Coordination in Federal Systems[1]

Antoine LOEPER[2]

April 2008

[1]A previous version of the paper circulated under the title "Heterogeneity and Coordination Externalities in Federal Systems". I wish to thank David Austen-Smith, Helmuth Cremer, Jacques Crémer, Daniel Diermeier, Bard Harstad, Michel Lebreton, seminar participants at Université de Sciences Sociales de Toulouse, CORE summer school 2005 in Louvain La Neuve, EEA congress 2005 in Amsterdam, AFSE congress 2005 in Paris and European Meeting of the Econometric Society 2005 in Istanbul for helpful comments. All remaining errors are mine.

[2]Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, a-loeper@kellogg.northwestern.edu

Abstract

We analyze the welfare properties of centralization versus decentralization in a heterogeneous confederation which exhibits interjurisdictional externalities. We take a different policy focus than the existing literature on fiscal and environmental federalism and consider the coordination problem that arises when jurisdiction members have to coordinate their laws, administrative rules or foreign policies. More concretely, we consider a large class of situations in which external costs are driven by the differences between local policies. Notwithstanding Oates' decentralization theorem, we show that the cost of decentralization vanishes as coordination becomes arbitrarily important. When coordination costs are sufficiently symmetric, decentralization is socially preferred to centralization whatever the magnitude of externalities and the heterogeneity of preferences.

JEL Classification Numbers: D61, D62, H73, H77

Keywords: Federalism, Decentralization, Coordination, Externalities.

1 Introduction

When discussing the pros and cons of federalism in the early times of American democracy, Tocqueville (1835) pointed out that the main failure of its high degree of decentralization was “a lack of uniformity” and “an appearance of disorder”. The lack of coherence in the management of local affairs was argued to result in chaotic local administration and inefficient federal interventions. Conversely, he didn’t consider complete centralization as fully satisfactory for a unique administration could not embrace the diversity of wishes and needs of its constituents. Nevertheless Tocqueville concluded that in the administrative sphere, the advantages of decentralization outweigh the harmony of centralization.

About two centuries later, the appropriate vertical allocation of political responsibilities in federal systems is still the subject of intense debate. The political economy literature has shifted its policy focus from administrative and political considerations to fiscal ones but the basic trade-off has the same flavor. In more modern terms, inter jurisdictional externalities cannot be properly internalized under decentralization while centralized policies are less responsive to local needs. Consequently, most scholars compare the costs and benefits of political integration by balancing the heterogeneity of local preferences versus the magnitude of synergies and cross-border externalities. This fundamental trade-off has first been highlighted by Oates (1972) and the subsequent literature on fiscal and environmental federalism. This literature analyzes the impact of *tax and expenditure externalities* generated by redistributive programs, tax competition on mobile factors or local public goods with spillovers. It has highlighted how fiscal competition and environmental spillovers generate a *prisoner's dilemma* which results in the well known *race to the bottom* when policies are decentralized.

So far, no attempt has been made to extend this formal analysis to the administrative, regulatory or legislative policy domains studied by Tocqueville. As noted by Ellingsen (1998), "the study of laws, regulations and standards may require a different model." This paper makes a first step in this direction and analyzes the pros and cons of decentralization when it comes to enacting laws and administrative rules rather than deciding the level of taxes and public good provision. Arguably, public policies of a different nature involve different externalities and thus distinct coordination failures.

We adopt a modeling strategy similar to the previously mentioned literature on federalism and take the presence of cross border externalities along with heterogeneous preferences as the primitives of our analysis. We depart from it in considering *coordination externalities* which embody the cost borne by the members of a federal systems for having inconsistent policies. Contrary to tax and expenditure externalities, they increase with the heterogeneity and incoherence of local policies. These costs take the form of transaction or compliance costs, conflict of laws, non tariff barriers to trade or fixed cost in policy setting and implementation. It has been estimated that removing the costs associated with heterogeneous regulations could boost bilateral trade of services within the E.U. by 30 to 60%.[1] A study from the European Commission has shown that small and medium transnational European firms have to pay 2.6% of their sales just to adapt to the different national accounting

[1] Kox, Lejour and Montizaan 2004.

norms in the union.[2] More generally, conflicting laws, incompatible standards and fragmented administrations deter economic competition, exchange and trust.[3] Coordination externalities are not confined to the business sphere: misaligned diplomatic efforts can jeopardize negotiations with other countries. Heterogeneous school curricula, time zones, administrative rules and official languages impede exchange, cooperation programs and federal interventions.

Disjoint preferences along with coordination externalities give rise to a conflict of interest reminiscent of the *battle-of-the-sexes*:[4] although local needs or intrinsic preferences differ from one region to another, each jurisdiction prefers the other policies to be as close as possible to its own. In our model, consistently with Tocqueville's observation and with the view of many federal administrations, the coordination failure of decentralization is an excessive policy heterogeneity: the fiscal race to the bottom becomes an administrative and political race to local particularism.

To reduce the costs generated by inconsistent laws and policies, a natural alternative is the uniformization of policies. As a member of the European Parliament puts it, "having a single set of rules for a single market, rather than 25 different national rules, is actually an exercise in cutting red tape and bureaucracy."[5] This view is pervasive in the recent proposal of the U.S. Treasury Secretary to streamline financial regulation: "regulated entities can

[2]Source: Le Figaro économie 13/09/2004

[3]See e.g. Anderson and van Wincoop 2004 and Guiso, Sapienza and Zingales 2006.

[4]See for instance Fudenberg and Tirole (1991, p. 18) for a presentation of the battle-of-the-sexes game.

[5]Richard Corbett, source: EU Reporter, 17-25/01/2005.

benefit from uniform oversight by not being subject to the potential inefficiency of having to deal with multiple state regulators, laws, regulations and standards."[6] The recent calls for a E.U. foreign minister to conduct a common foreign policy resort to the same line of arguments. A uniform policy obviously solve coordination problems. However, such benefits must be balanced against the impossibility to take into account the diversity of preferences across the confederation.

The main goal of the paper is to investigate under which conditions imposing a uniform rule can be socially beneficial. The welfare analysis highlights a trade-off between the internalization of externalities and the matching of local preferences which is reminiscent of the fiscal federalism setup. However, the comparative statics of the decentralization debate turn out to be quite different if not opposite. First, if external costs are sufficiently salient, i.e. elastic, then no uniform policy Pareto dominates decentralization: salient enough externalities have a self-disciplining effect which induces local jurisdictions to choose their policies close to each other. Second, if external effects are furthermore symmetric, decentralization is socially preferred to any uniform policy. Surprisingly, both results hold whatever the magnitude of externalities and the heterogeneity of preferences. The reason is that if externalities are reciprocal, local jurisdictions internalize half of the social cost they impose which limits the coordination failure under decentralization. Third, the cost of decentralization is not monotonic in the magnitude of externalities and vanishes as they become arbitrarily large. Fourth, when coordination externalities are not continuous and take the form of political

[6] U.S. Treasury's Blueprint for a Modernized Financial Regulatory Structure.

network effects, decentralization is Pareto efficient whenever external effects are not too asymmetric.

These results, derived under conservative institutional assumptions, are in sharp contrast with the previously mentioned literature on fiscal and environmental federalism and with Oates decentralization theorem in particular. They suggests that when the fiscal prisoner's dilemma turns into a coordination problem, the comparative statics of the decentralization debate are quantitatively more favorable to decentralization and qualitatively different: the symmetry of external effects plays a more important role than their magnitude. For instance, one should not think in the same terms about the vertical allocation of policy making authority for education expenditure versus school curriculum, or law enforcement expenditure versus penal legislation.

The paper is organized as follows. Section 3 lays out the basic model. Section 4 derives decentralization theorems for convex coordination costs and section 5 allows for non convexities and network effects. Section 6 develops two extensions: local elections with intrajurisdiction heterogeneity and mutual recognition of regulatory regimes. Section 7 concludes. All proofs are relegated to an appendix.

2 Related Literature

So far, economists have mostly dealt with *tax or expenditure externalities* as in the seminal contributions of Oates 1972, Gordon 1983 and Zodrow and

Mieszkowski 1986.[7] For the sake of comparison, a standard specification of payoffs in the local public good with positive spillovers scenario is:[8]

$$U_i(x) = \theta_i \times G\left(x_i + \beta \sum_{j \neq i} x_j\right) - x_i, \tag{1}$$

where x_d is the provision of public good in jurisdiction i, G is increasing and concave, $\beta \in [0, 1]$ is the magnitude of spillovers and $\theta \in \mathbb{R}_+^N$ allows for heterogeneity of willingness to pay for the public good across the confederation. The central normative result is Oates Decentralization Theorem (henceforth O.D.T.): whenever spillovers (β) are small and local preferences (θ) sufficiently heterogeneous, decentralization outperforms centralization. Conversely, if local preferences are sufficiently homogeneous and spillovers large enough, centralization – i.e. voting over a uniform policy – is Pareto better. Both parts of the theorem do not necessarily hold in our context. Moreover, our results stress the importance of allowing for a richer topology of external effects rather than a uniform β.

In a similar setup, Besley and Coate 2003 and Lockwood 2002 have taken a more political economy approach and found that the comparative advantage of centralization may not be monotonic in the size of spillovers. Their results are driven by strategic voting at local and federal elections while ours are driven solely by the structure of interjurisdictional externalities. Moreover, they do not look at the impact of the spatial distribution of externalities.

Our model shares some similarities with the literature on coordination in

[7] See the reviews of Revesz 1996, Farber 1997, Oates 1999 or Wilson 1999. See also, among others, Persson and Tabellini 1992 or Ellingsen 1998 for more recent contributions.

[8] Of course, the specification for negative spillovers or tax externalities would be different but comparative statics are similar.

multi-divisional firms. Most prominently, Alonso, Dessein and Matouschek (2007) have found as well that the cost of decentralization is not monotonic in the importance of coordination while in Rantakari (2006), asymmetries in coordination costs calls for an asymmetric form of centralization. Their model features incomplete information, cheap-talk communication and their institutional assumptions are not directly comparable to ours. Nevertheless, their payoff specification is a very special case of ours and our results suggest that their conclusions might be driven to some extent by the particular profit function they assume and not only by strategic communication between managers.

3 The Basic Model

We consider a federal system composed of I jurisdictions which can be variously interpreted as states, districts, business units or agents in a social network. Each jurisdiction has the same number of identical residents.[9] An alternative x is a vector of unidimensional local policies,[10] x_i being the policy of jurisdiction i and x_{-i} the policies of the other jurisdictions.

3.1 Citizens' Preferences

As argued in the introduction, we consider policies which generate a conflict of interest similar to the battle of the sexes: each jurisdiction would prefer

[9] Subsection 6.1 allows for local heterogeneity and different population sizes.

[10] Our results are extended to multidimensional policy spaces in subsection 4.3 and to finite policy spaces in section 10.

the other policies to be as close as possible to its own. Formally, if $U_i(x)$ is the utility of the citizens of jurisdiction i, given x_i, $U_i(x)$ is single peaked in x_{-i} with a peak at $x_{-i} = (x_i, .., x_i)$. If we denote $V_i(x_i) = U_i(x_i, .., x_i)$ and $W_i(x) = U_i(x_i, .., x_i) - U_i(x)$, we can decompose U_i as follows:

$$U_i(x) = V_i(x_i) - W_i(x). \tag{2}$$

The first term $V_i(x_i)$ can be viewed as the *internal effect* of the local policy of jurisdiction i: it measures whether the local policy meets the specific needs of jurisdiction i and the preferences of its residents absent any coordination costs with the rest of the confederation. The second term $W_i \geq 0$ is the *coordination external effect*: it embodies the cost born by jurisdiction i for having a policy which is different from the ones of the other members of the confederation.

For the clarity of exposition, we will first present the main insights of our model using the following parametric special case of (2):

$$U_i(x) = -|x_i - \theta_i|^2 - \sum_{j \neq i} \beta_{i,j} |x_i - x_j|^\alpha. \tag{3}$$

This more simple, although quite rich and flexible formulation will allow us to derive clean comparative statics and single out the effect each parameter. The vector θ, which we call the profile of type, allows for *idiosyncratic* heterogeneity. The matrix of interjurisdictional interactions $\beta = (\beta_{i,j})_{j \neq i}$ allows for *relational* heterogeneity. Most models of federalism introduce only for the first kind of heterogeneity.[11] A key insight of our model is that the

[11] Lockwood 2002 is an exception, but the paper deals with local public goods with spillovers and doesn't provide comparative statics w.r.t. the relational heterogeneity.

structure of the relational heterogeneity is actually a key determinant of the comparative statics.

