n = \frac{\theta_n + \beta \bar{x}}{1 + \beta} \). So a state equilibrium \( x \) is a fixed point of the best-response function:

\[
\forall n, \ f_n(x) = \max \left( L, \min \left( R, \frac{\theta_n + \beta \bar{x}}{1 + \beta} \right) \right).
\]
where $\overline{x}$ is the mean of the vector $x$. In particular, for any two policy vector $x$ and $y$ and for any state $n$,

$$|f_n(y) - f_n(x)| \leq \frac{\beta}{1 + \beta} |\overline{y} - \overline{x}| \leq \frac{\beta}{1 + \beta} \max_n |y_n - x_n|.$$ 

Hence $f$ is a contraction for the sup norm and has a unique fixed point on $[L, R]^N$.

The fact that $x_n(\theta, L, R)$ is weakly increasing in $n$ is obvious from (10) and implies immediately (8). One can see from (10) that $f$ is weakly increasing in $x$ and in $(\theta, L, R)$. From Villas-Boas 97 (theorem 4), its fixed point $x(\theta, L, R)$ is weakly increasing in $(\theta, L, R)$.

Differentiating (8) with respect to $L$ (resp. $R$), summing over $n$ and solving for $\frac{\partial \tau_n(\theta, L, R)}{\partial L}$ (resp. $\frac{\partial \tau_n(\theta, L, R)}{\partial R}$), we get (9). □

**Lemma 2** For all $L \leq L', R \leq R'$,

$$V_n(\theta, L', R') - V_n(\theta, L, R)$$

is weakly increasing in $n$. In particular, simple majority rule preferences between $[L, R]$ and $[L', R']$ coincide with the preferences of the median state.

**Proof.** Observe that the induced utility function $V_n(\theta, L, R)$ of state $n$ can be written as

$$W_n(t_n, x) = \max_{y \in [L, R]} |y - t_n|^2 - \frac{\beta}{N} \sum_{m=1}^{N} |y - x_m|^2, \quad (11)$$

for $x = x(\theta, L, R)$ and $t_n = \theta_n$. Let $y^*(t_n, x)$ be the maximizer of (11). From the envelope theorem, for all $t_n \in \mathbb{R},$

$$\frac{\partial W_n}{\partial t_n}(t_n, x) = 2(y^*(t_n, x) - t_n). \quad (12)$$

25
Since (11) depends on $n$ only via $t_n$, (12) implies
\[
V_n(L, R) - V_m(L, R) = 2 \int_{\theta_n}^{\theta_m} (y^*(t, x(\theta, L, R)) - t) \, dt,
\]
which in turns implies
\[
[V_n(L', R') - V_n(L, R)] - [V_m(L', R') - V_m(L, R)]
= 2 \int_{\theta_n}^{\theta_m} (y^*(t, x(\theta, L', R')) - y^*(t, x(\theta, L, R))) \, dt.
\]
Observe that if $W(t_n, x, y)$ denotes the maximand of (11), for all $m$, $\frac{\partial^2 W}{\partial x \partial y_m} > 0$, hence $W$ has increasing differences in $(y, x_m)$. Hence, from Topkis theorem (see e.g. Milgrom and Shannon 1994, theorem 4), since (11) has a unique maximizer, $y^*$ is weakly increasing in $x$. Moreover, from lemma 1, $x(\theta, L', R') \geq x(\theta, L, R)$ for all $n$. So the integrand above is non negative, which proves that preferences over $[L, R]$ and $[L', R']$ satisfies the order-restriction property with respect to $\theta_n$, that is for all $L' \geq L$, $R' \geq R$ and $\theta_m \geq \theta_n$
\[
V_n(L', R') \geq V_n(L, R) \Rightarrow V_m(L', R') \geq V_m(L, R).
\]
From the representative voter theorem (see e.g. Grandmont 1978), simple majority rule preferences are exactly the preferences of the median voter. ■

8.1 Proofs in Subsection 4.1

8.1.1 Proof of proposition 2

Lemma 3 For all $L \leq R$, the state equilibrium $x(\theta, L, R)$ is equivalent to the unconstrained decentralized equilibrium $x_{dec}(t)$ as defined in (2) where $t$
is defined as
\[
\forall n, t_n = \max \left( t, \min \left( \bar{t}, \theta_n \right) \right),
\]
and
\[
t = (1 + \beta) L - \beta \pi(\theta, L, R),
\]
\[
\bar{t} = (1 + \beta) R - \beta \pi(\theta, L, R).
\]

At the decentralized equilibrium, for all \( m \neq n \),
\[
\frac{\partial V_{n}^{\text{dec}}}{\partial \theta_m} = \frac{2\beta}{N(1 + \beta)} (\theta_n - x_m^{\text{dec}}).
\]

**Proof.** The decentralization map \( \theta \rightarrow x^{\text{dec}}(\theta) \) defined in (2) is inverted as follows: \( \theta_n = (1 + \beta) x_n^{\text{dec}} - \beta x^{\text{dec}} \). By substituting \( x^{\text{dec}} = x(\theta, L, R) \) in the previous expression and by using (10), we get the “as if” profile of type \( t \).

The welfare of state \( n \) under decentralization \( V_{n}^{\text{dec}} \) is given by the program:
\[
V_{n}^{\text{dec}}(\theta) = \max_y \left| y - \theta_n \right|^2 - \frac{\beta}{N} \sum_{m=1}^{N} \left| y - x_m \right|^2,
\]
for \( x = x^{\text{dec}}(\theta) \). Using the envelope theorem, we have that for \( p \neq n \),
\[
\frac{\partial V_{n}^{\text{dec}}}{\partial \theta_p} = \frac{2\beta}{N} \sum_{m \neq p} (x_n^{\text{dec}} - x_m^{\text{dec}}) \frac{\partial x_m^{\text{dec}}}{\partial \theta_p} + \frac{2\beta}{N} (x_n^{\text{dec}} - x_p^{\text{dec}}) \frac{\partial x_p^{\text{dec}}}{\partial \theta_p},
\]
\[
= \frac{2\beta}{N(1 + \beta)} \left[ \sum_{m \neq p} (x_n^{\text{dec}} - x_m^{\text{dec}}) \frac{\beta}{N} + (x_n^{\text{dec}} - x_p^{\text{dec}}) \left( 1 + \frac{\beta}{N} \right) \right],
\]
where \( \frac{\partial x_m^{\text{dec}}}{\partial \theta_p} \) and \( \frac{\partial x_p^{\text{dec}}}{\partial \theta_p} \) are derived from (2). The first-order condition of (15) gives
\[
\frac{\beta}{N} \sum_m (x_m^{\text{dec}} - x_m^{\text{dec}}) = \theta_n - x_n^{\text{dec}}.
\]
Substituting (17) in (16), we get (14). □

Since $\Gamma - \Delta$ and $\Gamma + \Delta$ are increasing in $\Gamma$, lemma 2 implies that the majority preferences on $\Gamma$ are exactly the preferences of the median state, which proves the first part of proposition 2.

Using the notation of lemma 3, for all $\theta, \Delta$ and for almost all $\Gamma$ at which the median state is unconstrained,\(^{17}\)

$$\frac{\partial V_{\mu}}{\partial \Gamma} (\theta, \Gamma, \Delta) = \frac{\partial \left[ V_{\mu}^{\text{dec}} \left( \left( \max \left( \frac{1}{N} \min \left( \frac{\Gamma}{\theta_n} \right) \right) \right)_n \right) \right]}{\partial \Gamma} = \sum_{n > N-r} \frac{\partial V_{\mu}^{\text{dec}}}{\partial \theta_n} \frac{\partial \bar{\pi}}{\partial \Gamma} + \sum_{n \leq l} \frac{\partial V_{\mu}^{\text{dec}}}{\partial \theta_n} \frac{\partial \bar{t}}{\partial \Gamma}. \tag{18}$$

Substituting (9) in (13),

$$\frac{\partial \bar{\pi}}{\partial \Gamma} = \frac{\partial \bar{t}}{\partial \Gamma} = 1 + \beta - \beta \frac{(1 + \beta)(l + r)}{N + \beta(l + r)} = \frac{(1 + \beta)N}{N + \beta(l + r)}. \tag{19}$$

Substituting (14) and (19) in (18), we have that whenever the median state is unconstrained, for almost all $\Gamma$,

$$\frac{\partial V_{\mu}}{\partial \Gamma} (\theta, \Gamma, \Delta) = \frac{\beta}{N(1 + \beta)} \left[ l \left( \theta_{\mu} - \Gamma + \Delta \right) \frac{\partial \bar{t}}{\partial \Gamma} + r \left( \theta_{\mu} - \Gamma - \Delta \right) \frac{\partial \bar{\pi}}{\partial \Gamma} \right]. \tag{20}$$

$$\frac{\partial V_{\mu}}{\partial \Gamma} (\theta, \Gamma, \Delta) = \frac{\beta}{N(1 + \beta)} \left[ l \left( \frac{\theta_{\mu} - \Gamma + \Delta}{\theta_{\mu} - \Gamma - \Delta} \right) \right]. \tag{21}$$

If the policy of the median state was not subject to the federal constraints, we see from (21) that for almost all $\Gamma$, $\frac{\partial V_{\mu}}{\partial \Gamma} > 0$ (resp. < 0) whenever $\Gamma + \Delta < \theta_{\mu}$ (resp. $\Gamma - \Delta > \theta_{\mu}$). The same is true a fortiori when the median state is subject to the federal constraints.\(^{18}\) So the most preferred $\Gamma$ of the voters of

\(^{17}\) $V_{\mu}$ may not be differentiable at the (finitely many) $\Gamma$ at which $l (\theta, \Gamma, \Delta)$ or $r (\theta, \Gamma, \Delta)$ are not locally constant in $\Gamma$.

