THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 23_14

Marco Alderighi
Università della Valle d'Aosta
Università Bocconi
The Rimini Centre for Economic Analysis, Italy

Christophe Feder
Università degli Studi di Torino, Italy

POLITICAL COMPETITION, POWER ALLOCATION AND WELFARE IN UNITARY AND FEDERAL SYSTEMS

The Rimini Centre for Economic Analysis
Legal address: Via Angherà, 22 – Head office: Via Patara, 3 - 47921 Rimini (RN) – Italy
www.rcfea.org - secretary@rcfea.org

Political competition, power allocation and welfare in unitary and federal systems*

Marco Alderighi †

Università della Valle d'Aosta, Italy
Università Bocconi, Italy
Rimini Centre for Economic Analysis, Italy

Christophe Feder

Università degli Studi di Torino, Italy

November 7, 2014

Abstract

The paper studies how political competition among self-interested parties affects welfare and power allocation between government levels. We find that the unitary and the federal systems of government are not welfare-maximizing, leading to a higher and a lower than optimal centralization level, respectively. Second best is achieved under the system which induces larger potential welfare losses in response to similar deviations from the first best: the risk of large losses force parties to stay closer to the optimum in order to keep the likelihood of winning the elections high. The federal system yields to the second-best solution when inter-jurisdictional spillovers are weak; vice versa a unitary system is better when such spillovers are strong.

JEL classification: H11, D72, H41, H70, H77.

Keywords: Unitary and federal systems; self-interested politicians; local public goods; allocation of power; voting; institutional design.

*Christophe Feder gratefully acknowledges the financial support from the European Social Fund Community grant of Regione Autonoma Valle d'Aosta.

†Corresponding author. Università della Valle d'Aosta, Grand Chemin 75/77, Saint Christophe 11020, Italy; and CERTeT, Università Bocconi, Roentgen 1, 20136 Milan, Italy; telephone +39 0165 066 723; fax: +39 0165 066 748. E-mail: m.alderighi@univda.it.

1 Introduction

When talking about institutional design, there are at least two main aspects to be taken into account: firstly, how much power is in the hands of central and local governments; and secondly, the way in which the allocation of power is put into practise (Lijphart, 2012). Most constitutions contain a variety of rules on the allocation of power which limit, but do not exclude, the possibility of modifications (Kessler et al., 2009). Referenda and constitutional, or even ordinary, laws are two such possibilities. While the former is largely analyzed in the literature, the latter is less considered (Lockwood, 2005; Lorz and Willmann, 2013).

This paper focuses on this second possibility, and investigates the allocation of power between central and local governments as a result of a political competition, where, depending on the system of government, central or local parties declare their electoral programme for the allocation of power (i.e. they present a binding programme detailing the set of local public goods that they want to retain under their competence and those which they want to delegate to the other government level). Parties are self-interested, and their objective is win the elections and to obtain decisional power at their own government level.

We consider two cases which refer to opposite systems of government: the unitary system, and the federal system. In the unitary system (say Japan), the central government chooses how much power to transfer to the local governments. In the federal system (say the European Union), the local governments choose the level of autonomy that they want to retain.[1] We further assume that the allocation of power among government levels embeds a trade-off. By decentralizing, policies are closer to voters' needs; by centralizing, there is a better coordination of policies and an internalization of spillovers.

We find that neither the unitary nor the federal system leads to the first-best outcome, the former inducing too much centralization and the latter too little. The departure from the optimal allocation of power and the consequent welfare loss is here caused by the self-interested behaviour of parties.

Moreover, we show that higher welfare levels are reached under the system of government which induces larger potential welfare losses in response to similar deviations (i.e. same size and different sign) from the optimal level of decentralization. That is, the fed-

[1]The ways in which a federal system can be articulated are much more diverse than those of a unitary system, and, in some cases, local governments cannot freely choose how to allocate power, but this decision is variously shared between the central and the local governments (Watts, 1998). In order to simplify the analysis, and to maintain the symmetry with the unitary system case, we exclude these intermediate situations. Our extreme form of 'federation' is often referred to as 'confederation' or 'league'.

eral system dominates the unitary one when the welfare loss caused by a given increase of decentralization (from the optimal level) is larger than the welfare loss caused by an equivalent reduction of decentralization. This result has the following explanation: larger losses force parties to stay close to the optimal level in order not to greatly reduce their likelihood of winning the elections. Hence, with larger potential welfare losses, voters are more prone to change their electoral preferences in response to an electoral programme which deviates from the optimal level, and, consequently, their resulting chance of punishment is increased. This fact, therefore, reduces the incentive of parties to modify their electoral programmes in their interest.

In the second part of the paper, we focus more deeply on the role of spillovers. Assume, for example, that a country is formed by two jurisdictions where a set of local public goods are produced, and jurisdictions are weakly connected. In this case, it is preferable to assign most of the competencies (power) to the local governments, leaving only the strictly necessary ones at central level. Here, opposite deviations from the optimal allocation of power have an asymmetric impact on welfare.