For all $i \neq j$, the coefficient $\beta_{i,j}$ stands for the magnitude of the external cost imposed to jurisdiction i by the policy implemented in jurisdiction j, while the magnitude of the internal effect is normalized to 1 without loss of generality. We refer to β as the externality matrix and by convention, its diagonal is null. The parameter α represents the salience of the external effect compared to the salience of the internal effect, normalized to 2.[12] Put differently, $\alpha - 1$ is the elasticity of the coordination cost.

Section 4 deals with the case of convex coordination externalities, i.e. $\alpha > 1$. It corresponds to coordination costs which increase smoothly and gradually in the political or administrative distance between local policies. This could be the case for instance of time zones, traffic codes, cultural or foreign policies. Section 5 look at the parameter region $0 < \alpha \leq 1$ where incompatibility costs increase rapidly when local policies deviate slightly from full standardization, and to the limit case $\alpha \to 0$ which corresponds to pure network effects and economies of scale. Such effects can be generated for instance by technological standards and regulations or economies of scale in policy setting and implementation. In both sections, we first derive our result under the parametric specification in (3) and then generalize them to a broader class of preferences.

[12] We will show in subsection 4.3 that this normalization is innocuous.

3.2 Centralization and Decentralization

In order to draw *ceteris paribus* conclusions, we follow the institutional assumptions of the "first generation" literature on fiscal and environmental federalism. Under decentralization, each jurisdiction chooses its local policies to maximize the welfare of its constituency taking the other policies as given. In other words, decentralization is modelled as a non cooperative game between jurisdictions.[13]

Under centralization, a uniform policy is implemented. By definition, it removes all coordination costs. Our results characterize sets of parameters under which there is no uniform policy which is socially preferred to decentralization. For this purpose, we do not need to specify which or how the uniform policy is chosen.

3.3 The Second Best Approach

The basic model follows closely Oates' institutional assumptions (1972): centralization generates preferences frustration through a one-size-fits-all policy while decentralization entails a coordination failure since externalities are not internalized. This second best scenario has been the workhorse for most of literature on federalism.[14] The shortcomings of this modeling strategy have been discussed in numerous papers.[15] In this paper, we stick to it in order

[13]In subsection 6.1, we allow for intradistrict heterogeneity and model local policy making through local elections.

[14]Some recent papers are Cremer and Palfrey 1996, Alesina and Spolaore 1997, Bolton and Roland 1997 and Ellingsen 1998.

[15]See for instance Cremer, Estache and Seabright 1996, Epple and Nechyba 2004 or Oates 2005 for a dicsussion.

to see whether Oates' conclusions are still valid under the same institutional assumptions but with a different policy focus. Taking a step back and endogenizing the respective cost of centralization and decentralization for this new collective choice problem is an interesting direction which do not pursue here for the sake of brevity.[16]

In any case, the uniformity assumption is arguably more plausible in our legislative and regulatory context than in the public finance setup. First of all, contrary to taxation and public good provision, uniformization is beneficial *per se* in our case for it removes all external effects. Second, empirically, many laws and regulations are in fact standardized by central authorities, either indirectly through directives as in the E.U. – e.g. the Common Technical Regulation in telecommunication – or directly by centralized legislation – e.g. the common time zone in China. See also Strumpf and Oberholzer-Gee (2002) for an example in liquor control in the U.S. Third, theoretical arguments point toward uniformity as well. A public choice school model of centralization would predict a uniform policy since the most preferred policy of any citizen is indeed uniform.[17] From a more contractual perspective, in a companion paper (Loeper 2008), we show in a similar setup that if the information about local preferences is privately held by local jurisdictions, the incentive compatibility problem may constrain the federal planner to implement a uniform mechanism.

[16] Seabright 1996, Qian and Weingast 1997, Lockwood 2002, Besley and Coate 2003, Harstad 2007 and Loeper 2007 propose alternatives models of a more political economy flavor.

[17] This argument can be made more formal, for instance by modelling federal elections with a citizen-candidate model *a la* Besley and Coate (1996).

4 Convex Coordination Costs

Subsection 4.1 and 4.2 derive comparative statics on the relative benefits of centralization and decentralization under the parametric specification in (3) in order to single out the effect of each of the parameters (α, β, θ). As we focus on convex coordination costs, we assume throughout those subsections that $\alpha > 1$ and this assumption will be implicit in all corresponding propositions. We generalize our results on convex coordination costs to a larger class of preferences in subsection 4.3.

Proposition 1 *Under the specification in (3), for all $\alpha > 1$ and all (β, θ), there exists a unique decentralized equilibrium which we denote $x^{dec}(\alpha, \beta, \theta)$.*

As the game between local jurisdictions is supermodular with a unique Nash equilibrium, it satisfies some appealing properties that make it a robust non cooperative prediction under mild rationality and informational assumptions.[18]

4.1 Pareto Improving Centralization and the Salience of Externalities

Proposition 2 *The decentralized equilibrium is Pareto optimal if and only if the profile of type θ is uniform. A uniform policy is Pareto optimal if and only if it is the most preferred policy of some jurisdiction.*

[18] See Milgrom and Roberts 1990 and Vives 1990.

The coordination failure of decentralization takes the form of an excessive heterogeneity of local policies.[19] As externalities are not internalized, each jurisdiction chooses its policy too close to its type, at the expenses of coordination.

Contrarily, a uniform policy removes external costs and can be Pareto optimal, but the gains from uniformization may not be large enough to reconcile the interests of jurisdictions with extreme types. We end up with a trade-off between the match of local needs and the internalization of externalities which is reminiscent of the fiscal federalism scenario. A natural question is then whether a uniform policy can improve unanimously on the decentralized equilibrium.

Proposition 3 *Given a profile of externalities* (α, β)*, there exists a profile of type* θ *and a uniform policy which strictly Pareto dominates the decentralized equilibrium if and only if* $\alpha < 2$.

Proposition 3 is at odd with the common wisdom on the relative benefits of centralization and decentralization in two respects. First if externalities are more than quadratic, then however large the externality matrix β and however homogeneous the profile of types θ, no uniform policy unanimously improves on the decentralized equilibrium. Second, there exists a uniform policy which Pareto improves on the decentralized equilibrium only if the salience of external effects is large enough. It is surprising in the sense that a uniform policy can be viewed as a full insurance against external costs while

[19] For instance, in the case $I = 2$, one can show that for all Pareto optimal policy vectors (x_1^*, x_2^*), we have $|x_1^* - x_2^*| < \left|x_1^{dec} - x_2^{dec}\right|$.

decentralization is a way to limit internal costs. If elasticities are interpreted as degrees of relative risk aversions,[20] proposition 3 suggests that the more risk averse the jurisdictions are to coordination costs, the less they will benefit from being insured against the latter through a uniform rule.

The intuition behind these results is that externalities act as a self-discipline device: salient enough externalities make jurisdictions less inclined to set their local policy far away from each other which mitigate the coordination failure of decentralization.

Notice that proposition 3 does not imply that some degree of harmonization is not worthwhile. Indeed as the magnitude of external effects increases, the decentralized equilibrium tends towards a uniform policy. But one cannot leap from the premise that some policy harmonization is desirable to the conclusion that complete uniformization is the best way to effect it. Modern federal systems provide numerous examples of partial policy harmonization without any coercive federal intervention: many federal regulatory bodies or international organizations such as the U.N. or the I.M.F. promote uniform policy through model laws which can be modified by states if and when adopted.[21] Similarly, international agreements and standards result in some degree of law harmonization but bind states only with their own consent.[22]

When $\alpha < 2$, a Pareto improving uniform policy may exists. Obviously,

[20]If the utility representation in (3) is a Von Neumann-Morgenstern one, the degree of relative risk aversion w.r.t. $|x_i - x_j|$ is $\alpha - 1$.

[21]For instance, in the U.S., the National Conference of Commissioners on Uniform State Laws ("NCCUSL"), a group that aims to promote uniformity of state law, has drafted over 200 model laws since its founding in 1892.

[22]See for instance Farber 1997 or Posner 2006 for such an account of international law.

this can happen only if β is sufficiently large: if externalities are negligible, decentralization is almost optimal. However, as the next proposition shows, the comparative advantage of centralization is not monotonic in the size of externalities.

Proposition 4 *Given (α, θ), let $(\beta^n)_{n\geq 0}$ be a sequence such that for all $i \neq j$, $\beta^n_{i,j} \to +\infty$, and $\beta^n_{i,j} = O\left(\beta^n_{j,i}\right)$,[23] let x^{dec}_n and x^*_n be respectively the decentralized equilibrium and a Pareto improving policy for $(\alpha, \beta^n, \theta)$ and denote U^{dec}_n and U^*_n the corresponding profile of utility, then $U^{dec}_n - U^*_n \to 0$.*

Observe that we do not restrict the Pareto improving policy to be uniform, so proposition 4 shows that as externalities become arbitrarily severe, the decentralized equilibrium gets arbitrarily close to the Pareto frontier. Contrarily to the case of local public goods with positive spillovers as described in (1), the welfare loss due to the coordination failure of decentralization is not always increasing in the size of externalities. The discrepancy is due to the way the gap between the local cost and the federal cost of political decision varies with the magnitude of externalities. An increase in the intensity of spillovers accruing from other jurisdictions decreases the private returns of providing the local public good – the usual free riding problem – while it increases their social returns. On the contrary, an increase in the intensity of coordination externalities increases both the private and the social returns from policy coordination.

When coordination becomes crucial, the rationality of local units alone induces self-discipline and voluntary harmonization, enough so that a central

[23] Let $(u_n)_n$ and $(v_n)_n$ be two positive series, $u_n = O(v_n)$ if $\limsup_n \frac{u_n}{v_n} < +\infty$.

intervention, whatever it be, is of limited help. For instance, the laws on international sales of good of most legal systems of the world have rapidly converged as trade soared after the second world war without coercive intervention by a supranational body.[24] Similarly, most languages have developed without any well-defined single authority. Arguably, a language is a social convention which requires a lot of coordination. As Friedman (1980) puts it, "It is a complex structure that is continually changing and developing. It has a well defined order, yet no central body planned it."

Proposition 4 imposes some symmetry conditions on the externality matrix. In the appendix we show that the latter condition is not necessary in the two jurisdictions case but for more than two jurisdictions, the result may fail if we drop this assumption. As the next subsection illustrates, the degree of reciprocality of externalities turns out to be a key determinant of the cost of decentralization.

Notice that the Pareto principle cannot rank the two regimes when externalities are more than quadratic: from proposition 3 centralization does not Pareto dominate nor does decentralization since there always exists a Pareto optimal uniform policy. The next subsection solves the indeterminacy by taking a welfarist approach.

4.2 Federal Welfare and the Symmetry of Externalities

In order to obtain more discriminative conclusions, social preferences are derived from a Bergson-Samuelson social welfare function $S(U)$ where U is as defined in (3). For the clarity of exposition, we first consider the case

[24] See for instance Berman 1965.

of two jurisdictions, and any social welfare function S which is symmetric, concave and neutral with respect to the alternatives.[25] Hence the following results do not depend on the relative weight assigned to efficiency versus redistributive concerns. We use the following notations:

$$\beta_S = \frac{\beta_{1,2} + \beta_{2,1}}{2} > 0, \ \beta_A = \frac{\beta_{1,2} - \beta_{2,1}}{\beta_{1,2} + \beta_{2,1}} \in \left]-1, 1\right[. \tag{4}$$

The terms β_S and β_A stand respectively for the average magnitude of externalities and their degree of asymmetry. Most of the literature on federalism derives comparative statics with respect to the size of externalities – i.e. β_S – and the degree of *idiosyncratic heterogeneity – i.e.* $|\theta_1 - \theta_2|$ – assuming away the possibility of *relational heterogeneity* – i.e. $\beta_A = 0$. The following propositions deliver two important results. First, even under our standard institutional assumption, the comparative advantage of decentralization may be decreasing in the degree of idiosyncratic heterogeneity and increasing in the magnitude of externalities. Second, the relational heterogeneity of preferences turns out to be a key determinant of the cost of decentralization so that focusing on the degenerate case $\beta_A = 0$ would be misleading.