\(^{18}\) For instance, if the constraint $x_{\mu} \leq \Gamma + \Delta$ is binding, then necessarily $\Gamma + \Delta < \theta_{\mu}$ so (21) is positive. If we apply the envelope theorem to the program (15) for $y = x (1, \Gamma, \Delta)$
the median state is such that $\theta_\mu \in [\Gamma - \Delta, \Gamma + \Delta]$. Since their most preferred policy is given by $\frac{\theta_\mu + \beta \pi(\theta, \Gamma, \Delta)}{1 + \beta}$ and $\pi(\theta, \Gamma, \Delta) \in [\Gamma - \Delta, \Gamma + \Delta]$, at their most preferred $\Gamma$, the policy of the median state is unconstrained.

Finally, to prove (6), observe that $l(\theta, \Gamma, \Delta)$ is weakly increasing in $\Gamma$ while $r(\theta, \Gamma, \Delta)$ is weakly decreasing in $\Gamma$. So if the expression in bracket in (21) changes sign at a point of discontinuity, it can only be an upward jump discontinuity. Hence, if $V_\mu$ is not differentiable at some $\Gamma^\circ$, $V_\mu$ has a convex kink at $\Gamma^\circ$. Hence, the local maxima of $V_\mu$ can only happen at a differentiability point. In particular, the most preferred $\Gamma$ of the voters of the median state are the zeros of the right hand-side of (21), which implies $\Gamma = \theta_\mu + \frac{l-r}{l+r} \Delta$.

8.1.2 Proof of proposition 3

From proposition 2, for all $\Gamma \in G(\theta, \Delta)$, $\Gamma$ maximizes $V_\mu (\Gamma, \Delta)$. So from theorem 4 in Milgrom and Roberts (1994), to prove that $G(\theta, \Delta)$ is weakly decreasing, it suffices to show that $\frac{\partial V_\mu}{\partial n}$ is weakly decreasing in $\theta_n$ for $n \neq \mu$. Furthermore, from proposition 2, it suffices to show this at any $(\theta, \Gamma, S)$ such that the median state is unconstrained.

From (21), for all $n \neq \mu$, $\frac{\partial V_\mu}{\partial \theta}$ depends on $\theta_n$ only though $r$ or $l$. Therefore, $\frac{\partial V_\mu}{\partial \theta}$ is weakly decreasing in $\theta_n$ if (21) is decreasing in $r$ and increasing in $l$. If we call $A(l, r + 1) - A(l, r) = \beta \frac{N (\theta_\mu - \Gamma - \Delta) - 2 \beta l \Delta}{(N + \beta (l + r)) (N + \beta (l + r + 1))}$, \hfill (22)

as in lemma 3 under the additional constraint $x_\mu \leq \Gamma + \Delta$, $\frac{\partial V_\mu}{\partial \theta}$ is given by (21) plus the Lagrange multiplier $\lambda$, which is positive. Hence, $\frac{\partial V_\mu}{\partial \theta}$ is greater than if the median state was not subject to the federal constraint, i.e. greater than (21) and thus positive as well.
For a fixed \((\Gamma, \Delta)\) and \(\theta_{-\mu}\), since \(l\) is decreasing in \(\theta_{\mu}\), the numerator of (22) is increasing in \(\theta_{\mu}\). Let \(t_{\mu}\) be the largest \(\theta_{\mu}\) at which state \(\mu\) is unconstrained. Since the median state is unconstrained in equilibrium, from what precedes, to prove that \(A(l, r)\) is weakly decreasing in \(r\), it suffices to prove that the numerator of (22) is negative at \(\theta_{\mu} = t_{\mu}\). From (10), at \(\theta_{\mu} = t_{\mu}\), \(x_{\mu} = \frac{\theta_{\mu} + \beta \varpi}{1 + \beta}\) and by definition of \(t_{\mu}\), \(x_{\mu} = \Gamma + \Delta\), so

\[
t_{\mu} - \Gamma - \Delta = \beta (\Gamma + \Delta - \varpi). \tag{23}
\]

Moreover, we always have that

\[
\varpi \leq \frac{(N - l) (\Gamma + \Delta) + l (\Gamma - \Delta)}{N}. \tag{24}
\]

Substituting (24) in (23), we get \(t_{\mu} - \Gamma - \Delta \leq 2\beta \frac{\varpi}{N}\), which shows that (22) is negative. Similar algebra shows that \(A(l, r)\) is decreasing in \(l\) and therefore \(G(\theta, \Delta)\) is weakly decreasing in \(\theta_{-\mu}\).

To prove the second point of proposition 3, suppose by contradiction that \(t\) is a profile of type which is skewed to the right, \(G \in G_t(\Delta)\) and \(G > t_{\mu}\). Let \(t^*\) be such that for all \(n\), \(t^n_\mu = t^{n+2(nu-t_\nu-n)}\). By construction, \(t^*\) is symmetric around \(t_{\mu}\) (i.e. for all \(n\), \(t^n_\mu = 2t^* - t^{2n_\mu-n}\) and for all \(n\), \(t^n_\mu \leq t_n\), the inequality being strict for \(n \neq \mu\). By symmetry, \(G(t^*, \Delta)\) is symmetric around \(t_{\mu}\) so there exists \(G' \in G_{t^*}(\Delta)\) such that \(G' \leq t_{\mu}\). Since \(G' < G\) and \(t^* \leq t\) from the first part of the proposition, \(G' \in G_t(\Delta)\) and \(G \in G_{t^*}(\Delta)\). By symmetry, \(2\theta_{\mu} - G \in G_{t^*}(\Delta)\) and since \(2\theta_{\mu} - G \leq G\), the first part of the proposition implies as well that \(2\theta_{\mu} - G \in G_t(\Delta)\).

Since \(G\) and \(2\theta_{\mu} - G\) both belong to \(G_t(\Delta)\) and \(G_{t^*}(\Delta)\), the median state is indifferent between \(\Gamma = G\) and \(\Gamma = 2\theta_{\mu} - G\) at \(\theta = t\) and \(\theta = t^*\), i.e. \(\int_{2\theta_{\mu} - G}^{G} \frac{\partial \nu}{\partial \Gamma} (\theta, \Gamma, \Delta) d\Gamma = 0\) for \(\theta \in \{t, t^*\}\). From (21), \(\frac{\partial \nu}{\partial \mu}\) depends on
non median types \( \theta_{-\mu} \) only through \( l \) and \( r \). From what precedes, \( \frac{\partial V_\mu}{\partial \Gamma} \) is increasing in \( l \) and decreasing in \( r \). Since \( t^* \leq t \),

\[
\begin{align*}
l(t^*, \Gamma, \Delta) & \geq l(t, \Gamma, \Delta), \\
r(t^*, \Gamma, \Delta) & \leq r(t, \Gamma, \Delta),
\end{align*}
\]

which implies that for all \( \Gamma \),

\[
\frac{\partial V_\mu}{\partial \Gamma}(t^*, \Gamma, \Delta) \leq \frac{\partial V_\mu}{\partial \Gamma}(t, \Gamma, \Delta),
\]

the inequality being strict for all \( \Gamma \) such that (25) or (26) is strict. Hence, for the median state to be indifferent between \( \Gamma = G \) and \( \Gamma = 2\theta_{\mu} - G \) both at \( \theta = t \) and \( \theta = t^* \), it must be that both (25) and (26) hold with equality for all \( \Gamma \) in \([2\theta_{\mu} - G, G]\).