Since most of the power has been already decentralized, a further allocation of competencies from the central to the local governments induces a large reduction of welfare: the only public goods under the control of the central government are those which provide significant returns from coordination. On the contrary, the reallocation of competencies from the local to the central governments provokes a lower reduction of welfare: local governments previously retained also those competencies which only have small benefits with respect to centralization.

Based on the previous conclusion on the behaviour of parties in the presence of asymmetric welfare losses, we therefore find a smaller departure from the optimal allocation of power in the electoral programmes of local parties and, hence, a lower welfare loss in a federal system. Although this conclusion is well-known in the literature, the above explanation differs from the usual ones. In our case, the result is driven by the self-interested behaviour of parties, while the general explanation is based on the fact that federal systems lead to much more decentralization (Janeba and Wilson, 2011; Bloch and Zenginobuz, 2012).

The current paper contributes to the literature on the allocation of power by offering additional results in this research field (Lockwood, 2005). Redoano and Scharf (2004), Lorz and Willmann (2005) and Rota Graziosi (2009) consider the welfare implications of an alternative political procedure (direct vs representative democracy), while the current paper analyses different systems of government (unitary vs federal). Caplan (2001a) focuses only on the federal case, while Panizza (1999) studies the delegation of power through elections in unitary systems. Crémer and Palfrey (1996) and Lockwood (2004) analyse the delegation of power in direct but not in representative democracy. Other papers consider the delegation of power between private citizens and the public sector (Aghion et al., 2004; Hayo and Voigt, 2013) and the relationship between the size of a country and the division of power among different levels of government (Wrede, 2004; Alesina et al., 2005; Lorz and Willmann, 2013), but do not explicitly focus on the distortionary effects of the self-interested behaviour of parties.

The rest of the paper is organized as follows. Section 2 presents the model. In Section 3, we compute the optimal amount of public spending at the two levels of government. Sections 4 discusses the impact of different government systems on the power allocation. Sections 5 and 6 analyse, respectively, the consequences of a government system choice in the cases of welfare asymmetric losses and of the different strength of inter-jurisdictional spillovers. Section 7 focuses on the effects of the self-interested behaviour of local and central parties and Section 8 concludes. All proofs of Propositions are presented in Appendix A. Some extensions of the model can be found in Appendix B.

2 The model

We consider a country formed by two jurisdictions, $j = 1, 2$, where central and local governments concur to define the intensity of a set of local public goods (e.g. schools, hospitals, harbours, etc.) which have a different degree of inter-jurisdictional spillovers. Let $g_{aj} \geq 0$ be the intensity of the public good produced in jurisdiction j, with the index of spillover $a \in [0, 1)$; and $k(a) \in [0, 1]$ be the intensity of inter-jurisdictional spillovers, with $k'(a) \geq 0$. In addition, let $\gamma_{aj} \in (0, \infty)$ be the (jurisdiction-specific) voters' evaluation of public good a. We assume that: $\bar{\gamma} = E(\gamma_{aj}) = 1$; $\sigma^2 = Var(\gamma_{aj})$; and $Cov(\gamma_{aj}\gamma_{a,-j}) = Cov(\gamma_{aj}\gamma_{bj}) = 0$, with $a \neq b$, for any $a, b \in [0, 1)$ with $j = 1, 2$. Thus, σ is a measure of the heterogeneity between jurisdictions.

The decision on the intensity of local public good $a \in [0, 1)$ in jurisdiction j: namely, g_{aj}, could be *centralized* (g^c_{aj}) i.e. taken by the central government; or *decentralized* (g^d_{aj}), i.e. taken by the local governments. The choice of how to allocate power between government levels embeds a trade-off. Decentralization provides policies closer to local people's needs, while centralization leads to a better coordination of policies and a full internalization of spillovers. A reasonable allocation of decisions between the central and the local governments should therefore assign the power over public goods with lower spillovers to the local governments and over those with higher spillovers to the central one (Oates, 1972; Besley and Coate, 2003; Lorz and Willmann, 2013). Let $\hat{a} \in [0, 1)$ be a measure of the degree of *autonomy* of the local governments.[2] In line with previous reasoning, we assume that the decision on public goods indexed by $a \in A \equiv [0, \hat{a})$ is managed by the local governments, and that indexed by $a \in A^c \equiv [\hat{a}, 1)$ by the central government.[3]

The cost of producing g is $C(g)$, with $C'(0) = 0$, $C'(\infty) = \infty$ and $C'(g), C''(g) > 0$ for $g > 0$. It is the same in the two jurisdictions; it is independent from the decisional level; and it is entirely financed by a non-distortionary local tax, τ_j.[4] The budget-balancing condition holds at the local level, and transfers from one jurisdiction to the other are not permitted. Thus, for jurisdiction $j = 1, 2$, we have:

$$\int_0^1 C(g_{aj})\, da = \int_0^{\hat{a}} C\left(g_{aj}^d\right) da + \int_{\hat{a}}^1 C\left(g_{aj}^c\right) da = \tau_j, \tag{1}$$

where g_{aj}^c refers to public good intensity chosen at the central level, and g_{aj}^d refers to that chosen at the local level.