Proposition 5 *In the case $I = 2$, given $(\alpha, \beta_A, \theta)$ with $\beta_A \neq 0$,*

(i) Decentralization is socially preferred to any uniform policy if and only if

[25] Concavity corresponds to redistributive concerns. Symmetry amounts to assume that districts' welfare are ordinally comparable under their cardinal representation in (3). See subsection 6.1 for non symmetric social welfare functions.

Neutrality w.r.t. the alternatives means that the social welfare ordering is a function of ordinal preferences only – see e.g. Moulin 1989. As proved in the appendix, in our model, it is equivalent to the scale invariance of S. The corresponding class of social welfare functions contains among others the isoelastic ones $S(U) = \sum_i -|U_i|^\gamma$ for any $\gamma \geq 1$ and hence the Benthamite and the Rawlsian ones.

$\beta_S \leq \widetilde{\beta}$ *for some finite positive* $\widetilde{\beta}$.

(ii) The utility of both jurisdictions is monotonically decreasing in β_S *if and only if* $|\beta_A| > \frac{2-\alpha}{\alpha}$. *Otherwise it is increasing for* $\beta_S \geq \widetilde{\beta}$ *for some finite positive* $\widetilde{\beta}$ *and for some jurisdiction.*

Point (i) is consistent with O.D.T.: large enough externalities make the case for centralization, provided that the uniform policy is chosen optimally. However, consistently with proposition 4, point (ii) suggests that whenever they are sufficiently symmetric, larger externalities provide self-disciplining incentives that can mitigate the coordination failure of non cooperative behavior: the latter is not monotonic in the size of externalities.[26]

Proposition 6 *In the case* $I = 2$, *given* $(\alpha, \beta_A, \beta_S)$ *with* $\beta_A \neq 0$, *if* $\alpha < 2$ *(resp.* $\alpha > 2$*) decentralization is socially preferred to any uniform policy if and only if* $|\theta_1 - \theta_2| \geq \Delta$ *(resp.* $|\theta_1 - \theta_2| \leq \Delta$*) for some finite positive* Δ.

Proposition 6 highlights another departure from O.D.T.: while more homogeneous preferences make the case for centralization if external effects are less salient than internal effects, the opposite holds in the other case. The reason is that the heterogeneity of preferences increases as well the cost of decentralization since it increases the polarization of policies in equilibrium. When external effects are more elastic than internal effects, the coordination failure of decentralization increases more rapidly than the cost of a one-size-fits-all policy.

[26] In particular, if the uniform policy is not chosen optimally as in (i), decentralization may be socially preferred if and only if $\beta_S \in [\underline{\beta}, \overline{\beta}]$ for some finite positive $\underline{\beta}, \overline{\beta}$.

Proposition 7 *In the case $I = 2$,*
(i) Given $(\alpha, \beta_S, \theta)$, decentralization is socially preferred to any uniform policy if and only if $|\beta_A| \leq \widetilde{\beta}$ for some $\widetilde{\beta}$.
(ii) Decentralization is socially preferred to any uniform policy for all (β_S, θ) if and only if $\beta_A = 0$ and $\alpha \geq 2$.

More symmetry in external effects makes the case for decentralization. Contrary to O.D.T., in the case of symmetric and sufficiently salient externalities, decentralization is socially preferred whatever the size of externalities and the heterogeneity of preferences. The intuition is that if externalities are symmetric, the social cost imposed by jurisdiction i on jurisdiction j, $\beta_{j,i} |x_j - x_i|^{\alpha}$ is equal to the private cost suffered by jurisdiction i, $\beta_{i,j} |x_i - x_j|^{\alpha}$. The reciprocality of externalities imposes some congruence between the private and the social costs of their decisions which limits the cost of decentralization.
Proposition 7(ii) can be generalized to any number of jurisdictions under the usual Benthamite social welfare function $B(x) = \sum_i U_i(x)$ where U_i's are as defined in (3).

Proposition 8 *If β is symmetric and $\alpha \geq 2$, decentralization is Benthamite preferred to any uniform policy for all θ.*
In the case $\alpha = 2$, decentralization is Benthamite preferred to any uniform policy for all θ if and only if for all i,

$$\sum_{j=1}^{I} \left(\beta_{i,j} - \beta_{j,i}\right) \geq -1.$$

If externalities are symmetric, the decentralized game between jurisdictions admits a potential function:[27]

$$P(x) = -\sum_i |x_i - \theta_i|^2 - \frac{1}{2}\sum_{i,j} \beta_{i,j} |x_i - x_j|^\alpha ,$$

which differs from the Benthamite social welfare function B only through the weight of the second sum – the external effects – which is twice as small as in B. Roughly speaking, the reciprocality of externalities makes jurisdictions internalize at least half of the Benthamite social cost.

From proposition 3, the condition $\alpha \geq 2$ cannot be dropped. The second part of the proposition shows that some degree of symmetry in external effects is necessary as well. Observe that for all β, we have $\sum_i \sum_j \left(\beta_{i,j} - \beta_{j,i}\right) = 0$, so that to satisfy the I conditions of proposition 8, the average asymmetry between externalities imposed and suffered by each jurisdiction $\sum_{j=1}^{I} \left(\beta_{i,j} - \beta_{j,i}\right)$ should not be too large in absolute value compared to the weight of the local effect – equal to 1.[28]

The symmetry of externalities may be a reasonable assumption in many contexts, for instance if the main determinant of coordination costs is geographic proximity. Similarly, transaction costs due to conflict of laws are often borne to a comparable extent by all parties. It is by no means a general rule. Local regulatory policies are often designed to be more stringent on foreign competitors than on national champions. Such discrimination is theoretically ruled out by the dormant Commerce Clause in the U.S. or the article 30 and 36 of the Treaty Establishing the European Community. As the European

[27] See Monderer and Shapley 1996 for a definition of potential games.

[28] In the two jurisdictions case, the conditions of proposition 8 amount to $|\beta_1 - \beta_2| \leq 1$.

Court of Justice puts it, local regulations are legal "so long as they affect in the same manner, in law and in fact, the marketing of domestic products and of those from other Member States."[29] However, even non discriminatory policies may have an asymmetric impact: cross-border flows of goods, capital, services, workers and students are rarely balanced. Consequently, the costs generated by incompatible accounting rules, regulatory policies, labor laws and academic curricula will not be borne evenly on each side of the border.

4.3 A Generalization

Suppose now that for each jurisdiction, the policy space is $\mathbb{R}^n$ for some $n \geq 1$. The space of alternatives is then $(\mathbb{R}^n)^I$. Let U be a differentiable profile of utility function, let V and W be the profile of internal and external effects as defined in (2), let v and w denote the sum of internal and external effects respectively: $w(x) = \sum_i W_i(x)$, $v(x) = \sum_i V_i(x)$.

Definition 1 *A map $w : \mathbb{R}^{nI} \rightarrow \mathbb{R}$ is invariant by uniform translation if for all $x_0 \in \mathbb{R}^n$, $w(x) = w\left((x_i - x_0)_i\right)$.*

The assumption of invariance by translation is natural in our setup since what drives coordination costs is the difference between policies rather than the level of policies as in the public good case. It is a relatively mild assumption in the sense that by construction, W is null at any uniform policy.

In the next definitions, the symbol "." stands for the internal product in $\mathbb{R}^n$

[29] See for instance Farber 1997 on discriminatory regulations.

and $\nabla_{x_i} w$ denotes the vector of partial derivatives of w with respect to the entries of the vector x_i.

Definition 2 *A map $w : \mathbb{R}^{nI} \to \mathbb{R}$ is more than quadratic if for all $x \in \mathbb{R}^n$, $\sum_i x_i.\nabla_{x_i} w(x) \geq 2w(x)$.*

Using Euler's theorem,[30] a non negative map W which is homogeneous of degree α is more than quadratic if and only if $\alpha \geq 2$. However, a map can be more than quadratic without being homogeneous of any order.[31]
Obviously, under our notation, if W_i is more than quadratic – resp. invariant by uniform translation – for all i, so is w.

Definition 3 *A map $W : \mathbb{R}^{nI} \to \mathbb{R}^I$ is reciprocal if for all i and $x \in \mathbb{R}$, $\nabla_{x_i} W_i(x) = \sum_{j \neq i} \nabla_{x_i} W_j(x)$.*

In words, a profile of external costs W is reciprocal if by changing its policy, each jurisdiction hurts itself as much as it hurts the other jurisdictions. It is satisfied for instance if for all i, $W_i(x) = \sum_j W_{i,j}(x_i, x_j)$ and for all $i \neq j$, $W_{i,j}(x_i, x_j) = W_{j,i}(x_j, x_i)$. In particular, it does not impose any restriction on the magnitude of external effects. We can now state a generalization of propositions 1, 3 and 8:

Proposition 9 *Suppose that v is concave, w is invariant by uniform translation, more than quadratic and W is reciprocal then there exists a decentralized*

[30] The Euler's theorem states that for any differentiable function f which is homogeneous of order r, $x.\nabla f(x) = rf(x)$.

[31] For instance, it should be clear that a sum of homogeneous functions of different order all greater than two is more than quadratic.

equilibrium and any of them is Benthamite superior to any uniform policy. In particular, no uniform policy can Pareto dominate decentralization.

The concavity of V_i is a relatively mild assumption which does not impose any restriction on the degree of heterogeneity. Once again, decentralization, although generically suboptimal, dominates centralization whatever the magnitude of externalities and the heterogeneity of local preferences.

5 Non Convexities and Network Effects

In section 4, we have assumed that external costs are smooth enough to guarantee the convexity of preferences. However, if coordination costs decrease only when local policies are sufficiently close, or if policies involve fixed costs, the continuity and convexity assumptions may be violated. Such patterns arise when dealing with technical standards and regulations or as a result of network effects in the production and consumption of public services. Formally, the model could encompass non convexities by allowing α to take values in $]0,1]$ in (3). The limit case $\alpha \rightarrow 0$ corresponds to network effects or economies of scale in policy development and implementation. Consistently with the aforementioned analogy with the battle of the sexes, the decentralized equilibrium may not be unique and could be uniform, which complicates the comparison with unitarian centralization. However, the comparative statics remain qualitatively similar. If anything, non convexities and network effects make the case for decentralization even stronger.

5.1 Non convexities

The next proposition considers a confederation of two jurisdictions and use the notations introduced in (4).

Proposition 10 *In the case* $I = 2$*, for all* $\alpha \in]0,1]$*, for all* (β_A, θ)*, there exists* $\underline{\beta}, \overline{\beta}$ *with* $0 < \underline{\beta} \leq \overline{\beta} < +\infty$ *such that:*[32]

(*i*) *For* $\beta_S < \underline{\beta}$*, the decentralized equilibrium is unique, non uniform and not Pareto optimal. There exists a uniform policy which Pareto dominates the decentralized equilibrium if and only if* β_S *is close enough to* $\underline{\beta}$*.*

(*ii*) *For* $\underline{\beta} \leq \beta_S \leq \overline{\beta}$*, the set of decentralized equilibria is comprised of a single non uniform policy and a set of uniform policies. The latter Pareto dominate the former and are Pareto optimal.*

(*iii*) *For* $\beta_S > \overline{\beta}$*, the set of decentralized equilibria is comprised of uniform policies only. Any Pareto suboptimal equilibrium is Pareto dominated by a Pareto optimal one.*

In case (*ii*) *and* (*iii*)*, if* $\beta_A = 0$ *the optimal Benthamite policy is an equilibrium.*

An immediate corollary is that decentralization entails policy uniformization only when it is efficient: in case (*ii*) and (*iii*), even if there is a suboptimal equilibrium, there is always a uniform one which is optimal and Pareto better. Hence, a Pareto efficient outcome under decentralization is a reasonable

[32] As we show in the appendix, the comparative statics with respect to θ are similar: for all $(\alpha, \beta_A, \beta_S)$, there exists $\underline{\Delta\theta}, \overline{\Delta\theta}$ such that $0 < \underline{\Delta\theta} < \overline{\Delta\theta} < +\infty$, case *i*) prevails for $|\theta_1 - \theta_2| > \overline{\Delta\theta}$, case (*ii*) for $\underline{\Delta\theta} \leq |\theta_1 - \theta_2| \leq \overline{\Delta\theta}$ and case (*iii*) for $|\theta_1 - \theta_2| < \overline{\Delta\theta}$.

conjecture when externalities are sufficiently large.