For all \( \theta \in \mathbb{R}^N \) and \( n < \mu \), let \( \Gamma_n(\theta, \Delta) \) be a value of \( \Gamma \) at which \( l \) jumps from \( n - 1 \) to \( n \) for some \( n < \mu \), that is \( \frac{\theta_n + \theta \pi(\theta, \Gamma_n, \Delta)}{1 + \beta} = \Gamma_n - \Delta \). Since both \( \pi \) and \( \theta_n \) are greater at \( \theta = t \) than at \( \theta = t^* \), strictly so for \( \theta_n \), \( \Gamma_n(t^*, \Delta) < \Gamma_n(t, \Delta) \). Likewise, one can show that the value of \( \Gamma \) at which \( r \) jumps from \( n \) to \( n+1 \) for some \( n > \mu \) is strictly greater at \( \theta = t \) than at \( \theta = t^* \). Therefore, for (25) and (26) to hold with equality for all \( \Gamma \) in \([2\theta_{\mu} - G, G]\), \( l \) and \( r \) must to be constant on \([2\theta_{\mu} - G, G]\). From proposition 3, \( 2\theta_{\mu} - G \) and \( G \) should both be equal to \( \theta_{\mu} + \frac{t-r}{t+r} \Delta \) so \( G = \theta_{\mu} \), a contradiction.

### 8.2 Proofs in Subsection 4.2

The following lemma proves the equations in (7):

**Lemma 4** The induced preferences of the voters of the median state on \( x(\theta, G(\theta, \Delta), \Delta) \) are decreasing in \( \Delta \). If \( G(\theta, \Delta) \) is single valued on \([D, D']\)
the same is true for the voters of the states \( \{l + 1, \ldots, N - r\} \). Moreover, for all \( \Delta \in [D, D'] \),

\[
\begin{align*}
\text{for } n & \leq l, \quad x_n(\theta, G(\theta, \Delta), \Delta) = \theta - \frac{2r}{l + r} \Delta, \\
\text{for } n & > N - r, \quad x_n(\theta, G(\theta, \Delta), \Delta) = \theta + \frac{2l}{l + r} \Delta, \\
\text{for } l & < n \leq N - r, \quad \frac{\partial [x_n(\theta, G(\theta, \Delta), \Delta)]}{\partial \Delta} = 0. 
\end{align*}
\]

**Proof.** Suppose that \( G(\theta, \Delta) \) is single valued on \([D, D']\). From (6), \( l \) and \( r \) are constant on \([D, D']\). The first two lines of (7) are immediate consequences of (6). For all \( m \in \{l + 1, \ldots, N - r\} \),

\[
x_m \left( \theta, \theta + \frac{l - r}{l + r} \Delta, \Delta \right) = \frac{\theta_m + \beta x_m(\theta, \theta + \frac{l - r}{l + r} \Delta, \Delta)}{1 + \beta}.
\]

From (9),

\[
\frac{\partial x_m}{\partial \Delta} = x_m = \frac{\beta}{1 + \beta} \frac{(1 + \beta)(l + r) \frac{l - r}{l + r} + (1 + \beta)(r - l)}{N + \beta(l + r)} = 0.
\]

From (1) and what precedes, the preferences of unconstrained states are decreasing on \([D, D']\).

From proposition 2, the median state is indifferent between all elements of \( G(\theta, \Delta) \). Let \( \Gamma(\Delta) \) be a selection of \( G(\theta, \Delta) \) such that \( l(\theta, \Gamma(\Delta), \Delta) \) and \( r(\theta, \Gamma(\Delta), \Delta) \) are piecewise constant on \([D, D']\). The preceding reasoning shows that \( U_\mu(x(\theta, \Gamma(\Delta), \Delta)) \) is decreasing on any continuity interval of \( \Gamma(\Delta) \), and by Berge maximum theorem, it is continuous at the points of non linearity, which proves the first part of the lemma. \( \blacksquare \)

We denote \( (\mathcal{L}, \mathcal{R})(\theta, \Delta) = \{(l, r)(\theta, \Gamma, \Delta) : \Gamma \in G(\theta, \Delta)\} \).
Lemma 5  There exists $D_1, ..., D_I$ (with the convention that $D_0 = 0$, $D_{I+1} = +\infty$) such that for all $i = 0, ..., I$, $(\mathcal{L}, \mathcal{R}) (\theta, \Delta)$ is constant on $]D^i, D^{i+1}[$. Moreover, $\Phi (\theta, \Delta) \equiv \{ \frac{r}{l+r} : (l, r) \in (\mathcal{L}, \mathcal{R}) (\theta, \Delta) \}$ is single valued and constant on $]D^i, D^{i+1}[$ and for all $i = 1, ..., I$

$$\lim_{\Delta \to D_i} \Phi (\theta, \Delta) < \lim_{\Delta \to D_i} \Phi (\theta, \Delta),$$

and for all $i = 0, ..., I$,

$$(\mathcal{L}, \mathcal{R}) (\theta, D^i) \supset \left\{ \lim_{\Delta \to D^i} (\mathcal{L}, \mathcal{R}) (\theta, \Delta), \lim_{\Delta \to D^i} (\mathcal{L}, \mathcal{R}) (\theta, \Delta) \right\}.$$  \hfill (29)

Proof. Let $\Delta^\prime, \Delta^\prime\prime$ and $(l, r)$ be such for all $\Delta \in [\Delta^\prime, \Delta^\prime\prime]$, if $\Gamma = \theta_\mu + \frac{r}{l+r} \Delta$, $(l, r) = (l, r) (\theta_\mu, \Gamma, \Delta)$. From lemma 4, $\frac{\partial x_n (\theta_\mu, \theta_\mu + \frac{r}{l+r} \Delta)}{\partial \Delta} = 0$ for all $n \in \{l+1, N-r\}$ so

$$\frac{\partial}{\partial \Delta} [U_\mu \left( x (\theta_\mu, \theta_\mu + \frac{r}{l+r} \Delta, \Delta) \right)] = -\frac{4rl}{l+r} \frac{\beta}{N} \left( x_\mu - \theta_\mu + \frac{2r}{l+r} \Delta \right) + \frac{4lr}{l+r} \frac{\beta}{N} \left( x_\mu - \theta_\mu - \frac{2l}{l+r} \Delta \right) = -\frac{8\beta rl}{N(l+r)} \Delta.$$  

Hence, for all $(l, r)$ as described above, $U_\mu \left( x (\theta_\mu, \theta_\mu + \frac{r}{l+r} \Delta, \Delta) \right)$ is quadratic in $\Delta$ with a slope proportional to $-\frac{r}{l+r} \Delta$. This shows that $\Phi (\theta, \Delta)$ must be weakly decreasing in $\Delta$ in the strong set order sense. Since $\frac{r}{l+r}$ can take only a finite number of values, $\Phi$ has at most a finite number of point of discontinuity. Moreover, from what precedes, in between any two points of discontinuity, $\Phi$ must be single valued, which proves (28). From Berge maximum theorem, $\mathcal{G} (\theta, \Delta)$ is upper hemi-continuous in $\Delta$, and from (6), so must be $(\mathcal{L}, \mathcal{R}) (\theta, D^i)$, which implies (29). \hfill \blacksquare
8.2.1 Proof of proposition 4

Lemma 6 If $(\Gamma^e, \Delta^e)$ is a LFE, then $G(\theta, \Delta)$ is single valued on $[\Delta^e - \varepsilon, \Delta^e]$ for some positive $\varepsilon$ and $\Gamma^e = \lim_{\Delta^e \to \Delta^e} G(\theta, \Delta)$.

Proof. Suppose that $G(\theta, \Delta)$ is not single valued on $[\Delta^e - \varepsilon, \Delta^e]$. From (29), $G(\theta, \Delta^e)$ is not single-valued. Let $\Gamma^o \in G(\theta, \Delta^e)$ such that $\Gamma^o \neq \Gamma^e$, say $\Gamma^o < \Gamma^e$ for concreteness. From proposition 2, $\Gamma^o = \frac{l^o - r^o}{l^o + r^o} \Delta^e$ where $(l^o, r^o) = (l, r) (\theta, \Gamma^o, \Delta^e)$ and the voters from the median state are indifferent between $(\Gamma^o, \Delta^e)$ and $(\Gamma^e, \Delta^e)$. From lemma 2, the voters from leftist states strictly prefer $(\Gamma^o, \Delta^e)$ to $(\Gamma^e, \Delta^e)$. By continuity, they strictly prefer $(\frac{l^o - r^o}{l^o + r^o} (\Delta^e - \varepsilon), \Delta^e - \varepsilon)$ for $\varepsilon$ sufficiently small and positive. From lemma 4, so do the voters from the median state. From lemma 5, $(\frac{l^o - r^o}{l^o + r^o} (\Delta^e - \varepsilon))$ belongs to $G(\theta, \Delta^e - \varepsilon)$, hence $(\Gamma^e, \Delta^e)$ cannot be a LFE.

If $G(\theta, \Delta)$ is single valued on $[\Delta^e - \varepsilon, \Delta^e]$ for some positive $\varepsilon$ and $\Gamma^e \neq \lim_{\Delta^e \to \Delta^e} G(\theta, \Delta)$, the same reasoning shows that a majority of voters will prefer $(\frac{l^o - r^o}{l^o + r^o} (\Delta^e - \varepsilon), \Delta^e - \varepsilon)$ to $(\Gamma^e, \Delta^e)$ where $\Gamma^o = \lim_{\Delta^e \to \Delta^e} G(\theta, \Delta)$ and $(l^o, r^o) = (l, r) (\theta, \Gamma^o, \Delta^e)$. ■

The following lemma uses the notations of lemma 5.