2.1 Voters

In each jurisdiction, there is a continuum of voters with total mass equal to 1. Voters have the same income I, but differ in their preferences about parties. This form of heterogeneity implies that the result of elections is uncertain. The utility of the representative voter, who lives in jurisdiction j, is:

$$U_j = \int_0^1 u_{aj}(g_{aj}, g_{a,-j})\, da + x_j, \tag{2}$$

where $u_{aj}(g_{aj}, g_{a,-j})$ is the utility that the representative voter living in j receives from the consumption of the public good a, when g_{aj} units of a are produced in jurisdiction j, and $g_{a,-j}$ units are produced in jurisdiction $-j$; and x_j is the utility that s/he receives

[2] This variable is not indexed by j, as we assume that local governments have same power, i.e. $\hat{a}_j = \hat{a}_{-j}$.

[3] This assumption will be relaxed in Appendix B, where we show that this partition can be endogenously derived within the model.

[4] We suppose that the production of public goods occurs at the local level. Assuming different cost functions between central and local choices is a way of explaining different performance of the central and the local governments (Lorz and Williams, 2005). In order to better identify other key determinants for the decentralization choice, we make the simplifying assumption that cost functions are identical in the two cases. This choice is also motivated by the fact that it is not easy to identify in which case we observe the larger expenses (Breton and Scott, 1978; Lockwood, 2000).

from the consumption of a bundle of private goods. We include in u_{aj} both jurisdiction idiosyncratic characteristics, γ_{aj}, and spillovers between jurisdictions, $k(a)$, in a standard way (Redoano and Scharf, 2004; Lorz and Willmann, 2013):

$$u_{aj} = \gamma_{aj} \left(g_{aj} + k(a) \, g_{a,-j} \right). \tag{3}$$

We also assume that voters cannot move from their jurisdiction. The voter budget constraint is:

$$x_j + \tau_j = I. \tag{4}$$

Since the mass of voters in jurisdiction j is 1, using (1), (2) and (4), the indirect utility function of the representative voter, as well as the welfare function of jurisdiction j, is:

$$U_j = \int_0^1 \left[u_{aj} \left(g_{aj}, g_{a,-j} \right) - C \left(g_{aj} \right) \right] da + I. \tag{5}$$

2.2 Systems of government and politicians

We consider two different systems of government: the unitary system where the central government determines the allocation of power; and the federal system where the local governments are empowered to set the degree of decentralization (Lockwood, 2005). There are central and local elections. Voters are represented by two parties: a left-wing party and a right-wing party. We use L (l) to refer to the central (local) representatives of the left-wing party, and R (r) to refer to the central (local) representatives of the right-wing party.[5]

Politicians cannot switch from one party to another, and, more importantly, they cannot move inside the same party from one parliament to another. Thus, we assume that there is horizontal competition between the left and the right parties and, vertical competition between the central and the local parties.

Although central or local politicians are involved in day-to-day decisions concerning the provision of public goods, our primary interest is in the strategic decisions concerning the allocation of political power between the central and the local governments. Therefore, as a short cut, we assume that operative decisions on the level of public goods are made by the administrative staff (The details are presented in Section 3). More precisely, central administrative staff decide the intensity of those public goods whose choice is in the hands of the central government, while local administrative staff deal with local choices. The

[5] Henceforth, variables referring to the central election are in capital letters, while those referring to local elections are in small letters.

political programme consists, therefore, only of a binding proposal on the allocation of power, i.e. the decision on the level of local autonomy $\hat{a}$.

We model the electoral competition in a political economy tradition, supposing that politicians are self-interested. Since central and local government are alternatively involved in the decision of the local autonomy level, we limit our analysis to two cases: the utility of central parties in the unitary case and the utility of local parties in the federal case. In both cases, we assume that the utility of parties is positively affected by winning elections (getting 50 per cent plus one vote or more), and by retaining the decisional power at their particular government level. Thus, in the unitary system, the utility function of central party $i = L, R$, when it proposes $\hat{a}_i^U$, is:

$$V_i\left(\hat{a}_i^U\right) = \Pi_i + \Lambda\left(1 - \hat{a}^U\right), \tag{6}$$

where Π_i is the probability that party i wins the election; $1 - \hat{a}^U$ is the power assigned to the central government, with $\hat{a}^U = \hat{a}_i^U$ if the party i wins the election, and $\hat{a}^U = \hat{a}_{-i}^U$ if the elections are won by the other party; Λ is the value that central parties assign to having power, normalized to the value that they assign to winning the competition.