In any case, consistently with proposition 4 and contrary to O.D.T., when externalities are sufficiently severe (and as we switch from case (i) to (ii) or (iii)), the potential gains from centralization vanish. In particular the ranking of centralization and decentralization is not monotonic in the magnitude of externalities. From the last point of proposition 10, the symmetry of externalities still makes the case for decentralization.

5.2 Pure Network Effects

As α tends to 0, coordination externalities become pure network effects – or positive externalities as termed in Konishi, Lebreton and Weber (1997): the welfare of a given jurisdiction i depends only on its policy x_i and on the set of jurisdictions which have the same policy $\{j : x_j = x_i\}$. In this limit case, the coordination failure of decentralization may actually disappear: Konishi *et al.* (1997) have shown that if the policy space has two alternatives, there is always an efficient Nash equilibrium. In the appendix, we show that this result can be extended to a compact policy space in the case of two jurisdictions.[33] The generalization to more than two jurisdictions and two alternatives requires more structure on the profile of preferences. Indeed, as shown in Konishi *et al.* (1997), in general a Nash equilibrium may fail to exist. We assume that the preferences of each jurisdiction i is given by the

[33] Upper semi-continuity of preferences has to be assumed to guarantee the existence of best responses. A function f is upper semi-continuous on X if for all $x_0 \in X$, $\limsup_{x \to x_o} f(x) \leq f(x_0)$.

following utility function:

$$U_i(x) = V(\theta_i, x_i) + W_i(\{j : x_j = x_i\}), \tag{5}$$

where the policy space X is compact (possibly multidimensional) V is continuous in x and W is increasing (in the inclusion sense). The function V introduces idiosyncratic heterogeneity through the vector of type $\theta \in \mathbb{R}^I$ as in (3) while W embodies the network effects.

Definition 4 *A function $V : X \times \Theta$ has monotonic differences if for all $x, y \in X$ and $\theta \in \Theta$, $V(x, \theta) - V(y, \theta)$ is monotonic in θ.*

Monotonic differences is the equivalent of the single crossing property but we do not need to impose any differentiability requirement. Observe that it does not impose any restriction on the degree of heterogeneity of local preferences since the set of admissible types Θ can be any subset of the real line.[34]

Definition 5 *A profile of network effects W is weakly congruent if for all $i, j \in \{1, .., I\}$, for all $S, T \subset \{1, .., I\} \setminus \{i, j\}$,*

$$W_j(T \cup \{i, j\}) - W_i(T \cup \{i\}) + W_i(S \cup \{i, j\}) - W_j(S \cup \{j\}) \geq 0.$$

As W is increasing, W is weakly congruent in particular if it is uniform, i.e. if $W_i = W_j$ for all $i \neq j$. Hence, a profile of network effects is weakly congruent if it is not too asymmetric. Observe that we do not impose any restriction on the magnitude of these effects. Indeed if W is weakly congruent,

[34] Our results can be generalized to multidimensional types under some intermediate preferences assumption as in Demange 1990.

for any $\gamma \geq 0$, γW is weakly congruent as well. Using a result of Greenberg and Weber (1986), we can prove the following:

Proposition 11 *Under the specification in (5), if V has monotonic differences and if W is weakly congruent then there exists a strong Nash equilibrium.*[35]

Under the condition of proposition 11, if coalitions can form easily, decentralization will result in a stable and Pareto optimal outcome. notice that in our context, the amount of coordination required to get to a Strong Nash equilibrium may be achieved by the federal regulatory bodies that foster partial uniformity through model laws or recommendations – see subsection 4.1. In any case, a coercive intervention is not a necessary condition for efficiency. Similarly to the case of smooth coordination costs, the symmetry of external effects makes the case for decentralized policy making whatever the importance of coordination and the heterogeneity of local preferences.

6 Extensions

6.1 Local Heterogeneity

In this subsection, we model local policy making through local elections in an extended model which encompasses intrajurisdiction heterogeneity and different population sizes. Each jurisdiction i is now comprised of a finite number P_i of residents indexed by a set V_i. The preferences of voter v in

[35] A strong Nash equilibrium is a profile of strategy immune to deviations by coalitions of players (Aumann 1959). In particular, it is Pareto efficient.

jurisdiction i are given by (3) with parameters $\left(\beta^v_{i,j}\right)_{j\neq i}$ and θ^v_i while $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ refer to the corresponding profiles. We assume that under decentralization, decisions within each jurisdiction are taken through local referenda by local majority rule (henceforth abbreviated LM). Decentralization is then defined as follows:

Definition 6 *A policy vector x is a local majority decentralized equilibrium (LMDE) if for all i, x is LM preferred in jurisdiction i to (x'_i, x_{-i}) for any x'_i.*

Both the heterogeneity and the space of alternatives are multidimensional, so LM preferences over policy vectors are generically intransitive. Nevertheless, we prove in the appendix that for all $(\alpha, \boldsymbol{\beta}, \boldsymbol{\theta})$, there exists a unique LMDE. However, heterogeneity within jurisdictions makes the Pareto welfare analysis vacuous: if types are sufficiently spread, any policy can be Pareto optimal. To overcome this problem, we refine the Pareto domination criterion and make further assumptions on the distribution of local preferences.

Definition 7 *A policy vector x LM Pareto dominates y if x is LM preferred to y in all jurisdictions.*

Definition 8 *A profile of preferences is locally symmetry if for all i, the (D+1)-tuples $\left(\left(\beta^v_{i,j}\right)_{j\neq i}, \theta^v_i\right)_{v\in V_i}$ are symmetrically distributed around some $\left(\beta_{i,j}\right)_{j\neq i}, \theta_i$ which we call the local median voter of jurisdiction i.*

Observe that as defined in (3), the preferences depend linearly on (β, θ) – modulo a constant term. Hence this class of preferences belongs to the class of intermediate preferences introduced by Grandmont (1978): under the assumption of local symmetry, the majority preferences of each jurisdiction are exactly the preferences of its median voter. From what precedes, LM Pareto domination is equivalent to Pareto domination among local median voters. Therefore, the results of the Pareto welfare analysis of subsection 4.1 and section 5 can be readily extended to the LM Pareto criterion by replacing jurisdiction preferences by local median preferences: in particular, from proposition 3, if $\alpha \geq 2$, for any uniform policy there is a jurisdiction in which a majority of voters prefer the LMDE. The Benthamite welfare analysis of subsection 4.2 carries over as well since under the local symmetry assumption, the Benthamite welfare of each jurisdiction is an affine function of the welfare of its median voter. However, the social weight of each jurisdiction i is now equal to P_i. For instance from proposition 8, if $\alpha \geq 2$ and if for all $i \neq j$, $P_i \times \beta_{i,j} = P_j \times \beta_{j,i}$, i.e. if external costs are reciprocal in aggregate terms, the LMDE is Benthamite preferred to any uniform policy.[36]

6.2 Conflict of Laws and the Country of Origin Principle

As an alternative to federal uniformization of legal systems, some federations have put forward a more flexible solution to mitigate conflict of laws and regulatory compliance costs. It consists in the mutual recognition of regulatory regimes, sometimes referred to as the Country of Origin Principle (thereafter

[36]The necessary and sufficient conditions of proposition 8 become: for all i, $\sum_{j=1}^{I} \left(\Pi_i \beta_{i,j} - \Pi_j \beta_{j,i}\right) \geq -\Pi_i$.

CoOP).[37] Under the CoOP exporting firms have to conform to the laws and regulations of the country of origin, while under the traditional Country of Reception Principle (CoRP) they have to conform to the requirements of the country which they export to.

Clearly, the CoOP facilitates trade since it avoids manufacturers and service providers from the obligation to accommodate multiple regulatory regimes when trading across borders. However, as an indirect effect, it changes the incentives faced by states when enacting their regulation: the CoOP shifts the burden from the exporting firms to the importing consumers. Under the CoRP, a state has incentive to improve the compatibility of its regulation with the laws of the countries where its firms are exporting: it wishes to minimize their compliance costs. Contrarily, under the CoOP, it has less incentive, if any at all, to take into account the suitability of its regulation to the specific needs of the states to which its firms are exporting since the corresponding costs and benefits are mostly borne outside its borders.[38]

[37] See Kox, Lejour and Montizaan 2004 and the Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000.

[38] One could argue that a regulation enacted by a state A which would not meet the specific needs of a state B would be detrimental to the quality of the product of the exporting firms of state A, which would lower their profit. Pursuing the reasoning, this loss should be internalized by state A. However, if it were the case, such a regulation would be useless in the first place. Indeed the aim of most regulatory requirements is to remedy asymmetries of information between producers and consumers, or consumption externalities, i.e. market imperfection that cannot be solved through the price system. Hence, most of the costs and benefits of the corresponding regulation will be borne by economic agents of state B rather than the firms of state A. Naturally, this is not true of laws regulating production externalities such as pollution, which costs are borne mostly

Under the CoOP, the externalities suffered by some jurisdiction i are no longer a function of the differences between its policy x_i and the policies x_{-i} of the jurisdictions which its firms are exporting to – since there are no compliance costs. Instead, externalities are function of the differences between the policies x_{-i} of the jurisdictions it is importing from and its most preferred policy θ_i. Formally, the utility function in (3) would get closer to the following:

$$U_i(x) = -|x_i - \theta_i|^\alpha - \sum_{j \neq i} \beta_{i,j} |x_j - \theta_i|^\alpha \tag{6}$$

where $\beta_{i,j}$ depends on the amount of services imported from jurisdiction j by jurisdiction i. The decentralized equilibrium would then be $x^{dec} = \theta$: in words, the CoOP removes any incentives to harmonize policies. As the externality parameters (α, β) do not have the same interpretation under both principles, we do not venture into a direct welfare comparison. Nevertheless, to assess the impact of the CoOP, policy makers should balance the direct effect – i.e. reduced barriers to trade – with the indirect strategic effect – i.e. reduced incentives for regulation harmonization.

Furthermore, the determinants of the relative advantage of policy uniformization versus decentralization under the CoOP would dramatically change. The following proposition shows that they are more in line with O.D.T. and the degree of symmetry of externalities becomes irrelevant.

Proposition 12 *Given θ, the Benthamite welfare under any uniform policy and under decentralization is a function of the symmetric part of β only.*[39] *Decentralization is Benthamite preferred to any uniform policy if $\|\beta\|$ is small*

in state A.

[39]The symmetric part β^S of a matrix β is defined by $\beta^S_{i,j} = \frac{1}{2}\left(\beta_{i,j} + \beta_{j,i}\right)$ for all i, j.

enough while it is Pareto dominated by a uniform policy if $\|\beta\|$ is large enough.

More generally, similar comparative statics apply to all policies which generate cross-border externalities strategically equivalent to the one described in (6). This is in particular the case of what has been termed as third party or psychological externalities (see Mueller 1997): the laws on abortion, religion and other ideological issues enacted in a given jurisdiction may have no physical cross-border effects yet they may exert an acute psychological pain on the residents of the other member of the confederation. Contrarily to coordination externalities, a citizens of ideology θ_i prefers the policies of the other jurisdictions x_{-i} to be as close as possible to θ_i rather than to their actual policy x_i, which is consistent with (6).

7 Concluding Remarks

We have studied the welfare properties of unitarian centralization versus non cooperative decentralization in an heterogeneous confederation with coordination externalities. Even under very conservative institutional assumptions, the comparative statics of the decentralization debate exhibit important discrepancies with respect to the usual fiscal federalism scenario. The coordination failure of decentralization vanishes as externalities become arbitrarily large. More surprisingly, whenever coordination costs are sufficiently elastic and reciprocal, decentralization is socially preferred to any uniform policy whatever the magnitude of externalities and the heterogeneity of local preferences.

Hence, the determinants of the decentralization debate are quite different whether we consider fiscal policies or administrative rules, laws and foreign policies. More generally our analysis applies to any policy for which the interjurisdictional conflict of interest is closer a battle of the sexes than to a prisoner's dilemma. The case for decentralization appears to be stronger in the earlier case.

The behavior of local and federal governments have been modelled as simply as possible in order to make our results easily comparable to the existing literature. The next step is naturally to endogenize the respective cost of centralization and decentralization by modelling in more detail the political process at the local and federal level.

8 Appendix

8.1 Proof of proposition 1

The following lemma proves proposition 1 along with some additional properties of the decentralized equilibrium that will be useful in the sequel.