Lemma 7 A pair $(\Gamma^e, \Delta^e)$ is a LFE if and only if there exists $i > 0$ such that one of the following is true:

(i) $\Delta^e \in [D^{i-1}, D^i]$, $G(\theta, \Delta)$ is single-valued on $[D^{i-1}, D^i]$, $G(\theta, \Delta^e) = \{\Gamma^e\}$ and majority preferences on $\Delta$ are single peaked on $[D^{i-1}, D^i]$ with a peak at $\Delta^e$.

(ii) $\Delta^e = D_i$, $G(\theta, \Delta)$ is single-valued on $[D^{i-1}, D^i]$, $\Gamma^e = \lim_{D^i \to D_i} G(\theta, \Delta)$ and majority preferences on $\Delta$ are increasing on $[D^{i-1}, D^i]$. 34
Proof. If (i) is satisfied, \((\Gamma^e, \Delta^e)\) is clearly a LFE. Reciprocally, suppose that \((\Gamma^e, \Delta^e)\) is a LFE and \(\Delta^e \in ]D^{i-1}, D^i[\). From lemma 6 and 5, \(G(\theta, \Delta)\) is single-valued on \(]D^{i-1}, D^i[\). Therefore, the induced preferences of all voters on \(x(\theta, G(\theta, \Delta), \Delta)\) are well-defined, quadratic and concave. Therefore, majority preferences are quasi-concave on \(]D^{i-1}, D^i[\), and the conclusion follows from the median voter theorem.

If \(\Delta^e = D^i\), from lemma 6 \(\Gamma^e = \lim_{\Delta^e} G(\theta, \Delta)\) and \(G(\theta, \Delta)\) is single-valued on \(]D^{i-1}, D^i[\). Suppose majority preferences are not increasing on \(]D^{i-1}, D^i[\). As argued earlier, majority preferences are quasi-concave, so they must be decreasing on \(]D^i - \epsilon, D^i[\) for some \(\epsilon > 0\). Since \(\Gamma^e = \lim_{\Delta^e} G(\theta, \Delta)\), \((\Gamma^e, \Delta^e)\) cannot be a LFE.

Reciprocally, suppose (ii) is satisfied. Then since \(\Gamma^e = \lim_{\Delta^e} G(\theta, \Delta)\), for some \(\epsilon > 0\), for all \(\Delta \in ]\Delta^e - \epsilon, \Delta^e[, (G(\theta, \Delta), \Delta)\) is not majority preferred to \((\Gamma^e, \Delta^e)\). For all \(\Gamma \in G(\theta, \Delta^e)\), the median voter is indifferent between \((\Gamma^e, \Delta^e)\) and \((\Gamma, \Delta^e)\) so from lemma 2, voters from rightist and leftist states have opposite preferences, so \((\Gamma, \Delta^e)\) is not majority preferred to \((\Gamma^e, \Delta^e)\). Finally, let \(\Delta^k \setminus \Delta^e\) and \(\Gamma^k \in G(\theta, \Delta^k)\). One can restrict attention to sequences such that \((l, r)(\theta, \Gamma^k, \Delta^k) = (l^o, r^o)\) for all \(k\) and some \((l^o, r^o)\). From (28), \((l^o, r^o) \neq (l, r) (\theta, \Gamma^e, \Delta^e)\). From (6), \(\Gamma^o \equiv \lim \Gamma^k\) exists and from what precedes, \(\Gamma^o \neq \Gamma^e\). Suppose to fix ideas that \(\Gamma^e < \Gamma^o\). For \(k\) sufficiently large, \(\Gamma^e + \Delta^e < \Gamma^k + \Delta^k\). From lemma 4, the median state strictly prefer \((\Gamma^e, \Delta^e)\) to \((\Gamma^k, \Delta^k)\) so from lemma 2 and what precedes, \((\Gamma^e, \Delta^e)\) is strictly preferred to \((\Gamma^k, \Delta^k)\) by leftist states. So \((\Gamma^e, \Delta^e)\) is preferred by simple majority rule to \((\Gamma^k, \Delta^k)\) for \(k\) sufficiently large. ■
We now prove proposition 4. Using the notations of lemma 5, let \( j \) be the lowest integer such that for all \( i \geq j \), either \( G(\theta, \Delta) \) is not single valued on \( ]D^i, D^{i+1}[ \) or majority preferences are weakly decreasing on \( ]D^i, D^{i+1}[ \).

As \( \Delta \to \infty \), either \( l = 0 \) and \( r > 0 \) or \( r = 0 \) and \( l > 0 \) so from (6), for all \( \Delta \in ]D^I, +\infty[ \), \( G(\theta, \Delta) \subset \{ \theta_\mu + \Delta, \theta_\mu - \Delta \} \). If \( G(\theta, \Delta) \) is single-valued, from what precedes, for all \( n \), \( V_n(\theta, G(\theta, \Delta), \Delta) \) is constant on \( ]D^I, +\infty[ \) so \( j \leq I \). As \( \Delta \to 0 \), from (6), all states \( n \) such that \( \theta_n \neq \theta_\mu \) must be constrained and \( G(\theta, \Delta) \) is single valued. In this case, one can easily see from (1) that \( V_n(\theta, G(\theta, \Delta), \Delta) \) must be increasing in a neighborhood of \( \Delta = 0 \). Since a majority of states have a type different from \( \theta_\mu \), \( \Delta = 0 \) cannot be a LFE and \( j > 0 \).

We now prove the existence of a LFE. By definition of \( j \), \( G(\theta, \Delta) \) is single valued on \( ]D^{j-1}, D^j[ \) and the induced preferences of all voters on \( x(\theta, G(\theta, \Delta), \Delta) \) are well-defined and quasi-concave on \( ]D^{j-1}, D^j[ \). Therefore, majority preferences are quasi-concave. Suppose first that majority preferences are not weakly increasing on \( ]D^{j-1}, D^j[ \). By definition of \( j \), they are neither weakly decreasing on \( ]D^{j-1}, D^j[ \). Since they are quasi-concave, they must be single-peaked and the peak is a LFE. If majority preferences are weakly increasing on \( ]D^{j-1}, D^j[ \), from lemma 7, \( \lim_{\Delta \to D^j} G(\theta, \Delta) \) exists and \( \lim_{\Delta \to D^j} G(\theta, \Delta) \) is a LFE.

Finally, from lemma 7, at any LFE \( (\Gamma^e, \Delta^e) \), a majority of voters have preferences which are weakly increasing in \( \Delta \) to the left of \( \Delta^e \). From lemma 4, these must be voters from states which are constrained by the federal bounds,

\[ -\frac{\partial}{\partial \Delta} \sum_{m \neq n} |x_n - x_m|^2 \text{ is zero.} \]
which proves that a majority of states must be constrained at $(\Gamma^e, \Delta^e)$.

8.3 Proofs in Section 5

8.3.1 Proof of proposition 5

We introduce the following notation: for $\Delta > 0$, $n > \mu$ (resp. $n < \mu$), let $\Gamma_n$ be such that for $\Gamma < \Gamma_n$, the bound $\Gamma + \Delta$ (resp. $\Gamma - \Delta$) is binding for state $n$ at $(\Gamma, \Delta)$ while it is not binding (resp. binding) for $\Gamma > \Gamma_n$.

**Lemma 8** There exists $c > 0$ such that for all $\Delta$, all $\Gamma \in \mathcal{G}(\theta, \Delta)$ and all $n \neq \mu$, we have $|\Gamma - \Gamma_n| \geq c\Delta + O(\beta)$ as $\beta \to 0$.

Proof. Using the notation of lemma 5, let $(l, r) \in (\mathcal{L}, \mathcal{R})(\theta, \Delta)$, that is $\Gamma = \theta + \frac{l-r}{l+r-1} \Delta \in \mathcal{G}(\theta, \Delta)$. From proposition 2,

$$U_\mu \left( x \left( \theta, \theta + \frac{l-r+1}{l+r-1} \Delta, \Delta \right) \right) - U_\mu (x(\theta, \Gamma, \Delta)) \leq 0. \quad (30)$$

Using the notation of the lemma, let $\Gamma^o$ be the closest $\Gamma_n$ to $\Gamma$. Suppose to fix ideas that $\Gamma^o = \Gamma_{N-r+1}$ and in particular $\Gamma^o > \theta + \frac{l-r}{l+r-1} \Delta$ (the proof in the three other cases $\Gamma_{N-r-1}$, $\Gamma_l$ and $\Gamma_{l+1}$ is identical). If $\Gamma^o > \theta + \frac{l-r+1}{l+r-1} \Delta$, then the lemma holds since one can take $c = \frac{l-r+1}{l+r-1} - \frac{l-r}{l+r}$. If not, then using $f = O(\beta)$ as $\Delta \to 0$ means that there exists $C > 0$ and $\varepsilon > 0$ such that for all $\beta \in [0, \varepsilon]$, $|f(\varepsilon)| \leq C\beta$.