There are different interpretations for (6). A first view is that politicians are self-interested (Brennan and Buchanan, 1980; Caplan, 2001a). A second view, proposed by Dixit and Londregan (1998) and empirically validated by Caplan (2001b), is that the politicians have two interests: a utilitarian one (winning the election) and an ideological one (reaching a certain level of centralization or decentralization). A last interpretation comes from Panizza (1999), who relates Λ to the degree of democracy: high values of Λ refer to dictatorships (politicians maximize their power without considering the result of the election), while low values of Λ mean full democracy (politicians consider only the result of the political elections).

In the federal system, we suppose that the level of autonomy is chosen at the end of a bargaining phase between the winning parties of the two electoral competitions. We assume that the utility function of the local party $i = l, r$ in jurisdiction j, when it proposes $\hat{a}_{ij}^F$, is:

$$v_{ij}\left(\hat{a}_{ij}^F\right) = \pi_{ij} + \lambda \hat{a}^F, \tag{7}$$

where π_{ij} is the likelihood that party i wins the election in jurisdiction j; $\hat{a}^F$ is the power assigned to the local governments which comes from the negotiation between the winning parties of the local competitions;[6] λ is the value that local politicians assign to the power

[6] More precisely, we assume that it coincides with the average value proposed by the two winning parties in their electoral programmes.

of the local parliament, normalized to the value that local politicians assign to winning the competition. Additional interpretations are the same of those for Λ.

Following Persson and Tabellini (2002), we assume that the probability of winning the elections depends on the appeal of the electoral proposal and on a random component (the details will follow in the next sections): Δ in the unitary case, and δ_j or δ_{-j} in the federal case.

2.3 Timing of the game

We analyse two different versions of the same three-stage game. Firstly, we consider the unitary case, where the central government is empowered to decide the degree of local autonomy. The timing of the game is as follows:

- Stage 1 - Nature determines Δ and γ_{aj}, that are unknown to voters and politicians. Central parties L and R simultaneously announce their binding electoral programme, i.e. $\hat{a}_L$ and $\hat{a}_R$, respectively;

- Stage 2 - Voters of both jurisdictions based on the electoral programmes and Δ choose their preferred party. The party which has more support wins the election and the level of autonomy $\hat{a}^U$ is set in accordance with the winning programme. In the case of a tie, the winner is randomly chosen.

- Stage 3 - Nature reveals γ_{aj} at the local but not at the central level. Local public goods are provided in accordance to the indication of the central and local staffs.

Secondly, we consider the federal case, where local governments are empowered to set the degree of local autonomy. The timing of the game is as follows:

- Stage 1 - Nature determines δ_j and γ_{aj}, that are unknown to voters and politicians. Local parties l and r of the two jurisdictions simultaneously and independently announce their preferred level of local autonomy to voters (their electoral programmes), i.e. $(\hat{a}_{lj}, \hat{a}_{rj})$, $j = 1, 2$;

- Stage 2 - The voters of each jurisdiction based on the electoral programme of the contenders, δ_j, and anticipating the electoral outcome in the other jurisdiction choose their preferred party. In each jurisdiction, the party which has more support (votes) wins the election. In the case of a tie, the winner is randomly chosen. The level of autonomy $\hat{a}^F$ is set at the end of a bargaining process between the two winning parties.

- Stage 3 - Nature reveals γ_{aj} at the local but not at the central level. Local public goods are provided in accordance with the choice of the central and local staff.

We solve the models backwards to characterize the perfect Bayesian equilibria (see also fn. 10).

3 Local public goods provision

We now consider the third stage of the game. We suppose that politicians delegate to the administrative staff the decision on the intensity of local public good provision; and that the administrative staff work in the interest of voters and choose local public good intensities that maximize voters' utility (Lorz and Willmann, 2005).[7]

The allocation of power among government levels embeds a trade-off between better information (at local level) and better coordination (at central level). We capture the fact that local staff are better informed on the voters' needs by assuming that the local administrative staff know γ_{aj}, while the central administrative staff only know the expected values of γ_{aj}, i.e. $E\left(\gamma_{aj}\right) = \bar{\gamma} = 1$. On the contrary, we capture the fact that the central administrative staff better coordinate the production by assuming that the central administrative staff keep into account inter-jurisdictional spillovers in determining the intensity of local public goods i.e. $\max_{g^c_{aj}, g^c_{a,-j}} u_{aj}\left(g^c_{aj}, g^c_{a,-j}\right) + u_{a,-j}\left(g^c_{a,-j}, g^c_{a,j}\right) - C\left(g^c_{aj}\right) - C\left(g^c_{a,-j}\right)$, while the local administrative staff only account for local needs, i.e. $\max_{g^d_{aj}} u_{aj}\left(g^d_{aj}, g^d_{a,-j}\right) - C\left(g^d_{aj}\right)$.