Lemma 1 *For all $\alpha > 1$ and for all (β, θ), there exists a unique decentralized equilibrium. The corresponding map $\theta \to x^{dec}(\alpha, \beta, \theta)$ is onto and one to one.*

Proof. As $\alpha > 1$, for all $i \neq j$, the term $|x_i - x_j|^\alpha$ has increasing differences in (x_i, x_j), hence the decentralization game between jurisdictions exhibits strategic complementarities. Moreover, the payoffs in (3) are continuous in x and the strategy space can be restricted to $[\min_i \theta_i; \max_i \theta_i]^I$

w.l.o.g.[40] Finally, the game is supermodular and a pure strategy Nash equilibrium exists.

To prove the rest of the proposition, fix (α, β) and consider the correspondence $\Gamma : \mathbb{R}^I \rightrightarrows \mathbb{R}^I$ which maps every policy vector x to the set of vectors θ such that x is an equilibrium for (α, β, θ). As an equilibrium exists for all θ, Γ is onto. Since utility functions are concave and differentiable, the Nash equilibria are exactly the solutions of the first order conditions for the local median voters:

$$\text{for } i = 1..I,\ x_i - \theta_i + \sum_{j \neq i} \frac{1}{2} \alpha \beta_{i,j} (x_i - x_j)^{\alpha - 1} = 0, \tag{7}$$

with the understanding that $(x_i - x_j)^{\alpha-1} < 0$ whenever $x_i < x_j$. From (7), one can easily express θ as a function of x which shows that Γ is well defined on $\mathbb{R}^I$ and single-valued. The corresponding map is differentiable on $\mathbb{R}^I \backslash \Theta$, where $\Theta = \{x : \exists i, j : x_i = x_j\}$ and for all $x \in \mathbb{R}^I \backslash \Theta$, for all i:

$$\begin{aligned} \frac{\partial \Gamma_i}{\partial x_j}(x) &= -\frac{1}{2} \alpha (\alpha - 1) \beta_{i,j} |x_i - x_j|^{\alpha - 2} \quad \text{for } j \neq i, \\ \frac{\partial \Gamma_i}{\partial x_i}(x) &= 1 + \sum_{\alpha \neq i} \left| \frac{\partial \Gamma_i}{\partial x_j}(x) \right|. \end{aligned} \tag{8}$$

From (8) the Jacobian matrix $\frac{\partial \Gamma}{\partial x}$ has a strictly dominant positive diagonal. Now let $x, y \in \mathbb{R}^I$, and let $[x, y]$ be the corresponding line segment. As Θ is a finite union of linear subspace, either $[x, y] \subset \Theta$ or $[x, y] \cap \Theta$ is a finite set. Suppose the latter case. The map $u \rightarrow \Gamma(x + u.(y - x))$ is continuous and

[40] Any equilibrium strategy profile in $[\min_i t_i; \max_i t_i]^I$ is an equilibrium in $\mathbb{R}^I$ and no strategy profile outside this set can be an equilibrium.

differentiable almost everywhere on $[0,1]$ and we have:

$$\begin{aligned}\Gamma(y)-\Gamma(x) &= \int_0^1 \left[\frac{\partial\Gamma}{\partial x}(x+u.(y-x))\times(y-x)\right]du \\ &= \left[\int_0^1 \frac{\partial\Gamma}{\partial x}(x+u.(y-x))\,du\right]\times(y-x). \qquad (9)\end{aligned}$$

The matrix $\int_0^1 \frac{\partial\Gamma}{\partial x}(x+u.(y-x))\,du$ is strictly diagonally dominant as a sum of positive strictly diagonally dominant matrices. As such it is invertible (see Taussky 1949) and from (9), $\Gamma(x)=\Gamma(y)$ implies $x=y$.

Suppose else that $[x^1,x^2]\subset\Theta$. Although Γ is not differentiable on Θ, one can easily check that Γ is differentiable along Θ. Integrating Γ along $[x^1,x^2]$ as above yields the same conclusion. Finally, we have proven that Γ is onto and one to one. Therefore its inverse $\theta\rightarrow x^{dec}(\alpha,\beta,\theta)$ is well defined, onto and one to one. ■

8.2 Proofs in Subsection 4.1

The proof of proposition 2 follows directly from the fact that externalities are not internalized under decentralization while a non dictatorial uniform policy can be Pareto improved by infinitesimally moving local policies towards local their respective type. We ommit the detail for brevity.

Lemma 2 *Given (α,β,θ), let x be the decentralized equilibrium, $(u,...,u)$ a uniform policy, U^{dec} and U^{c} the corresponding profile of utility, then*

$$\begin{aligned}U_i^{dec}-U_i^{c} &= |x_i-u|^2+\sum_{j\neq i}W_{i,j}, \\ \textit{where } W_{i,j} &= \frac{\alpha.\beta_{i,j}}{2}|x_i-x_j|^{\alpha}\left[1-\frac{2}{\alpha}+\frac{|x_i-u|^2-|x_j-u|^2}{|x_i-x_j|^2}\right],\end{aligned}$$

with the convention that $W_{i,j}=0$ if $x_i=x_j$.

Proof. From (7):

$$2\left(x_i-\theta_i\right)=-\sum_j \alpha\beta_{i,j}\left(x_i-x_j\right)^{\alpha-1},$$

with the understanding that $(x_i - x_j)^{\alpha-1} < 0$ whenever $x_i < x_j$. Decomposing U_i^c and using the identity above, we get

$$-U_i^c=\left|u-\theta_i\right|^2=\left|u-x_i\right|^2+\left|x_i-\theta_i\right|^2-\left(u-x_i\right)\sum_j \alpha\beta_{i,j}\left(x_i-x_j\right)^{\alpha-1},$$

Adding U_i^{dec}, we obtain

$$U_i^{dec}-U_i^c=\left|x_i-u\right|^2-\sum_j \beta_{i,j}\left(x_i-x_j\right)^{\alpha-1}\left[\alpha\left(u-x_i\right)+\left(x_i-x_j\right)\right].$$

Finally, simple algebra shows that whenever $x_i \neq x_j$,

$$\alpha\left(u-x_i\right)+\left(x_i-x_j\right)=-\frac{\alpha}{2}\left(x_i-x_j\right)\left[1-\frac{2}{\alpha}+\frac{\left|x_i-u\right|^2-\left|x_j-u\right|^2}{\left|x_i-x_j\right|^2}\right].$$

■

8.2.1 Proof of Proposition 3

For a given salience matrix α, let us define the following propositions:

$P^1(\alpha,\beta)$:= "There exists a vector of type θ and a real number u such that the uniform policy $(u,..,u)$ strictly Pareto dominates the decentralized equilibrium for the profile of preferences (α,β,θ)."

$P^2(\alpha,\beta)$:= "There exists a vector of policies x and a real number u such that for all i

$$\left|x_i-u\right|^2+\sum_{j\neq i}\frac{\alpha.\beta_{i,j}}{2}\left|x_i-x_j\right|^{\alpha}\left[1-\frac{2}{\alpha}+\frac{\left|x_i-u\right|^2-\left|x_j-u\right|^2}{\left|x_i-x_j\right|^2}\right]<0, \quad (10)$$

with the convention that the summand is null whenever $x_i = x_j$."

From lemma 1 we know that for all (α, β) the mapping $\theta \to x^{dec}(\alpha, \beta, \theta)$ is onto. So from lemma 2, P^1 and P^2 are equivalent and it suffices to show that $P^2(\alpha, \beta)$ is true i.f.f. $\alpha < 2$.

If $\alpha \geq 2$, the first term in the left-hand side of (10) is non-negative, and the second term is a sum of negative term for the jurisdiction i such that $i \in \arg\max\left(|x_i - u|^2\right)$. If $\alpha < 2$, let x be a vector of policies and u, h be two scalars such that $x_i - u$ is equal to h for some jurisdictions and to $-h$ for the others. For any jurisdiction i, the second term of the left-hand side of (10) can be written $|h|^2 + \gamma_i |h|^\alpha$ for some $\gamma_i > 0$. As $\alpha < 2$, one can choose h sufficiently small so that the latter expression is negative for all i.

8.2.2 Proof of Proposition 4

Let α, $(\beta^n)_{n\geq 0}$ and θ be as defined in proposition 4. We denote $x^n = x^{dec}(\alpha, \beta^n, \theta)$.

Lemma 3 *For all $j \neq i$, $\beta^n_{i,j} \left|x^n_i - x^n_j\right|^{\alpha-1}$ is bounded and $\beta^n_{i,j} \left|x^n_i - x^n_j\right|^{\alpha-1} \to 0$.*

Proof. To simplify notations, throughout the proof we reorder the indexes of jurisdictions so that $i \leq j \Rightarrow x^n_i \leq x^n_j$. Observe that the reordering may not be independent of n. With a slight abuse of notation, we define the sequences of preferences profiles $(\alpha, \beta^n, \theta^n)_{n\geq 0}$ and the sequence of decentralized equilibria $(x^n)_{n\geq 0}$ consistently with this reindexation. It amounts to a symmetric permutation of rows and columns of β^n. So for all $n \geq 0$, we still have that for all $i \neq j$, $\beta^n_{i,j} \to +\infty$ and $\beta^n_{i,j} = O\left(\beta^n_{j,i}\right)$. Moreover, for all n, θ^n_i and x^n_i lies within $[\min_i \theta_i, \max_i \theta_i]$. We will prove the lemma with the new indexation. One can easily see that it implies the lemma with the original one. We prove the lemma by induction on i.

For any n, by construction $x^n_1 \leq x^n_j$. So from (7),

$$2\left|x^n_1 - \theta^n_1\right| = \sum_{j>1} \alpha\beta^n_{1,j}\left|x^n_1 - x^n_j\right|^{\alpha-1}. \tag{11}$$

From what precedes, the left-hand side of (11) is bounded, so the right-hand side as well. It implies that for all j, $\beta^n_{1,j}\left|x^n_1 - x^n_j\right|^{\alpha-1}$ is bounded. In particular, for all $i \neq j$, $\left|x^n_i - x^n_j\right| \to 0$, which in turn implies $\beta^n_{1,j}\left|x^n_1 - x^n_j\right|^{\alpha} \to 0$. This proves the lemma for $i = 1$. In the case of two jurisdictions, a symmetric argument for $i = 2$ completes the proof.

Now assume that it is true for $1, .., i-1$. For any n, by construction $x^n_i \geq x^n_j$ for $i > j$ and $x^n_i \leq x^n_j$ for $i < j$. So from (7),

$$\pm 2\left|x^n_i - \theta^n_i\right| + \alpha\sum_{j<i} \beta^n_{i,j}\left|x^n_i - x^n_j\right|^{\alpha-1} = \alpha\sum_{j>i} \beta^n_{i,j}\left|x^n_i - x^n_j\right|^{\alpha-1}. \tag{12}$$

From the induction hypothesis, for all $j < i$, $\alpha\beta^n_{j,i}\left|x^n_j - x^n_i\right|^{\alpha-1}$ is bounded. As $\beta^n_{i,j} = O\left(\beta^n_{j,i}\right)$, necessarily $\alpha\beta^n_{i,j}\left|x^n_i - x^n_j\right|^{\alpha-1}$ is bounded as well. Hence the left-hand side of (12) is bounded, which proves that for all $j > i$, $\beta^n_{i,j}\left|x^n_i - x^n_j\right|^{\alpha-1}$ is bounded. As $\left|x^n_i - x^n_j\right| \to 0$, we have $\beta^n_{i,j}\left|x^n_i - x^n_j\right|^{\alpha} \to 0$ for all $j \neq i$, which completes the induction.

■

Lemma 3 implies that at the decentralized equilibrium, coordination costs vanish. Therefore, we have that for all i, $U^{dec}_{n,i} + (x^n_i - \theta^n_i)^2 \to 0$, so U^{dec}_n is bounded below.

Let $(y^n)_{n\geq 0}$ be a sequence of policy vectors which Pareto improves on $(x^n)_{n\geq 0}$ and denote $(U_n^*)_{n\geq 0}$ the corresponding sequence of utility profile. As U_n^{dec} is bounded below, so is U_n^* and necessarily, for all $i \neq j$, $\left|y_i^n - y_j^n\right| \to 0$.