---

20 The notation $f = O(\beta)$ as $\Delta \to 0$ means that there exists $C > 0$ and $\varepsilon > 0$ such that for all $\beta \in [0, \varepsilon]$, $|f(\varepsilon)| \leq C\beta$. 

37
(21), the left hand-side of (30) can be rewritten as

\[
\int_{\Gamma_0}^{\Gamma_0} \frac{\partial U_\mu}{\partial \Gamma} (\theta, \Gamma, \Delta) d\Gamma + \int_{\Gamma_0}^{\theta_0 + \frac{l-r-1}{l+r} \Delta} \frac{\partial U_\mu}{\partial \Gamma} (\theta, \Gamma, \Delta) d\Gamma
\]

\[
= \frac{\beta}{N + \beta (l + r)} \int_{\Gamma = \theta_0 + \frac{l-r-1}{l+r} \Delta}^{\Gamma_0} \left[ l (\theta_\mu - \Gamma + \Delta) + r (\theta_\mu - \Gamma - \Delta) \right] d\Gamma
\]

\[
+ \frac{\beta}{N + \beta (l + r - 1)} \int_{\Gamma = \Gamma_0}^{\theta_0 + \frac{l-r-1}{l+r} \Delta} \left[ l (\theta_\mu - \Gamma + \Delta) + (r - 1) (\theta_\mu - \Gamma - \Delta) \right] d\Gamma.
\]

Taking the limit as \( \beta \to 0 \), the above integrals can be approximated as follows

\[
\int_{\Gamma = \theta_0 + \frac{l-r-1}{l+r} \Delta}^{\Gamma_0} \left( \theta_\mu + \frac{l-r-1}{l+r} - \Gamma \right) d\Gamma
\]

\[
= \frac{\beta (l + r)}{N} \int_{\Gamma = \theta_0 + \frac{l-r-1}{l+r} \Delta}^{\Gamma_0} \left( \theta_\mu - \Gamma - \Delta \right) d\Gamma + O (\beta^2),
\]

\[
= - \frac{\beta (l + r)}{2N} \left( \frac{2l \Delta}{(l + r - 1)(l + r)} \right)^2 + O (\beta^2),
\]

(31)

where \( O (\beta^2) \) can be chosen independent of \( \Delta \) (for \( \Delta \in [0, \theta_N - \theta_1] \)). Rearranging (31), (30) implies

\[
-(l + r) \left( \frac{2l \Delta}{(l + r - 1)(l + r)} \right)^2 + \left( \frac{2l \Delta}{l + r - 1} \right)^2 \leq (\theta_\mu - \Gamma_0 - \Delta)^2 + O (\beta)
\]

which is equivalent to

\[
\frac{4l^2}{(l + r - 1)(l + r)} \Delta^2 \leq (\Gamma_0 - \theta_\mu + \Delta)^2 + O (\beta).
\]

(32)

\(^{21}\)Observe that in (31), we implicitly assumed that the number of constrained states \((l', r')\) for \( \Gamma \in \left[ \Gamma_0, \frac{l-r-1}{l+r} \Delta \right] \) is constant and equal to \((l, r + 1)\). It may be that \( l' > l \) or \( r' < r - 1 \) as \( \Gamma \to \frac{l-r-1}{l+r} \Delta \). One can easily check that since \( \frac{4l^2}{(l + r - 1)(l + r)} \) is increasing in \( l \) and decreasing in \( r \) whenever \( a > 0 \) and \( b < 0 \), the left hand-side of (30) is still bounded below by (31), which is all we need for the proof.
Since state $N - r + 1$ is constrained at $\Gamma = \theta_{\mu} + \frac{l - r}{l + r} \Delta$, $\Gamma^o > \theta_{\mu} + \frac{l - r}{l + r} \Delta$ so $\Gamma^o - \theta_{\mu} + \Delta$ is positive so (32) implies

$$\Gamma^o - \theta_{\mu} + \Delta + O(\beta),$$

which can be rewritten as

$$\Gamma^o - \left(\theta_{\mu} + \frac{l - r}{l + r} \Delta\right) \geq \frac{2l}{l + r} \left(\sqrt{\frac{l + r}{l + r - 1} - 1}\right) \Delta + O(\beta),$$

which completes the proof. 

Using the notation of lemma 8, for all $n \neq \mu$, $\Gamma_n$ is characterized by $\Gamma_n \pm \Delta = \frac{\theta_n + \beta \pi(\mu, \Gamma_n, \Delta)}{1 + \beta}$ where $\pm$ refers to whether $n < \mu$ or $n > \mu$. So for all $\Delta > 0$, as $\beta \to 0$, $\Gamma_n = \theta_{n} \pm \Delta + O(\beta)$. Hence, lemma 8 implies that for all $\Delta > 0$, for all $\Gamma \in \mathcal{G}(\theta, \Delta)$ and for all constrained states $n$ at $(\Gamma, \Delta)$, $|\Gamma(\Delta) \pm \Delta - \theta_{n}| \geq c\Delta + O(\beta)$. So for all $\Gamma \in \mathcal{G}(\theta, \Delta)$ and for $\beta$ sufficiently small, it should be clear from (1) that fixing $\Gamma$, all constrained states $n$ at $(\Gamma, \Delta)$ locally prefer a larger $\Delta$. Moreover, $V_n(\theta, \Gamma, \Delta) \leq -(c\Delta)^2 + O(\beta)$ for some $c > 0$ while from (1) and (2), under decentralization, $U_n^{dec} = O(\beta)$. From proposition 4, a majority of states are constrained at any LFE so decentralization is majority preferred and socially better than any LFE.\footnote{It should be clear from the proof of proposition 4 that at any LFE, $\Delta$ is bounded away from 0 as $\beta \to 0$.}

### 8.3.2 Proof of proposition 6

If $(\Delta^e, \Gamma^e)$ is a LFE, from lemma 6, $G(\theta, \Delta)$ must be single-valued on $|\Delta^e - \varepsilon, \Delta^e|$ for some $\varepsilon > 0$. Since $\theta$ is symmetric, this means that $G(\theta, \Delta) = \{\theta_{\mu}\}$ so $\Gamma^e = \{\theta_{\mu}\}$. Consider the family of policy vector $x(\theta, \theta_{\mu}, \Delta)$ and let $c(\Delta)$ be
the number of states constrained by the left bound at \((\Gamma, \Delta) = (\theta, \Delta)\). By symmetry, the number of states constrained by the right bound is \(c(\Delta)\) as well. From (1), for \(n \leq c(\Delta)\),

\[
U_n(x(\theta, \theta, \Delta)) = -(\theta - \theta_n - \Delta)^2 - \frac{\beta}{N} \sum_{m=c+1}^{N-c} (\theta - x_m - \Delta)^2 - \frac{4c\Delta^2}{N}.
\]

Since \(\theta\) is symmetric, \(\sum_{m=c+1}^{N-c} (\theta - x_m - \Delta) = 0\). Moreover, from lemma 4, for almost all \(\Delta\), \(\frac{\partial x_n}{\partial \Delta} = 0\) so

\[
\frac{\partial U_n(x(\theta, \theta, \Delta))}{\partial \Delta} = 2\left[(\theta - \theta_n - \Delta) + \frac{\beta}{N} \sum_{m=c+1}^{N-c} (\theta - x_m - \Delta)\right] - \frac{8\beta \Delta}{N},
\]

\[
= 2\left[(\theta - \theta_n) - \left(1 + \frac{N + 2c}{N}\right) \Delta\right].
\]  
(33)

By symmetry for \(n > N - c(\Delta)\), for almost all \(\Delta\)

\[
\frac{\partial U_n(x(\theta, \theta, \Delta))}{\partial \Delta} = 2\left[(\theta_n - \theta) - \left(1 + \frac{N + 2c}{N}\right) \Delta\right].
\]  
(34)

For \(n \in \{c(\Delta) + 1, N - c(\Delta)\}\), from lemma 4, for almost all \(\Delta\), \(\frac{\partial x_n}{\partial \Delta} = 0\) so

\[
\frac{\partial U_n(x(\theta, \theta, \Delta))}{\partial \Delta} = -2c\frac{\beta}{N} (x_n - \theta + \Delta) + 2c\frac{\beta}{N} (x_n - \theta - \Delta) = -4c\frac{\beta}{N} \Delta.
\]  
(35)

If \(W(\Delta) = \sum_n U_n(x(\theta, \theta, \Delta))\), from (33), (34) and (35), simple algebra yields that for almost all \(\Delta\)

\[
\frac{\partial W}{\partial \Delta} = 2c\left[\frac{1}{c} \sum_{n\leq c} (\theta_n - \theta) + \frac{1}{c} \sum_{n>N-c} (\theta_n - \theta) - (2 + 4\beta) \Delta\right].
\]  
(36)