It follows that, when the decisions are taken by the local administrative staff, the local public goods intensity is given by:

$$\gamma_{aj} = C'\left(g^d_{aj}\right). \tag{8}$$

When the decisions are, indeed, taken by the central administrative staff, the local public goods intensity is given by:

$$1 + k\left(a\right) = C'\left(g^c_{aj}\right). \tag{9}$$

Since the central administrative staff have no information on γ_{aj}, goods supplied in both jurisdictions with the same spillover index a have the same intensity: $g^c_{aj} = g^c_{a,-j} = g^c_a$ (Lockwood, 2002; Besley and Coate, 2003; Feld et al., 2007).

[7] We do not consider the possibility of different quality of provision or the self-interested behaviour of public directors (Lessmann and Marwardt, 2010; Strand, 2012).

In this section, (8) and (9) are obtained under the assumption that the administrative staff choose the intensity of local public goods. The same results emerge when this choice is directly in the hands of the parties in power provided that the parties explicitly include in their electoral programmes a binding promise on the way in which they want to choose the intensity of local public goods, if their decision is required, i.e. if they win the election and the local public-good intensity choice is under their control. The intuition behind this equivalence result is that, for any $\hat{a}$ chosen by the politicians, the higher the voters' utility, and the higher the probability of winning. So, a self-interested politician, when s/he decides the local public good intensity, has an incentive to choose it at its optimal level, since it maximizes her/his probability of winning the elections.

4 Allocation of power

We now move to consider the second stage of the game. Before entering the last stage of the game, there is no information on γ_{aj}, and, therefore, g^d_{aj} is a random variable. Indeed, the level of g^c_{aj} is already known, since, in the last stage, it is chosen without information at the central level. However, still in this case, the impact of a given g^c_{aj} on consumer utility is uncertain since the consumers' utility is also affected by γ_{aj}. Using (8) and (9), equation (5) can be expressed to account for the dependence of the welfare function on $\hat{a} \in [0, 1)$, the degree of autonomy of the local government:

$$U_j(\hat{a}) = \int_0^{\hat{a}} E\left(u_{aj}\left(g^d_{aj}, g^d_{a,-j}\right) - C\left(g^d_{aj}\right)\right) da + \int_{\hat{a}}^1 E\left(u_{aj}\left(g^c_{aj}, g^c_{a,-j}\right) - C\left(g^c_{aj}\right)\right) da + I, \tag{10}$$

where expectations are taken over γ_{aj}.

Two elements are in play in determining the welfare as a function of $\hat{a}$: the heterogeneity among jurisdictions captured the variability of γ_{aj}, and the strength of inter-jurisdictional spillovers, summarized by $k(a)$. The higher the jurisdiction heterogeneity, the better the decentralization of power, while the higher the inter-jurisdiction effects, the better the performance of the central administrative staff.

Before analysing the effects of government systems on autonomy level $\hat{a}$ and on welfare, we make the following assumptions.

Assumption 1 *$\hat{a}^B$, $\hat{a}^F$ and $\hat{a}^U$ exist and are internal.*

Assumption 2 *$U_j(\hat{a})$ is strictly concave.*

Assumption 1 helps to restrict the analysis, and to focus on the key aspects of the model. Assumption 2 is amply utilized in political economy studies (Crémer and Palfrey, 1996; Cheikbossian, 2008). A sufficient condition for strict concavity can also be obtained when $C''' \geq 0$.

Lemma 1 *If $C''' \geq 0$ then $U_j(\hat{a})$ is strictly concave.*

Lemma 1 shows that Assumption 2 is satisfied by a broad class of functions, e.g. a quadratic cost function. The intuition behind the concavity of $U_j(\hat{a})$ is provided in Figure 1. Let $w^d(a) = E\left(u_{aj}\left(g^d_{aj}, g^d_{a,-j}\right) - C\left(g^d_{aj}\right)\right)$ and $w^c(a) = E\left(u_{aj}\left(g^c_{aj}, g^c_{a,-j}\right) - C\left(g^c_{aj}\right)\right)$ be the indirect utility derived by local public good a, when g^d_{aj} and g^c_{aj} are chosen optimally at, respectively, the local and the central levels.

Note that w^d is increasing and linear in k, while w^c is increasing and convex (i.e. $d^2w^c/dk^2 = 1/C'' > 0$). In addition, when $C''' \geq 0$, the slope of w^c is always higher than that of w^d (see the proof of Proposition 2). To get a simple interpretation of the result, assume that $k(a) = a$. In this case, the consumer utility provided by (10) is simply the sum of the areas below w^d for the interval $[0, \hat{a})$, and below w^c for the interval $[\hat{a}, 1)$. A small change in the allocation of power, $d\hat{a}$, which moves the selected level of autonomy farther away from the optimal one generates an additional welfare loss equal to $\left(w^c(\hat{a}) - w^d(\hat{a})\right) d\hat{a}$.[8] Therefore, the larger the deviation from the optimal autonomy level, the higher the distance between w^c and w^d, and the higher the additional welfare loss. This guarantees the strict concavity of $U_j(\hat{a})$.