As $(x^n)_{n\geq 0}$, $(y^n)_{n\geq 0}$, $\left(U_n^{dec}\right)_{n\geq 0}$ and $(U_n^*)_{n\geq 0}$ are bounded, there exists a subsequence $f \in \mathbb{N}^{\mathbb{N}}$ such that these four sequences converges. Denote V^{dec} and V^* the limit of the two latter sequences. From lemma 3, $\left(x^{f(n)}\right)_{n\geq 0}$ converges to a uniform vector $(u, .., u)$ and for all i, $V_i^{dec} = -(u-\theta_i)^2$. From what precedes, $\left(y^{f(n)}\right)_{n\geq 0}$ converges to a uniform vector $(v, .., v)$ and for all i, $V_i^* \leq -(v-\theta_i)^2$. By construction $V^{dec} \leq V^*$, which implies that u and v are such that for all i, $(v-\theta_i)^2 \leq (u-\theta_i)^2$ which implies $u = v$ and $V^{dec} = V^*$.

We shall now construct an example to show that one cannot dispense with the condition $\beta_{i,j}^n = O\left(\beta_{j,i}^n\right)$ for all $i \neq j$. We set $\alpha = 2$, $t_1 = -1$, $t_I = 1$ and all other types equal 0. For all j, $\beta_{1,j} = \beta_{I,j} = \underline{\beta}$ and for all $i = 2, .., I-1$, $\beta_{i,j} = \overline{\beta}$. Solving for the decentralized equilibrium yields $x_I = -x_1 = \left(1 + 2(I-1)\underline{\beta}\right)^{-1}$ and $x_i = 0$ for the other jurisdictions. Solving for $\max(U_1 + U_I)$ yields $y_I = -y_1 = \left(1 + 4(I-1)\underline{\beta}\right)^{-1}$ and $y_i = 0$ for the other jurisdictions. By construction, jurisdictions 1 and I prefer y to x. For $1 < i < I$, $U_i(x) = -\frac{2\overline{\beta}}{\left(1+2(I-1)\underline{\beta}\right)^2}$ and $U_i(y) = -\frac{2\overline{\beta}}{\left(1+4(I-1)\underline{\beta}\right)^2}$ so $U_i(x) < U_i(y)$ and y Pareto dominates x. If we let $\frac{\overline{\beta}}{\underline{\beta}^2} \to +\infty$, from what precedes, $U_i(y) - U_i(x) \to +\infty$.

8.3 Proofs in Subsection 4.2

8.3.1 Proofs of Propositions 5 and 7

We use the notation $\Delta x = x_1^{dec} - x_2^{dec}$ and likewise for $\Delta\theta$.

Lemma 4 *i)* Δx *is constant in* β_A *and decreasing in* β_S*. When* $\beta_S = 0$*,* $\Delta x = \Delta\theta$*. When* $\beta_S \longrightarrow +\infty$*,* $\Delta x \longrightarrow 0$

ii) The utility of jurisdiction 1 under decentralization can be expressed as function V_1^{dec} *of* $\alpha, \beta_A, \Delta\theta$ *and* Δx *only:*

$$V_1^{dec} = -\frac{1}{4}(1+\beta_A)^2(\Delta\theta - \Delta x)^2 - \frac{1+\beta_A}{\alpha}(\Delta\theta - \Delta x)\Delta x, \quad (13)$$

$$\frac{\partial V_1^{dec}}{\partial \Delta x} = \frac{1}{2}(1+\beta_A)\left(1+\beta_A - \frac{2}{\alpha}\right)(\Delta\theta - \Delta x) + \frac{(1+\beta_A)}{\alpha}\Delta x. \quad (14)$$

The same expression holds for jurisdiction 2 by replacing β_A *by* $-\beta_A$*.*

iii) If $\beta_A \neq 0$*,* $\left|V_1^{dec} - V_2^{dec}\right|$ *is decreasing in* Δx*.*

iv) The socially optimal uniform policy is $\left(\frac{\theta_1+\theta_2}{2}, \frac{\theta_1+\theta_2}{2}\right)$ *and the corresponding profile of utilities is* $V_1^c = V_2^c = -\frac{1}{4}\Delta\theta^2$*, with*

$$\left(V_1^{dec} + V_2^{dec}\right)(\Delta x = 0) \leq V_1^c + V_2^c < \left(V_1^{dec} + V_2^{dec}\right)(\Delta x = \Delta\theta),$$

the first inequality being strict i.f.f. $\beta_A \neq 0$*.*

v) For all $\overline{V} \in \left]V_1^c + V_2^c, 0\right[$*,* $\left(V_1^{dec} + V_2^{dec}\right)(\Delta x) \geq \overline{V}$ *i.f.f.* $\Delta x \geq \widetilde{\Delta x}$ *for some* $\widetilde{\Delta x} \in \left]0, \Delta\theta\right[$*.*

vi) $\left|V_1^{dec} - V_2^{dec}\right|$ *(resp.* $V_1^{dec} + V_2^{dec}$*) is even in* β_A*, increasing (resp. decreasing) in* $|\beta_A|$*.*

Proof. Taking differences in (7), we have $\Delta\theta - \Delta x = a\beta_S \Delta x^{a-1}$ which proves *i*). From (7), $|\theta_i - x_i| = \frac{\alpha\beta_i|\Delta x|^{a-1}}{2}$ and

$$U_i^{dec} = -\left(\frac{\alpha\beta_i\left|\Delta x\right|^{a-1}}{2}\right)^2 - \beta_i\left|\Delta x\right|^{\alpha}.$$

As $\beta_i = \beta_S (1 \pm \beta_A)$ and $\Delta\theta - \Delta x = a\beta_S |\Delta x|^{a-1}$ we have

$$\begin{aligned} V_i^{dec} &= -\frac{1}{4}(1 \pm \beta_A)^2 (\Delta\theta - \Delta x)^2 - \frac{1 \pm \beta_A}{\alpha}(\Delta\theta - \Delta x)\Delta x, && (15) \\ V_1^{dec} + V_2^{dec} &= -\frac{1}{2}\left(1 + \beta_A^2\right)(\Delta\theta - \Delta x)^2 - \frac{2}{\alpha}(\Delta\theta - \Delta x)\Delta x, && (16) \\ V_1^{dec} - V_2^{dec} &= -\beta_A (\Delta\theta - \Delta x)^2 - \frac{2\beta_A}{\alpha}(\Delta\theta - \Delta x)\Delta x, && (17) \end{aligned}$$

which immediately proves *ii*) and *vi*). From (17),

$$\frac{\partial \left|V_1^{dec} - V_2^{dec}\right|}{\partial \Delta x} = -\beta_A \left[\left(2 - \frac{1}{\alpha}\right)(\Delta\theta - \Delta x) + \frac{\Delta x}{\alpha}\right],$$

which proves *iii*).

The uniform policy which maximizes $U_1 + U_2$ is $\left(\frac{\theta_1+\theta_2}{2}, \frac{\theta_1+\theta_2}{2}\right)$ and it entails $U_1 = U_2$. Therefore it is the socially optimal uniform policy and it yields $V_i^c = -\frac{1}{4}\Delta\theta^2$.

From (16),

$$\begin{aligned} \left(V_1^{dec} + V_2^{dec}\right)(\Delta x &= \Delta\theta) = 0, \\ \left(V_1^{dec} + V_2^{dec}\right)(\Delta x &= 0) = -\frac{1}{2}\left(1 + \beta_A^2\right)\Delta\theta^2. \end{aligned}$$

Since $\Delta\theta \neq 0$ and $V_1^c + V_2^c = -\frac{1}{2}\Delta\theta^2$, point *iv*) follows.

From (16),

$$\begin{aligned} \frac{\partial\left(V_1^{dec} + V_2^{dec}\right)}{\partial \Delta x} &= \left(\beta_A^2 + \frac{\alpha - 2}{\alpha}\right)(\Delta\theta - \Delta x) + \frac{2}{\alpha}\Delta x, \\ \frac{\partial^2\left(V_1^{dec} + V_2^{dec}\right)}{\partial \Delta x^2} &= \left(\frac{4 - \alpha}{\alpha} - \beta_A^2\right). \end{aligned}$$

So $V_1^{dec} + V_2^{dec}$ is either concave or convex and it is increasing in Δx at $\Delta x = \Delta\theta$. Therefore, it is either increasing or decreasing and increasing afterwards. Together with point *iv*), it implies point *v*).

■

In the notation of lemma $4i)$, given α, β_A, θ, V_i^{dec} depends on β_S only through Δx. As Δx is decreasing in β_S holding α, β_A, θ constant, comparative statics on U_i^{dec} w.r.t. $\beta_S \in \mathbb{R}_+^*$ can be deduced from comparative statics on V_i^{dec} w.r.t. $\Delta x \in [0, \Delta\theta]$.

From (14) V_1^{dec} (resp. V_2^{dec}) is increasing in Δx on $[0, \Delta\theta]$ i.f.f. $1+\beta_A - \frac{2}{\alpha} > 0$ (resp.$1 - \beta_A - \frac{2}{\alpha} > 0$). If the inequality does not hold, by differentiating (14) w.r.t. Δx, V_i^{dec} is convex. Necessary it is decreasing for $\Delta x \leq \widetilde{\Delta x}$ and increasing for $\Delta x \geq \widetilde{\Delta x}$ for some $\widetilde{\Delta x} \in]0, \Delta\theta[$. To complete the proof of proposition $5ii)$, notice that $1 + \beta_A - \frac{2}{\alpha}$ and $1 - \beta_A - \frac{2}{\alpha}$ are both positive i.f.f. $|\beta_A| < \frac{2-\alpha}{\alpha}$.

To prove proposition $5i)$, it suffices to show that under the notation of lemma 4, $S\left(V_1^{dec}, V_2^{dec}\right)$ is greater than $S\left(V_1^c, V_2^c\right)$ if and only $\Delta x > \widetilde{\Delta x}$ for some $\widetilde{\Delta x} \in]0, \Delta\theta[$. Assume $\beta_A \neq 0$, from lemma $4iv)$, $S^{dec}(\Delta x = 0) < S^c$ and $S^{dec}(\Delta x = \Delta\theta) > S^c$. By continuity there exists $\widetilde{\Delta x} \in]0, \Delta\theta[$ such that $S^{dec} > S^c$ for all $\Delta x > \widetilde{\Delta x}$, and $S^{dec} = S^c$ at $\Delta x = \widetilde{\Delta x}$. It remains to show that $S^{dec} < S^c$ for all $\Delta x < \widetilde{\Delta x}$. As $\left|V_1^{dec} - V_2^{dec}\right|\left(\widetilde{\Delta x}\right) > 0$ and $S^{dec}\left(\widetilde{\Delta x}\right) = S^c\left(\widetilde{\Delta x}\right)$, the concavity of S implies $\left(V_1^{dec} + V_2^{dec}\right)\left(\widetilde{\Delta x}\right) \geq V_1^c + V_2^c$. The latter inequality, together with lemma $4v)$ implies that for all $\Delta x < \widetilde{\Delta x}$, $\left(V_1^{dec} + V_2^{dec}\right)\left(\widetilde{\Delta x}\right) > \left(V_1^{dec} + V_2^{dec}\right)(\Delta x)$. From lemma $4iii)$, for all $\Delta x < \widetilde{\Delta x}$, $\left|V_1^{dec} - V_2^{dec}\right|\left(\widetilde{\Delta x}\right) < \left|V_1^{dec} - V_2^{dec}\right|(\Delta x)$. As S is increasing and concave, $S^{dec}(\Delta x) < S^{dec}\left(\widetilde{\Delta x}\right) = S^c$ for $\Delta x < \widetilde{\Delta x}$ which completes the proof of proposition $5i)$.

As the jurisdictions' welfare under a uniform rule is independent of β, lemma $4vi)$ together with the symmetry and concavity of S proves proposition $7i)$. If $\beta_A \neq 0$, from proposition $5i)$ centralization is socially preferred to decentralization as β_S tends to $+\infty$. If $\beta_A = 0$ and $\alpha < 2$, from lemma 4iv), $\left(V_1^{dec}, V_2^{dec}\right)$ tends toward (V_1^c, V_2^c) as Δx tend to 0. However, as both $\beta_A + \frac{\alpha-2}{\alpha}$ and $-\beta_A + \frac{\alpha-2}{\alpha}$ are negative, from the proof of proposition 5ii) above, $\left(V_1^{dec}, V_2^{dec}\right)$ is strictly decreasing in Δx at $\Delta x = 0$. So for Δx close enough to 0, $\left(V_1^{dec}, V_2^{dec}\right) < (V_1^c, V_2^c)$, which completes the necessary part of point ii) of proposition 7. The sufficiency part comes from proposition 8 together with the fact that if $\beta_A = 0$, $V_1^{dec} = V_2^{dec}$.