One can see from (36) that a discontinuity point of \(c(\Delta)\) corresponds to a convex kink of \(W\). So if \(\Delta^* = \arg\max_\Delta W(\Delta)\), \(W(\Delta)\) is differentiable at \(\Delta^*\) and

\[
\Delta^* = \frac{\frac{1}{2c(\Delta^*)} \left(\sum_{n\leq c(\Delta^*)} |\theta_n - \theta| + \sum_{n>N-c(\Delta^*)} |\theta_n - \theta|\right)}{1 + 2\beta},
\]  
(37)
which implies
\[
\Delta^* \geq \frac{1}{(N-1)} \sum_n |\theta_n - \theta_\mu| \frac{1 + 2\beta}{1 + 2\beta}.
\] (38)

As argued earlier, \( G(\theta, \Delta) = \{\theta_\mu\} \) on \( ]\Delta^* - \varepsilon, \Delta^*[ \) for some \( \varepsilon > 0 \). So \( x(\theta, G(\theta, \Delta), \Delta) \) and \( x(\theta, \theta_\mu, \Delta) \) coincide on \( ]\Delta^* - \varepsilon, \Delta^*[ \). Moreover, from lemma 7, majority preferences must be increasing on \( ]\Delta^* - \varepsilon', \Delta^*[ \) for some \( \varepsilon' > 0 \). That is, for a majority of states, \( \frac{\partial U_n(x(\theta, \theta_\mu, \Delta))}{\partial \Delta} (\Delta = \Delta^*) \geq 0 \). From (33) and (34), this implies that for at least \( \frac{N+1}{2} \) constrained states,
\[
|\theta_n - \theta_\mu| \geq \left(1 + \frac{\beta N + 2c(\Delta^*)N}{N} \right) \Delta^e.
\]

By symmetry, \( |\theta_n - \theta_\mu| \) is greater for the constrained states than for the unconstrained states, so from what precedes
\[
\Delta^e \leq \frac{\text{med}_n (|\theta_\mu - \theta_n|)}{1 + \beta \left(1 + 2\frac{c(\Delta^*)N}{N}\right)} \leq \frac{\text{med}_n (|\theta_\mu - \theta_n|)}{1 + \frac{3}{2} \beta},
\] (39)

which together with (38) proves the first point of proposition 6.

To prove the second point, notice that if \( \theta \) is maximally polarized, necessarily \( c(\Delta^*) = \left\lceil \frac{N+1}{4} \right\rceil \). So (37) implies \( \Delta^* = \frac{\delta}{1+2\beta} \). Under the notation of proposition 7, the symmetry of \( \theta \) implies \( D_\lambda = D_\rho \), \( D_\lambda^o = D_\rho^o \) and as \( \beta \to \infty \), \( D_\lambda < D_\rho^o \) so
\[
\Delta^e = D_\lambda = \frac{\delta}{1 + \frac{\beta}{N} \left(2 \left\lceil \frac{N+1}{4} \right\rceil + N\right)} > \frac{\delta}{1 + 2\beta}
\]
since for \( N \geq 5 \), \( \left\lceil \frac{N+1}{4} \right\rceil \leq \frac{N+4}{4} < \frac{N}{2} \).

If \( \theta \) is maximally polarized, any LFE is socially better and majority preferred to decentralization from proposition 8. Let \( \theta^k \) be a sequence of profile of type such that \( \pi(\theta^k) \to 0 \) and \((\Gamma^k, \Delta^k)\). Without loss of generality, we renormalize the sequence so that for all \( k \), \( \theta^k_{\mu} = 0 \) and \( \text{med}_n (|\theta^k_n - \theta^k_\mu|) = 1 \). By
definition of \( \pi \), necessarily, \( var(\theta^k) \to \infty \). From what precedes, \( \Gamma^k = 0 \) and from (39), \( \Delta^k \) is bounded. This implies that \( x(\theta^k, \Gamma^k, \Delta^k) \) is also bounded so \( W (x(\theta^k, \Gamma^k, \Delta^k)) = N \text{var}(\theta^k) + O(1) \). From (2), simple algebra shows that \( W (x_{dec}(\theta^k)) = \frac{2\beta + \beta^2}{1 + 2\beta + \beta^2} N \text{var}(\theta^k) \), which shows that decentralization is socially better for \( k \) sufficiently large. Under our assumptions, a majority of states have a bounded type, so since \( x(\theta^k, \Gamma^k, \Delta^k) \) is bounded and \( x_{dec}(\theta^k) \) is unbounded, a majority of voters strictly prefer \( (\Gamma^k, \Delta^k) \) to decentralization for \( k \) sufficiently large.

8.4 Proofs in Section 6

8.4.1 Proof of proposition 7

The following lemma characterizes the second stage voting equilibrium \( G(\theta, \Delta) \) for all \( \Delta \).

**Lemma 9** Under the notations of proposition 7:

\[
\Delta > \min \{ D^0_\lambda, D^0_\rho \} \Rightarrow G(\theta, \Delta) = \begin{cases} 
\{ \theta_\mu - \Delta \} & \text{if } D^0_\lambda < D^0_\rho \\
\{ \theta_\mu + \Delta \} & \text{if } D^0_\lambda > D^0_\rho \\
\{ \theta_\mu - \Delta, \theta_\mu + \Delta \} & \text{if } D^0_\lambda = D^0_\rho
\end{cases}
\]

\[
\Delta < \min \{ D^0_\lambda, D^0_\rho \} \Rightarrow G(\theta, \Delta) = \left\{ \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta \right\}.
\]

For \( \Delta \in [0, \min \{ D^0_\lambda, D^0_\rho \}] \), \( U_1(\theta, G(\theta, \Delta), \Delta) \) (resp. \( U_N(\theta, G(\theta, \Delta), \Delta) \)) is single-peaked in \( \Delta \) with a peak at \( D^0_\lambda \) (resp. \( D^0_\rho \)).

**Proof.** From proposition 2, for all \( \Delta \), there are three possible equilibria at the second stage: either \((l, r) = (0, \rho)\) and \( \Gamma = \theta_\mu - \Delta \), or \((l, r) = (\lambda, 0)\) and \( \Gamma = \theta_\mu + \Delta \), or \((l, r) = (\lambda, \rho)\) and \( \Gamma = \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta \).
If \((l, r) = (0, \rho)\) and \(\Gamma = \theta_\mu - \Delta\), the state equilibrium \(x(\theta, \theta_\mu - \Delta, \Delta)\) as characterized in (10) is given by

\[
\begin{align*}
x_1 &= \frac{\theta_1 + \frac{\beta}{N}(\lambda x_1 + (N-\lambda-\rho)x_2 + \rho x_N)}{1+\beta}, \\
x_\mu &= \frac{\theta_\mu + \frac{\beta}{N}(\lambda x_1 + (N-\lambda-\rho)x_2 + \rho x_N)}{1+\beta}, \\
x_N &= \frac{\theta_\mu}{1+\beta}
\end{align*}
\]

where the second equality is derived by solving the system described by the first equality. This solution is possible at \(\Delta\) if and only if the leftist states are indeed unconstrained at \(\Gamma\), i.e.

\[
x_1 \geq \theta_\mu - 2\Delta \iff \Delta \geq D' \equiv \frac{N + \beta \lambda + \beta \rho}{2(\beta + 1)(N + \beta \rho)}(\theta_\mu - \theta_1).
\]

In this case, substituting (40) in (1), simple algebra yields

\[
U_\mu (\theta, \theta_\mu - \Delta, \Delta) = -\frac{\beta \lambda (N + \beta \lambda + \beta \rho)}{N(\beta + 1)^2(N + \beta \rho)} (\theta_\mu - \theta_1)^2.
\]

Likewise, if \((l, r) = (\lambda, 0)\) and \(\Gamma = \theta_\mu + \Delta\), the state equilibrium \(x(\theta, \theta_\mu + \Delta, \Delta)\) as characterized in (10) is given by

\[
\begin{align*}
x_1 &= \frac{\theta_1 + \frac{\beta}{N}(\lambda x_1 + (N-\lambda-\rho)x_2 + \rho x_N)}{1+\beta}, \\
x_\mu &= \frac{\theta_\mu + \frac{\beta}{N}(\lambda x_1 + (N-\lambda-\rho)x_2 + \rho x_N)}{1+\beta}, \\
x_N &= \frac{\theta_\mu}{1+\beta}
\end{align*}
\]

In this case, substituting (42) in (1), simple algebra yields

\[
U_\mu (\theta, \theta_\mu + \Delta, \Delta) = -\frac{\beta \rho (N + \beta \lambda + \beta \rho)}{N(\beta + 1)^2(N + \beta \lambda)} (\theta_N - \theta_\mu)^2.
\]