4.1 Benchmark

In previous sections, we have assumed that operative decisions are taken by the administrative staff and that politicians are only involved in the choice of the autonomy level. In this subsection, we compute the first-best autonomy level. Since local public good intensity is chosen optimally by central and local staff, i.e. in accordance to equations (8) and (9), the optimal level of autonomy $\hat{a}^B$ (benchmark level) is obtained by maximizing the expected utility of voters: $\max_{\hat{a}} U_j(\hat{a}) + U_{-j}(\hat{a})$.[9] This is equivalent to the case where central or local parties maximize their probability of winning the elections without any interest in retaining power at their government level, i.e. $\lambda = \Lambda = 0$. By symmetry and concavity of U_j, we obtain:

[8] When we consider a more general $k(a)$, the welfare loss caused by $d\hat{a}$ is given by $\left(w^c(\hat{a}) - w^d(\hat{a})\right) \frac{dk}{d\hat{a}} d\hat{a}$.

[9] Since the central administrative staff has no information on γ_{aj}, we are computing the first-best constrained autonomy level.

Proposition 1 *Under Assumptions 1 and 2, the unique benchmark level of local autonomy* $\hat{a}^B$ *is:*

$$U_j'\left(\hat{a}^B\right) = 0. \tag{11}$$

This set-up can easily replicate some of the main results presented in the literature, which are summarized in the following Proposition.

Proposition 2 *Under Assumptions 1 and 2, the optimal level of local autonomy* $\hat{a}^B$ *is:*
a) decreasing in θ*, where* θ *is a positive shifter of* k*, such that* $\partial k(a, \theta) / \partial \theta > 0$*;*
b) increasing in σ*, if* C *is quadratic.*

Proposition 2 corresponds to a general finding in this field of research (Oates, 1972; Feidler and Staal, 2012). Figure 1 illustrates the result. On the one hand, larger inter-jurisdictional spillovers require stronger coordination, and, hence a higher centralization of power (part a). On the other hand, higher heterogeneity among jurisdictions makes local government choices more effective in dealing with the specific characteristics of jurisdictions, and, hence, yields a lower centralization of power (part b). Note that, in this second part of the Proposition, we have assumed a quadratic cost function, in order to have w^d only depending on the first two distribution moments. Analysing the link between inter-jurisdictional heterogeneity and power centralization with a more general specification of the cost function requires a broader definition of spread, which cannot be limited to the first two distribution moments (Rothschild and Stiglitz, 1971).

4.2 Unitary system

In this subsection, we consider the case in which the central government is empowered to decide the level of autonomy. Being self-interested, central parties want to centralize power. However, in order to increase their odds to win the elections, the two parties will be induced to transfer some power to the local governments. Following Persson and Tabellini (2002), the winning probability of party i is given by:

$$\Pi_i = \Pr\left(U_j\left(\hat{a}_i\right) > U_j\left(\hat{a}_{-i}\right) + \Delta\right). \tag{12}$$

We assume that Π_i depends on the appeal of the electoral programme $\hat{a}_i$ with respect to that of the other party $\hat{a}_{-i}$ and on a random component Δ, which is uniformly distributed on the interval $\left[-\frac{1}{\Psi}, \frac{1}{\Psi}\right]$. Simple computations imply that:

$$\Pi_i = \frac{1}{2} + \frac{1}{2}\Psi\left[U_j\left(\hat{a}_i\right) - U_j\left(\hat{a}_{-i}\right)\right]. \tag{13}$$