8.3.2 Proof of Proposition 6

Lemma 5 *For all* (α, β, θ), $h > 0$ *and* $\mu \in \mathbb{R}$, *let* $\overrightarrow{1} = (1, .., 1)$, $\theta' = h\theta + \mu\overrightarrow{1}$ *and for all* $i \neq j$, $\beta'_{i,j} = h^{2-\alpha}\beta_{i,j}$, *then*

$$h^2 U_{(\alpha,\beta,\theta)}(x) = U_{(\alpha,\beta',\theta')}\left(hx + \mu\overrightarrow{1}\right).$$

Proof. For all i,

$$\begin{aligned} U_{(\alpha,\beta',\theta'),i}(hx + \mu\overrightarrow{1}) &= -h^2 |x_i - \theta_i|^2 - \sum_{j \neq i} h^{2-\alpha}\beta_{i,j} h^\alpha |x_i - x_j|^\alpha \\ &= h^2 U_{(\alpha,\beta,\theta),i}(x). \end{aligned}$$

■

Under the notation of lemma 5, the preferences induced by the profile $(\alpha, \beta', \theta')$ composed with the map $x \to hx + \mu\overrightarrow{1}$ are exactly the preferences of (α, β, θ). Let S be a social welfare function which is neutral w.r.t. the alternatives. By definition, the corresponding social welfare ordering $\succeq^S$ is a function of ordinal preferences only, therefore:

$$x \succeq^S_{(\alpha,\beta,\theta)} y \Leftrightarrow hx + \mu\overrightarrow{1} \succeq^S_{(\alpha,\beta',\theta')} hx + \mu\overrightarrow{1}.$$

Together with lemma 5, this proves that

$$S(U) \geq S(V) \Leftrightarrow S(\gamma U) \geq S(\gamma V). \tag{18}$$

Nash equilibrium being an ordinal concept, $x^{dec}(\alpha, \beta', \theta') = hx^{dec}(\alpha, \beta, \theta) + \mu\overrightarrow{1}$, so $U^{dec}_{(\alpha,\beta',\theta')} = h^2 U^{dec}_{(\alpha,\beta,\theta)}$. Moreover, the same relation holds for the optimal uniform policy. Using (18), decentralization is socially preferred to any uniform rule at (α, β, θ) i.f.f. it is true at $(\alpha, \beta', \theta')$. Using the notation of proposition 6, we have that for all $h > 0$, $\mu \in \mathbb{R}$, decentralization is S-preferred to any uniform rule at $\left(\alpha, \beta_A, \beta_S, h\theta + \mu\overrightarrow{1}\right)$ i.f.f. it is true at $(\alpha, \beta_A, h^{\alpha-2}\beta_S, \theta)$. Hence, as far as the comparison between centralization and decentralization is concerned, increasing $\Delta\theta$ is tantamount to decreasing (resp. increasing) β_S if $\alpha < 2$ (resp. if $\alpha > 2$). Proposition 6 is then an immediate corollary from proposition 5*i*).

8.3.3 Proofs of Proposition 8

With the notations of lemma 2, if β is symmetric, for all $i \neq j$, $V_{i,j} + V_{j,i} = (\alpha - 2)\beta_{i,j}\left|x^{dec}_i - x^{dec}_j\right|^\alpha$ which is positive if $\alpha \geq 2$. If B^{dec} and B^c are the Benthamite welfare respectively under decentralization and under the uniform policy $(u, .., u)$,

$$B^{dec} - B^c = \sum_{i=1..I}\left|x^{dec}_i - u\right|^2 + \sum_{i=1..I}\sum_{j\neq i}\frac{V_{i,j} + V_{j,i}}{2},$$

which proves the first part of proposition 8.

If $\alpha = 2$, lemma 2 gives

$$U^{dec}_i - U^c_i = \left|x^{dec}_i - u\right|^2 + \sum_{j\neq i}\beta_{i,j}\left(\left|x^{dec}_i - u\right|^2 - \left|x^{dec}_j - u\right|^2\right).$$

Summing over all jurisdictions, we have

$$B^{dec} - B^{c} = \sum_{i=1..I} \left(1 + \sum_{j\neq i} \beta_{i,j} - \sum_{j\neq i} \beta_{j,i}\right) \left|x_i^{dec} - u\right|^2,$$

which gives the sufficiency part of proposition 8. Suppose that for some i, $\sum_{j\neq i} \beta_{i,j} - \sum_{j\neq i} \beta_{j,i} < -1$. From lemma 1, there exists a profile of type θ such that $x_i^{dec}(\beta, \theta) = 1$ and $x_j^{dec}(\beta, \theta) = 0$ for all $j \neq i$. From the equation above, the uniform policy $(0, .., 0)$ is Benthamite preferred to $x_i^{dec}(\beta, \theta)$, which establishes the necessary part.

8.4 Proof of proposition 9

Under the notations of proposition 9, we define $P(x) = v(x) - \frac{1}{2} w(x)$. As W is reciprocal, we have

$$\nabla_{x_i} P = \nabla_{x_i} V_i - \frac{1}{2} \sum_j \nabla_{x_i} W_j = \nabla_{x_i} V_i - \nabla_{x_i} W_i = \nabla_{x_i} U_i.$$

Hence P is a potential for the decentralized game and a decentralized equilibrium x^{dec} exists.[41]

Using successively the fact that for all $x^c \in \mathbb{R}^n$, $w(x^c, .., x^c) = 0$, and w is more than quadratic and reciprocal,

$$w(x^{dec}) - w(x^c, .., x^c) \leq \frac{1}{2} \sum_i \sum_j \nabla_{x_j} W_i \left(x^{dec}\right) .x_j^{dec} = \sum_i \nabla_{x_i} W_i \left(x^{dec}\right) .x_i^{dec}, \tag{19}$$

[41] See Monderer and Shapley 1994.

For all $x^c \in \mathbb{R}^n$ and $t \in [0,1]$, let $x(t) = (1-t)(x^c, .., x^c) + t x^{dec}$. Integrating $\frac{\partial [v(x(t))]}{\partial t}$ on $[0,1]$ and using the concavity of v,

$$v(x^{dec}) - v(x^c, .., x^c) \geq \int_0^1 \frac{\partial [v(x(t))]}{\partial t}(t=1)\, dt = \sum_i \nabla_{x_i} V_i \left(x^{dec}\right) . \left(x_i^{dec} - x^c\right). \tag{20}$$

As w is invariant by uniform translation, $\sum_i \nabla_{x_i} w = 0_{\mathbb{R}^n}$. From (19), we have

$$\sum_i \nabla_{x_i} W_i \left(x^{dec}\right) = \frac{1}{2} \sum_i \sum_j \nabla_{x_i} W_j \left(x^{dec}\right) = 0_{\mathbb{R}^n}. \tag{21}$$

Adding (19) and (20),

$$(v-w)\left(x^{dec}\right) - (v-w)(x^c, .., x^c) \geq$$
$$\sum_i \left(\nabla_{x_i} (V_i - W_i)\left(x^{dec}\right)\right) . \left(x_i^{dec} - x^c\right) + \sum_i \nabla_{x_i} W_i \left(x^{dec}\right) . x^c.$$

The first term of the right hand-side is null from the first order condition of the decentralized equilibrium, the second term is null as well from (21).

8.5 Proofs in Section 5

8.5.1 Proof of Proposition 10

Throughout this proof, $0 < \alpha \leq 1$ and we assume w.l.o.g. that $\theta_1 < \theta_2$. In particular, any decentralized equilibrium necessarily satisfies $x_1 \leq x_2$. The comparative statics w.r.t. θ can be derived immediately from the comparative statics w.r.t. to β_S by using lemma 5.

Characterization of the Set of Nash Equilibria Using the notations in (4), the decentralization game between the two jurisdictions admits the following potential function (See Monderer and Shapley 1996):

$$\Pi(x) = -(1-\beta_A)|x_1-\theta_1|^2 - (1+\beta_A)|x_2-\theta_2|^2 - \beta_S\left(1-\beta_A^2\right)|x_1-x_2|^\alpha, \tag{22}$$

which is continuous and clearly has a global maximum in $[\min_i \theta_i; \max_i \theta_i]^I$. Therefore a Nash equilibrium exists.

Suppose x^* is a non uniform Nash equilibrium, P is differentiable at x^* and x^* is a local maximum of P.[42] With a slight abuse of notation, the potential can be rewritten as a function $\Pi(x_1+x_2, x_2-x_1)$ which is strictly concave in x_1+x_2 and

$$\begin{aligned} \Pi^*(x_2-x_1) &= \max_{x_1+x_2} \Pi(x_1+x_2, x_2-x_1) \\ &= -\left(1-\beta_A^2\right)\left[|x_2-x_1-(\theta_2-\theta_1)|^2 + \beta_S|x_2-x_1|^\alpha\right] \end{aligned} \tag{23}$$

From (23), Π^* has at most one local maximum in x_2-x_1 on $]0,+\infty[$.[43] Therefore, there is at most one non uniform equilibrium. Finally, the set of Nash equilibria is either comprised of a unique non uniform Nash equilibrium – case *i*) – a set of uniform policies – case *iii*) – or both of them – case *ii*).

Comparative Statics The non uniform Nash equilibrium exists i.f.f. the following two conditions holds:[44]

[42] As the game admits a potential, x^* is a Nash equilibrium i.f.f. x_1^* maximizes $\Pi(., x_2^*)$ and x_2^* maximizes $\Pi(x_1^*, .)$.

[43] As $\frac{\partial \Pi^*}{\partial x_2-x_1} = -\left(1-\beta_A^2\right)\left[2+\alpha(\alpha-1)\beta_S|x_2-x_1|^{\alpha-2}\right]$, $\frac{\partial \Pi^*}{\partial x_2-x_1} > 0$ for $x_2-x_1 \to 0^+$ and $\frac{\partial \Pi^*}{\partial x_2-x_1} < 0$ for $x_2-x_1 \to +\infty$ so Π^* is convex and then concave. As such, it can have only one positive local maximum.

[44] Conditions 1 and 2 are necessary from what precedes. Reciprocally, let x^* satisfy condition 1. Necessarily, $x_1^* < x_2^*$. Simple calculus yields that $U_1(x_1, x_2^*)$ – and so $P(x_1, x_2^*)$ –

Condition 1 Π *has a non uniform local maximum* x^**.*

Condition 2 $\Pi(x^*) \geq \Pi(x_2^*, x_2^*)$ *and* $\Pi(x^*) \geq \Pi(x_1^*, x_1^*)$.

We shall show that holding $(\alpha, \beta_A, \theta)$ constant, either condition is satisfied i.f.f. β_S is small enough. From what precedes, condition 1 is equivalent to Π^* having a positive local maximand in $x_2 - x_1$. This is the case i.f.f. $x_2 - x_1 + \frac{\alpha}{2}\beta_S (x_2 - x_1)^{\alpha-1} = \theta_2 - \theta_1$ has a positive solution which is true i.f.f. β_S is small enough.[45] From the envelope theorem, $P(x^*)$ is decreasing in β_S while $P(x_2^*, x_2^*)$ and $P(x_1^*, x_1^*)$ are constant. So Condition 2 is satisfied i.f.f. β_S is small enough.

Conversely, there exists a uniform equilibrium i.f.f. β_S is large enough since the gains from any deviation from a uniform policy are decreasing in β_S. Therefore, as β_S increases, we can only switch from case $i)$ to $ii)$ or from $ii)$ to $iii)$. The inequalities are weak for case $ii)$ because the Nash equilibrium correspondence is upper hemicontinuous.

Moreover, it should be clear that as $\beta_S \to 0$, there is no uniform equilibrium while as $\beta_S \to +\infty$, all equilibria are uniform, i.e. under the notations of proposition 10, $\underline{\beta} > 0$ and $\overline{\beta} < +\infty$. The comparative statics w.r.t. $\theta_2 - \theta_1$ are derived from the comparative statics w.r.t. β_S as in the proof of proposition 6.

---

is increasing in x_1 to the left of x_1^*. To the right of x_1^*, it is decreasing and then increasing up to x_2^* and decreasing to the right of x_2^*. Hence, either x_1^* or x_2^* is a global maximand of $U_1(., x_2^*)$. So if x^* satisfies condition 2, it maximizes P and it is a Nash equilibrium.

[45] The left hand-side of the equation is decreasing and then increasing, and from the envelope theorem, its minimum is increasing in β_S.

Welfare Analysis In case $i)$, proposition 2 is still valid since it uses only the first order condition.