From proposition 2, for any \(\Delta\), \(\mathcal{G}(\theta, \Delta)\) are the most preferred \(\Gamma\) of the voters of the median state. From (41) and (43), they strictly prefer \(\Gamma = \theta_\mu - \Delta\) to \(\Gamma = \theta_\mu + \Delta\) if and only if

\[
\lambda (N + \beta \lambda) (\theta_\mu - \theta_1)^2 < \rho (N + \beta \rho) (\theta_N - \theta_\mu)^2 \iff D'_\lambda < D'_\rho.
\]
If \((l, r) = (\lambda, \rho)\) and \(\Gamma = \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta\), then

\[
x \left( \theta, \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta \right) = \frac{\theta_\mu - \frac{2\rho}{\lambda + \rho} \Delta}{\theta_\mu + \frac{2\lambda}{\lambda + \rho}},
\]

This solution is possible at \(\Delta\) if and only if states \(1\) and \(N\) are indeed constrained at \((\Gamma, \Delta) = \left( \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta \right)\), i.e.

\[
\frac{\theta_1 + \frac{\lambda}{N} (\lambda x_1 + (N - \lambda - \rho) x_\mu + \rho x_N)}{1 + \beta} \leq \Gamma - \Delta
\]

and

\[
\frac{\theta_N + \frac{\beta}{N} (\lambda x_1 + (N - \lambda - \rho) x_\mu + \rho x_N)}{1 + \beta} \geq \Gamma + \Delta,
\]

which is equivalent to

\[
\Delta \leq D'' \equiv \frac{\lambda + \rho}{2(1 + \beta)} \min \left\{ \frac{\theta_\mu - \theta_1}{\rho}, \frac{\theta_N - \theta_\mu}{\lambda} \right\}.
\]

In this case, substituting (45) in (1), we get

\[
U_\mu \left( \theta, \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta \right) = -4 \beta \frac{\lambda \rho}{N (\lambda + \rho)} \Delta^2,
\]

\[
U_1 \left( \theta, \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta \right) = - \left( \theta_1 - \theta_\mu + \frac{2\rho}{\lambda + \rho} \Delta \right)^2 - \frac{\beta}{N} \left( (N - \lambda - \rho) \left( \frac{2\rho}{\lambda + \rho} \right)^2 + 4\rho \right) \Delta^2,
\]

\[
U_N \left( \theta, \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta \right) = - \left( \theta_N - \theta_\mu - \frac{2\lambda}{\lambda + \rho} \Delta \right)^2 - \frac{\beta}{N} \left( (N - \lambda - \rho) \left( \frac{2\rho}{\lambda + \rho} \right)^2 + 4\lambda \right) \Delta^2.
\]

From (41) and (46), the voters of the median state prefer \(\Gamma = \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta\) to \(\Gamma = \theta_\mu - \Delta\) if and only if

\[
\Delta \leq \sqrt{\frac{(\lambda + \rho)(N + \beta\lambda + \beta\rho)}{4\rho(\beta + 1)^2(N + \beta\rho)} (\theta_\mu - \theta_1)} = D''_{\lambda}.
\]
A symmetric reasoning shows that the voters of the median state prefer \( \Gamma = \theta_\mu + \frac{\Delta - \rho}{\lambda + \rho} \Delta \) to \( \Gamma = \theta_\mu - \Delta \) if and only if \( \Delta \leq D^0_\rho \). Together with (44), this proves that if \( D^0_\lambda \leq D^0_\rho \) (resp. \( \geq \)), for \( \Delta > D^0_\lambda \) (resp. \( \Delta < D^0_\lambda \)), \( \mathcal{G}(\theta, \Delta) = \{\theta_\mu - \Delta\} \) (resp. \( \{\theta_\mu + \Delta\} \)) and for \( \Delta < D^0_\rho \) (resp. \( \Delta < D^0_\rho \)), \( \mathcal{G}(\theta, \Delta) = \{\theta_\mu + \frac{\Delta - \rho}{\lambda + \rho} \Delta\} \). Simple algebra shows that the feasibility constraints, i.e. \( D'' \geq \min \{D^0_\lambda, D^0_\rho\} \geq D' \), are always satisfied,23 which completes the proof of the first part of the lemma.

If \( \Delta < \min \{D^0_\lambda, D^0_\rho\} \), from (46) \( \frac{\partial U_1(\theta, \theta_\mu + \frac{\Delta - \rho}{\lambda + \rho} \Delta)}{\partial \Delta} = 0 \) if and only if
\[
-\frac{4\rho}{\lambda + \rho} \left( \theta_1 - \theta_\mu + \frac{2\rho}{\lambda + \rho} \Delta \right) - \frac{2\beta}{N} \left( N - \lambda - \rho \right) \left( \frac{2\rho}{\lambda + \rho} \right)^2 + 4\rho \Delta = 0.
\]
Solving for \( \Delta \) yields \( \Delta = D_\lambda \). Likewise, \( \frac{\partial U_N(\theta, \theta_\mu + \frac{\Delta - \rho}{\lambda + \rho} \Delta)}{\partial \Delta} = 0 \) if and only if \( \Delta = D_\rho \). ■

For the sake of brevity, we assume throughout the proof that \( D^0_\lambda > D^0_\rho \). The cases \( D^0_\lambda < D^0_\rho \) and \( D^0_\lambda = D^0_\rho \) can be derived identically. On \( ]D^0_\rho, +\infty[ \), from lemma 9, the state equilibrium is constant (c.f. (40)) and equal to \( x(\theta, \theta_\mu + D^0_\rho, D^0_\rho) \). Since \( \lim_{\Delta \to D^0_\rho} \mathcal{G}(\theta, \Delta) \neq \theta_\mu + D^0_\rho \), from lemma

23Suppose to fix ideas that \( D^0_\lambda \leq D^0_\rho \). If \( |\theta_N - \theta_1| \leq |\theta_\mu - \theta_1| \), \( D'' \geq D^0_\lambda \geq D' \) can be rewritten
\[
\sqrt{\frac{(\lambda + \rho)(N + \beta \rho)}{(N + \beta (\lambda + \rho) \rho}} \geq 1 \geq \sqrt{\frac{(N + \beta (\lambda + \rho) \rho}{(\lambda + \rho)(N + \beta \rho)}},
\]
which is satisfied since \( x \to \frac{N + \beta x}{x} \) is decreasing and \( \lambda + \rho > \rho \). If \( \frac{|\theta_N - \theta_1|}{\theta_\mu - \theta_1|} > |\theta_N - \theta_1| \), \( D'' \geq D^0_\lambda \geq D' \) holds whenever
\[
\frac{|\theta_N - \theta_\mu|}{|\theta_\mu - \theta_1|} \geq \sqrt{\frac{(N + \beta (\lambda + \rho))(N + \beta \lambda)(N + \beta \rho)}{\lambda \rho (N + \beta \rho)}}
\]
which is satisfied since \( D^0_\lambda \leq D^0_\rho \) implies \( \frac{|\theta_N - \theta_\mu|}{|\theta_\mu - \theta_1|} \geq \sqrt{\frac{(N + \beta \lambda)}{\rho (N + \beta \rho)}} \), and from what precedes,
\[
\sqrt{\frac{(N + \beta (\lambda + \rho))(N + \beta \lambda)}{(\lambda + \rho)(N + \beta \rho)}} < 1.
\]
6, \((\Gamma, \Delta) = (\theta_\mu + D_\rho^\sigma, D_\rho^\sigma)\) is not a LFE. From lemma 6, the only LFE candidates are therefore \(\left(\theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta\right)\) for \(\Delta \leq D_\rho^o\). On \([0, D_\rho^o]\), the preferences of all voters on \(x \left(\theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta\right)\) are quadratic and concave in \(\Delta\). Therefore, majority preferences are quasi-concave on \([0, D_\rho^o]\). Moreover, they are either single-peaked or strictly increasing on \(\Delta \leq D_\rho^o\). If they are increasing, from lemma 6 \((\Gamma^e, \Delta^e) \equiv \left(\theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} D_\rho^o, D_\rho^o\right)\) is the only LFE. By assumption, \((\Gamma^e, \Delta^e)\) is preferred by simple majority rule to any \((\Gamma, \Delta) = \left(\theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta\right)\) for \(\Delta < D_\rho^o\). From proposition 2, the voters of the median state are indifferent between \((\Gamma^e, \Delta^e)\) and \((\theta_\mu + D_\rho^o, D_\rho^o)\), from lemma 2 the rightist and leftist voters have opposite preferences between these two alternatives, so \((\theta_\mu + D_\rho^o, D_\rho^o)\) is not preferred by simple majority rule to \((\Gamma^e, \Delta^e)\). This shows that \((\Gamma^e, \Delta^e)\) is a Condorcet winner among all \((\Gamma, \Delta)\) such that \(\Gamma \in G(\theta, \Delta)\).