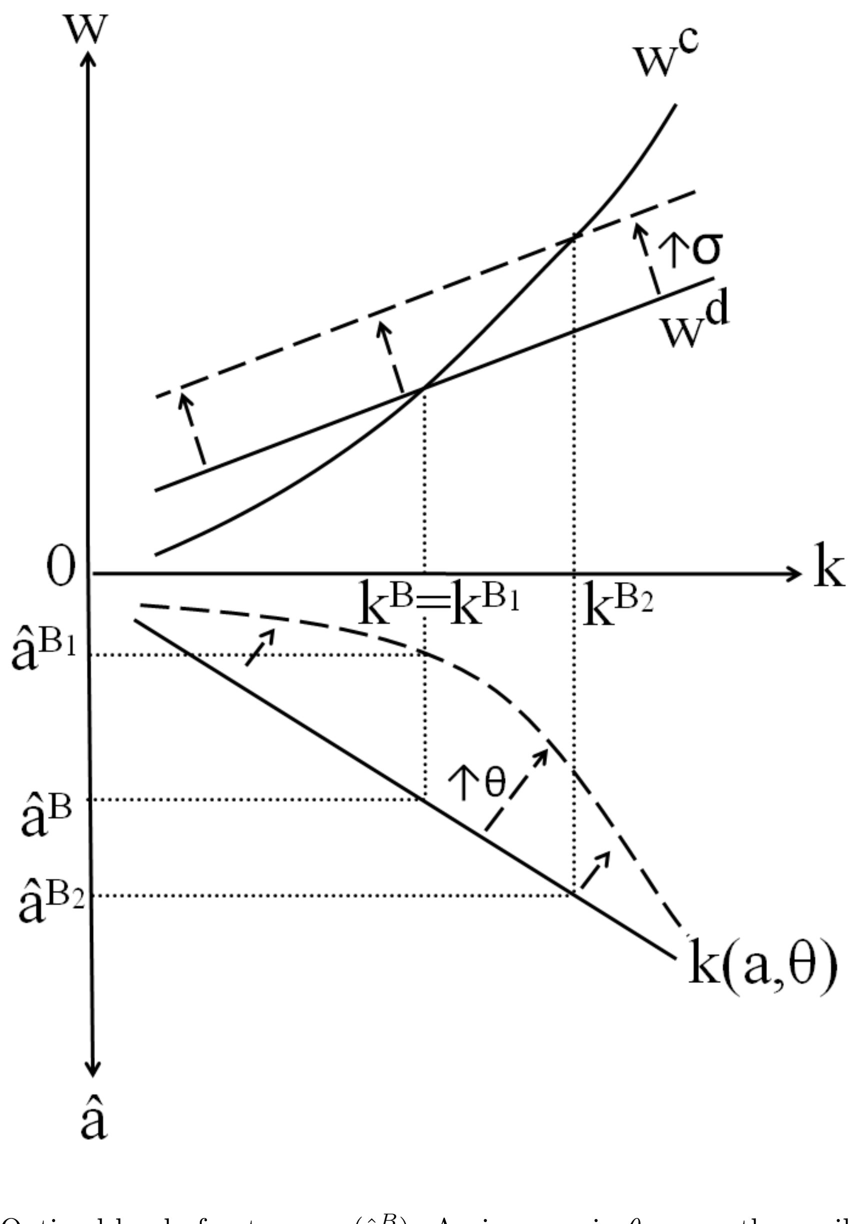

Figure 1: Optimal level of autonomy ($\hat{a}^B$). An increase in θ moves the equilibrium from $\hat{a}^B$ to $\hat{a}^{B_1}$ and an increase in σ from $\hat{a}^B$ to $\hat{a}^{B_2}$.

Using (6), and noting that $\hat{a}^u = \Pi_i \hat{a}_i + (1 - \Pi_i)\, \hat{a}_{-i}$, the utility function of party i is:

$$V_i\left(\hat{a}_i | \hat{a}_{-i}\right) = \left(1 - \Lambda\left(\hat{a}_i - \hat{a}_{-i}\right)\right) \Pi_i + \Lambda\left(1 - \hat{a}_{-i}\right), \tag{14}$$

where Π_i is given by (13). From the maximization of the utility (14) with respect to $\hat{a}_i$, we obtain the response function of party i to the programme of party $-i$. Proceeding in a similar way, we can find the response function of party $-i$, and, from this, the level of autonomy chosen in the unitary case. Proposition 3 formally presents the result:

Proposition 3 *When the system of government is unitary, the level of autonomy emerging from central elections, $\hat{a}^U$, is given by:*

$$U_j'\left(\hat{a}^U\right) = \frac{\Lambda}{\Psi}. \tag{15}$$

Therefore, there is an insufficient level of autonomy, i.e. $\hat{a}^U < \hat{a}^B$.

Note that in $\hat{a}^U$, the welfare function $U_j(\hat{a})$ has a positive slope (if $\Lambda \neq 0$) and, as expected, the level of autonomy is smaller than optimal, i.e. $\hat{a}^U < \hat{a}^B$. The *electoral bias* is the combination of two factors: Λ the value that central politicians assign to having power at the central level, and $\frac{1}{\Psi}$ a measure of the uncertainty of the elections.

Moreover, Proposition 3 replicates the Redoano and Scharf (2004) finding: the level of autonomy is lower under representative democracy than under direct democracy. Indeed, in our set-up, the level of autonomy chosen through a referendum is that which maximizes voters' utility U_j, i.e. $\hat{a}^B$, while that chosen through elections (under their assumption that there are central elections) is $\hat{a}^U$.

4.3 Federal system

We now consider the case where the local governments are empowered to set the level of autonomy. Being self-interested, local politicians want to decentralize power. However, in order to increase their odds of winning the political competition, local parties will be induced to transfer some power to the central government.