In case $ii)$, let x^* be the non uniform equilibrium and (u^*, u^*) be a uniform equilibrium. Necessarily $x_1^* \leq u^* \leq x_2^*$.[46] The uniform equilibrium is preferred by jurisdiction 1 to (x_1^*, u^*). As $u^* \leq x_2^*$, x^* is less preferred than (x_1^*, u^*) and than (u^*, u^*) by transitivity. A symmetric reasoning for jurisdiction 2 proves that any uniform equilibrium is Pareto preferred to the non uniform equilibrium. Moreover, since $x_1^* < x_2^*$, either $x_1^* < u^*$ or $u^* < x_2^*$, which implies that for at least one jurisdiction, (u^*, u^*) is strictly preferred to x^*. By continuity, there is a uniform policy which is strictly preferred to x^* by both jurisdiction. By continuity w.r.t. β_S, this is true in a neighborhood of $\underline{\beta}$. This proves that in case $i)$, for β_S sufficiently close to $\underline{\beta}$, there is a uniform policy which Pareto dominates the decentralized equilibrium.

In case $ii)$, for all uniform equilibria (u^*, u^*), from what precedes, $x_1^* \leq u^* \leq x_2^*$ so $\theta_1 \leq u^* \leq \theta_2$. In case $iii)$, if there is an equilibrium (u^*, u^*) such that $u^* < \theta_1$ (resp $\theta_2 < u^*$), then (θ_1, θ_1) (resp. (θ_2, θ_2)) is an equilibrium and it Pareto dominates (u^*, u^*). In cases $ii)$ and $iii)$, let (u^*, u^*) be an equilibrium such that $\theta_1 \leq u^* \leq \theta_2$ and let x be any policy. If $x_2 < u^*$, (u^*, u^*) is preferred to x by jurisdiction 2. If $x_1 > u^*$, (u^*, u^*) is preferred to x by jurisdiction 1. If $x_1 \leq u^* \leq x_2$, jurisdiction 1 prefers (x_1, u^*) to x. As it is a Nash equilibrium, jurisdiction 1 prefers (u^*, u^*) to (x_1, u^*), which proves that (u^*, u^*) is Pareto optimal. Finally, we have proven that all uniform equilibria

[46] As x^* and (u^*, u^*) are Nash equilibria, $U_1(x_1^*, x_2^*) \geq U_1(u^*, x_2^*)$ and $U_1(u^*, u^*) \geq U_1(x_1^*, u^*)$. Summing the two inequalities yields $|x_1^* - x_2^*|^\alpha \leq |x_2^* - u^*|^\alpha + |x_1^* - u^*|^\alpha$ which implies $x_1^* \leq u^* \leq x_2^*$.

in case ii) are Pareto optimal and any suboptimal equilibrium in case iii) is Pareto dominated by an optimal one.
In case ii) and iii), if $\beta_A = 0$, $\left(\frac{\theta_1+\theta_2}{2}, \frac{\theta_1+\theta_2}{2}\right)$ is a Nash equilibrium and it is the Benthamite optimal policy.

8.5.2 Proof of Proposition 11

Suppose $I = 2$. Denote $O_{1,2} = \left\{(V_i(x) + W_i(\{1,2\}))_{i=1,2}, x \in X\right\}$ and $U_i^* = \max_{x\in X} V_i(x) + W_i(\{i\})$ and x_i^* a maximizer. Let $O_{1,2}^*$ be the Pareto frontier of $O_{1,2}$ and X^* the corresponding policies. If $(U_1^*, U_2^*) \in O_{1,2}^*$, then as W is increasing, $x_1^* \neq x_2^*$ and (x_1^*, x_2^*) is a strong Nash equilibrium. If $(U_1^*, U_2^*) \notin O_{1,2}^*$, any uniform policy (x, x) with $x \in X^*$ is a strong Nash equilibrium.
Suppose now that $I \geq 3$, V has monotonic differences in θ and W is weakly congruent. We shall prove the existence of a strong equilibrium by using a theorem in Greenberg and Weber 1986 (p. 116, "A Generalization"). Their strong Tiebout equilibrium corresponds in our context to a strong Nash equilibrium.
As we have assumed that W is increasing in the inclusion sense, their condition (i) is satisfied. Hence, we just need to prove that their condition (ii) is satisfied. Let $S, T \subset \{1, .., I\}$ and let $i, k \in T \setminus S$ and $j \in S \setminus T$ with $\theta_i \leq \theta_j \leq \theta_k$. Suppose furthermore that for $l = i, k$

$$V(\theta_l, x) + W_l(T) > V(\theta_l, y) + W_l(S \cup \{l\}), \tag{24}$$

As V has monotonic differences, there exists $l \in \{i, j\}$ such that $V(\theta_j, x) - V(\theta_j, y) \geq V(\theta_l, x) - V(\theta_l, y)$. Using (24), $V(\theta_j, x) - V(\theta_j, y) > W_l(T) -$

$W_l(S \cup \{l\})$. Finally, as W is weakly congruent, $V(\theta_j, x) - V(\theta_j, y) > W_j(T \cup \{j\}) - W_j(S)$, which establishes the condition (ii) of Greenberg and Weber.

8.6 Proof in Section 6

The proof of proposition 12 follows directly from the fact that $x^{dec}(\theta) = \theta$. We ommit the detail for brevity. We now prove the existence and uniqueness of a LMDE.

For all i and all x_{-i}, let $BR_{i,v}(x_{-i})$ and $BR_i(x_{-i})$ be the best response correspondence respectively of voter v in jurisdiction i and of the LM preferences of jurisdiction i. By definition, a LMDE is a fixed point of $(BR_i(x_{-i}))_{1 \leq i \leq I}$. By concavity of individual preferences and the median voter theorem, $BR_i(x_{-i})$ is the median of $(BR_{i,v}(x_{-i}))_{v \in V_i}$. As the latter are continuous in x_{-i} and single-valued, so is BR_i which proves that a LMDE exists. Moreover, simple algebra shows that for all $v \in V_i$, $BR_{i,v}(.)$ is a contraction for the sup norm, and so is $BR_i(.)$, which proves that the LMDE is unique.

References

[1] Alesina, A.. and E. Spolaore, 1997, “On the Number and Size of Nations”, The Quarterly Journal of Economics, Vol. 112, No. 4, 1027-1056

[2] Alonso R., W. Dessein and N. Matouschek, 2007, “When Does Coordination Require Centralization?”, forthcoming, American Economic Review

[3] Anderson J. and E. van Wincoop, 2004, “Trade Costs”, Journal of Economic Literature, No. 3, pp. 691-751

[4] Aumann R., 1959, “Acceptable points in general cooperative n-person games”, Annals of Mathematics Studies, 40:287-294

[5] Berman H., 1965, “The Uniform Law on International Sale of Goods”, Law and Contemporary Problems, Vol. 30, No. 2, pp. 354-369

[6] Besley T. and S. Coate, 1996, “An economic model of representative democracy”, Quarterly Journal of Economics, 112 (1). pp. 85-114

[7] Besley T. and S. Coate, 2003, Journal of Public Economics 87, “Centralized versus Decentralized Provision of Local Public Good: a Political Economy Approach”

[8] Bolton P. and G. Roland, 1997, “The Break-up of Nations: a Political-economy Analysis”, Quarterly Journal of Economics 112: 1057-1090

[9] Crémer J., A. Estache and P. Seabright, 1996, Revue d’Economie Politique, 106, “Decentralizing Public Services: What Can We Learn from the Theory of the Firm ?”

[10] Cremer, J. and T. Palfrey, 1996, “In or Out? Centralization by Majority Vote,” European Economic Review 40: $43-60$

[11] Demange G. ,1994, “Intermediate preferences and stable coalition structures”, Journal of Mathematical Economics 23, 45–58

[12] Epple D. and Nechyba, T. 2004, “Fiscal Decentralization”, in Handbook of Regional and Urban Economics: Volume 4

[13] Ellingsen T. 1998 “Externalities vs internalities: a model of political integration”, Journal of Public Economics, Volume 68, Number 2, 1, pp. 251-268(18)

[14] Farber D., 1997, “Environmental Federalism in a Global Economy”, Virginia Law Review, Volume 83, 1283-1319

[15] Fudenberg, D. and J. Tirole, 1991, “Game Theory”, Cambridge and London: MIT Press

[16] Gordon R., 1983, "An Optimal Taxation Approach to Fiscal Federalism." Quarterly Journal of Economics, Vol. 98, No. 4, Nov. 1983, pp. 567-586

[17] Greenberg, J. and S. Weber, 1986, “Strong Tiebout equilibrium under restricted. preference domain”, Journal of Economic Theory 38, 101-117

[18] Grandmont J.M., 1978, “Intermediate Preferences and the Majority Rule”, Econometrica 46

[19] Guiso L., P. Sapienza and L. Zingales, “Does Culture Affect Economic Outcomes?”, The Journal of Economic Perspectives, Spring 2006, Vol. 20(2), pp. 23-48

[20] Harstad B., 2007, "Harmonization and Side Payments in Political Cooperation", American Economic Review, Vol. 97, No. 3

[21] Konishi, H., M. Le Breton and S. Weber, 1997, “Pure Strategy Nash Equilibrium in a Group Formation Game with Positive Externalities”, Games and Economic Behavior, 21(1-2):161–182

[22] Kox H., A. Lejour and R. Montizaan, 2004, "The Free Movement of Services Within the EU", C.P.B. document 69

[23] Lockwood B., 2002, "Distributive Politics and the Cost of centralization", The Review of Economic Studies, Volume 69, Number 2

[24] Milgrom P. and Roberts, 1990, "Rationnalizability, Learning and Equilibrium in Games with Strategic Complementarities", Econometrica 58(6)

[25] Friedman M. and R. Friedman, 1980, "Free to Choose: A Personal Statement", Harcourt Brace Jovanovich

[26] Loeper A., 2007, "Federal Directives and Local Discretion", mimeo

[27] Loeper A., 2008, "Decentralized Information, Coordination and the Cost of Centralization", mimeo

[28] Monderer D. and L.S. Shapley, 1996, "Potential Games", Games and Economic Behavior, 14, 124-143

[29] Moulin, H. 1989, "Axioms of Cooperative Decision Making", Econometric Society Monographs

[30] Mueller D., 1997, "Federalist Governments and Trumps", University of Pennsylvania Law Review, Volume 83, No. 7, pp. 1419-1432

[31] Oates W. E., 1972, "Fiscal Federalism", Harcourt-Brace, New York

[32] Oates W. E., 1999, "An Essay on Fiscal Federalism", Journal of Economic Literature, Vol. 37, No. 3, pp. 1120-1149

[33] Oates W.E., 2005, "Torward A Second-Generation Theory of Fiscal Federalism", International tax and Public Finance, 12, 349-373

[34] Persson T. and G. Tabellini, 1992, "The Politics of 1992: Fiscal Policy and European Integration", Review of Economic Studies 59

[35] Posner E., 2006, "International Law: A Welfarist Approach", University of Chicago Law Review, Vol. 73, No. 2, pp. 487-543

[36] Qian Y. and B.R. Weingast, 1997, "Federalism as a Commitment to Preserving Market Incentives", Journal of Economic Perspectives, no. 4: 83–92

[37] Rantakari, H. , "Governing Adaptation", mimeo

[38] Revesz, R., 1996, "Federalism and Interstate Environmental Externalities", University of Pennsylvania Law Review, Vol. 144, No. 6, 2341-2416.

[39] Seabright P., 1996, "Accountability and Decentralization in Government: An Incomplete Contract Model", European Economic Review, 40(1)

[40] Strumpf, K.S. anf F. Oberholzer-Gee, 2002 "Endogenous Policy Decentralization: Testing the Central Tenet of Economic Federalism", Journal of Political Economy 110, 1-36

[41] Taussky O., 1949, "A Recurring Theorem on Determinants", The American Mathematical Monthly, Vol. 56, No. 10, pp. 672-676

[42] Tocqueville A., 1835, “De la Democratie en Amérique I”

[43] Vives X., 1990, “Nash equilibrium with strategic complementarities”, Journal of Mathematical. Economics 19:305–321

[44] Wilson J., 1999 “Theories of tax competition”, National Tax Journal 52 (2): 269-304

[45] Zodrow G. and P. Mieszkowski, 1986. "Pigou, Tiebout, property taxation, and the underprovision of local public goods," Journal of Urban Economics, vol. 19(3), pages 356-370, May