If majority preferences are single-peaked on \([0, D_\rho^o]\), from lemma 9, \(\Delta^e = \min \{D_\lambda, D_\rho\} \) and \(\Delta^e \leq D_\rho^o\). From lemma 6 \(\left(\Gamma^e = \theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta^e, \Delta^e\right)\) is the unique LFE. By assumption, it is preferred by simple majority rule to any \((\Gamma, \Delta) = \left(\theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta, \Delta\right)\) for \(\Delta \leq D_\rho^o\). From lemma 4, the voters of the median state strictly prefer \((\Gamma^e, \Delta^e)\) to \((\theta_\mu + D_\rho^o, D_\rho^o)\). Moreover, since \(\Delta^e \leq D_\rho^o\),

\[
\theta_\mu + \frac{\lambda - \rho}{\lambda + \rho} \Delta^e \pm \Delta^e < \theta_\mu + D_\rho^o \pm D_\rho^o,
\]

so from lemma 2 \((\Gamma^e, \Delta^e)\) is strictly preferred by simple majority rule to \((D_\rho^o, \theta_\mu + D_\rho^o)\). Hence, \((\Gamma^e, \Delta^e)\) is a Condorcet winner among all \((\Gamma, \Delta)\) such that \(\Gamma \in G(\theta, \Delta)\).

\(^{24}\)They cannot be decreasing to the right of \(\Delta = 0\) from (46).
8.4.2 Proof of Proposition 8

As \( \beta \to \infty \), \( D_\lambda < D_\lambda^c \) and \( D_\rho < D_\rho^c \) so \( \Delta^e = \min \{ D_\lambda, D_\rho \} \). Suppose for the rest of the proposition that \( D_\lambda \leq D_\rho \), i.e. that \( (\rho^2 + \lambda \rho + N \lambda) |\theta_1 - \theta_\mu| \leq (\lambda^2 + \rho \lambda + N \rho) |\theta_N - \theta_\mu| \), the case \( D_\lambda \geq D_\rho \) is identical. From (2), the welfare under decentralization is given by

\[
U^\text{dec}_\mu = - \left( \frac{\beta \lambda (\theta_1 - \theta_\mu) + \rho (\theta_N - \theta_\mu)}{N (1 + \beta)} \right)^2,
\]

\[
U^\text{dec}_1 = - \left( \frac{\beta \rho (\theta_N - \theta_1) + (N - \lambda - \rho) (\theta_\mu - \theta_1)}{N (1 + \beta)} \right)^2,
\]

\[
U^\text{dec}_N = - \left( \frac{\beta \lambda (\theta_1 - \theta_N) + (N - \lambda - \rho) (\theta_\mu - \theta_N)}{N (1 + \beta)} \right)^2,
\]

As \( \beta \to \infty \),

\[
U^\text{dec}_\mu \to - \left( \frac{\lambda (\theta_1 - \theta_\mu) + \rho (\theta_N - \theta_\mu)}{N} \right)^2,
\]

\[
U^\text{dec}_1 \to - \left( \frac{\rho (\theta_N - \theta_\mu) + (N - \lambda) (\theta_\mu - \theta_1)}{N} \right)^2,
\]

\[
U^\text{dec}_N \to - \left( \frac{\lambda (\theta_1 - \theta_\mu) + (N - \rho) (\theta_\mu - \theta_N)}{N} \right)^2.
\]
Since $\Delta^c = D_\lambda$, from (46), the welfare at the federal equilibrium is given by

\begin{align*}
V_\mu &= -\frac{\beta}{N} \lambda \rho \left( \frac{\lambda + \rho}{\rho + \frac{\beta}{N} (N \rho + \lambda^2 + \lambda \rho)} \right)^2 (\theta_1 - \theta_\mu)^2 \\
V_1 &= -\frac{\beta}{N} \left( N \rho + \frac{\lambda^2 + \rho \lambda}{N \rho + \lambda^2 + \lambda \rho} \right) (\theta_1 - \theta_\mu)^2 \\
V_N &= -\left( \theta_N - \theta_\mu - \frac{\lambda (\theta_\mu - \theta_1)}{\rho + \frac{\beta}{N} (N \rho + \lambda^2 + \lambda \rho)} \right)^2 \\
& \quad -\frac{\beta}{N} \left( \frac{\lambda^3 + 2\lambda^2 \rho - \rho^3 + N \rho^2}{\rho + \frac{\beta}{N} (N \rho + \lambda^2 + \lambda \rho)} \right)^2 (\theta_\mu - \theta_1)^2.
\end{align*}

As $\beta \to \infty$,

\begin{align*}
V_1 &\to -(\theta_1 - \theta_\mu)^2, \ V_N \to -(\theta_N - \theta_\mu)^2 \text{ and } V_\mu \to 0. \quad (50)
\end{align*}

Let $W^{dec}$ the welfare under decentralization and $W$ the welfare at the federal equilibrium. Combining (48) and (50), we get,

\begin{align*}
\lim_{\beta \to \infty} (W^{dec} - W) &= \frac{1}{N^2} \left( \lambda |\theta_1 - \theta_\mu| - \rho |\theta_N - \theta_\mu| \right) ((2N - 3\lambda) |\theta_1 - \theta_\mu| + (-2N + 3\rho) |\theta_N - \theta_\mu|).
\end{align*}

If $|\theta_1 - \theta_\mu| = |\theta_N - \theta_\mu|$ and $\lambda \neq \rho$, then from (51),

\begin{align*}
\lim_{\beta \to \infty} (W^{dec} - W) &= -\frac{3}{N^2} (\lambda - \rho)^2 |\theta_1 - \theta_\mu|^2 < 0,
\end{align*}

which proves point (ii). If $\lambda = \rho$ and $|\theta_1 - \theta_\mu| = |\theta_N - \theta_\mu|$, then from (51),

\begin{align*}
\lim_{\beta \to \infty} (W^{dec} - W) &= \frac{\lambda}{N^2} (|\theta_1 - \theta_\mu| - |\theta_1 - \theta_\mu|)^2 (2N - 3\lambda) > 0,
\end{align*}

which proves point (iii).
The voters of the median state are always strictly better-off at the federal equilibrium than under decentralization.\textsuperscript{25} If $\lambda |\theta_1 - \theta_\mu| > \rho |\theta_N - \theta_\mu|$, then (48) and (50) imply $U_1^{\text{dec}} > V_1$ and $U_N^{\text{dec}} < V_N$. Likewise, if $\lambda |\theta_1 - \theta_\mu| < \rho |\theta_N - \theta_\mu|$, then $U_1^{\text{dec}} < V_1$ and $U_N^{\text{dec}} > V_N$, which proves the necessary part of (i). If $|\theta_1 - \theta_\mu| = \rho |\theta_N - \theta_\mu|$, the initial assumption $D_\lambda \leq D_\rho$ implies $|\theta_1 - \theta_\mu| \leq |\theta_N - \theta_\mu|$ and $\lambda \geq \rho$ and thus

$$\rho (\theta_1 - \theta_N)^2 = \rho |\theta_1 - \theta_\mu|^2 + 2\lambda |\theta_1 - \theta_\mu|^2 + \lambda |\theta_1 - \theta_\mu| |\theta_N - \theta_\mu|$$

$$\geq (\rho + 3\lambda) |\theta_1 - \theta_\mu|^2.$$ 

Substituting the above inequality in (43), we get

$$U_1^{\text{dec}} \leq -\left(\left(\frac{\beta}{1+\beta}\right)^2 + \frac{N + 2\lambda}{N} \frac{\beta}{(1+\beta)^2}\right) (\theta_\mu - \theta_1)^2,$$

Together with (47), this implies

$$V_1 - U_1^{\text{dec}} \geq \frac{\beta \lambda N \rho - N \lambda + \beta (N \rho + \lambda (2 \lambda + 2 \rho - N))}{N (\beta + 1)^2} (\theta_1 - \theta_\mu)^2$$

which is positive as $\beta \to \infty$ since $\lambda + \rho > N/2$. This proves (iv). If $|\theta_1 - \theta_\mu| = |\theta_N - \theta_\mu|$ and $\lambda = \rho$, the voters from both rightists and leftists states are pivotal and the same reasoning as above shows that $V_N \geq U_N^{\text{dec}}$ which proves the sufficiency part of (i).

\textsuperscript{25}From lemma 4, given that $\Gamma \in \mathcal{G}(\theta, \Delta)$, the median state prefers $\Delta^e$ to any $\Delta > \Delta^e$. From 2, holding $\Delta$ constant, the median state prefers $\Gamma \in \mathcal{G}(\theta, \Delta)$ to any other $\Gamma$. Since $x^{\text{dec}}(\theta) = x(\theta, \theta_\mu, |\theta_N - \theta_1|)$, the median state prefer the federal equilibrium to decentralization.

49
References