To proceed, we need to clarify the link between the electoral outcome of local elections and the allocation of power between the central and the local governments. Let $\left(\hat{a}_j^W, \hat{a}_{-j}^W\right)$ be the vector summarizing the electoral programme of the two winning parties of the local elections, where $\hat{a}_j^W$ is the programme of the winning party in jurisdiction j, and $\hat{a}_{-j}^W$ that of jurisdiction $-j$. Because the two winning parties may have different ideas concerning the transfer of power to the central government, a rule is necessary in order to reach a unique level of autonomy, $\hat{a}^F$. We assume that the winning parties concur to determine the level of autonomy, $\hat{a}^F$, as they have the same bargaining power:

$$\hat{a}^F = \frac{\hat{a}_j^W + \hat{a}_{-j}^W}{2}. \tag{16}$$

The voters of jurisdiction j have rational expectations of the winning electoral programme of the other jurisdiction, $\hat{a}_{-j}^W$.[10]

Recall that π_{ij} is the probability that party i wins in jurisdiction j. Similarly to the unitary case, we assume that π_{ij} depends on the appeal of the electoral programme, and on a random component δ_j, which is independent among jurisdictions, i.e. $Cov\left(\delta_j, \delta_{-j}\right) = 0$, and uniformly distributed on the interval $\left[-\frac{1}{\psi}, \frac{1}{\psi}\right]$, $\psi > 0$, $\forall i = l, r$:

$$\pi_{ij} = \Pr\left(U_j\left(\frac{1}{2}\hat{a}_{ij} + \frac{1}{2}\hat{a}_{-j}^W\right) > U_j\left(\frac{1}{2}\hat{a}_{-ij} + \frac{1}{2}\hat{a}_{-j}^W\right) + \delta_j\right). \tag{17}$$

From equation (17) and the uniform assumption on δ_j, we obtain:

$$\pi_{ij} = \frac{1}{2} + \frac{1}{2}\psi\left[U_j\left(\frac{1}{2}\hat{a}_{ij} + \frac{1}{2}\hat{a}_{-j}^W\right) - U_j\left(\frac{1}{2}\hat{a}_{-ij} + \frac{1}{2}\hat{a}_{-j}^W\right)\right]. \tag{18}$$

By noting that voters in j perfectly anticipate that the winning programme in the other jurisdiction, and that therefore $\hat{a}^F = \frac{1}{2}\left(\pi_{ij}\hat{a}_i + \left(1 - \pi_{ij}\right)\hat{a}_{-i}\right) + \frac{1}{2}\hat{a}_{-j}^W$, the utility function of party i is:

$$v_{ij}\left(\hat{a}_{ij}|\hat{a}_{-ij}\right) = \left(1 + \frac{1}{2}\lambda\left(\hat{a}_i - \hat{a}_{-i}\right)\right)\pi_{ij} + \frac{1}{2}\lambda\left(\hat{a}_{-j}^W + \hat{a}_{-i}\right), \tag{19}$$

where π_{ij} is given by (18).

Finally, from the maximization of the utility (19) with respect to $\hat{a}_{ij}$, we obtain the response function of party i to the programme of party $-i$ in jurisdiction j. Proceeding in a similar way, we can find the response function of party $-i$ in jurisdiction j, and from this, the level of autonomy chosen in jurisdiction j. After performing similar computations for jurisdiction $-j$, and thanks to (16), we find the level of autonomy chosen in the federal case. Proposition 4 formally presents the result:

[10] Because voters living in j, in order to evaluate the electoral proposals of their parties, must formulate a conjecture about the outcome of the competition in the other jurisdiction, the game is Bayesian. We restrict the analysis to symmetric equilibria, and we assume that both parties in $-j$ will choose the same electoral programme $\hat{a}_{-j}^W$.

Proposition 4 *When the system of government is federal, the level of autonomy chosen by local governments, $\hat{a}^U$, is given by:*

$$U_j'\left(\hat{a}^F\right) = -\frac{\lambda}{\psi}. \tag{20}$$

Therefore, there is an excessive level of autonomy, i.e. $\hat{a}^F > \hat{a}^B$.

Note that, in $\hat{a}^F$, the welfare function $U_j(\hat{a})$ has a negative slope (if $\lambda \neq 0$) and as expected the level of autonomy is larger than optimal, i.e. $\hat{a}^B < \hat{a}^F$. As before, the political bias is the combination of two factors: λ the value that local politicians assign to having power at the local level, and $\frac{1}{\psi}$ a measure of the uncertainty of the elections.

Before concluding, it is worth mentioning that the results presented in Propositions 1, 3 and 4 remain unchanged if we consider an arbitrary number of jurisdictions (Appendix B.2) and asymmetric bargaining power among jurisdictions (Appendix B.3). Thus, the analysis presented in Sections 5, 6 and 7 holds under these more general assumptions.

5 Asymmetric welfare losses

Propositions 3 and 4 show that both the federal and the unitary systems cannot yield to the first-best solution when politicians are self-interested. We now discuss how asymmetries in the utility of voters may induce them to prefer one system of government to the other.[11]

Before deriving some of the main results of the model, we formally provide a classification of welfare losses in relation to excessive or insufficient autonomy. Let $\Gamma = \{(\varrho, \mathring{\varrho}) : U_j\left(\hat{a}^B - \varrho\right) = U_j\left(\hat{a}^B + \mathring{\varrho}\right), \ \varrho, \mathring{\varrho} > 0\}$.

Definition 1 *Welfare losses are identical in the case of excessive and insufficient autonomy if, for every $(\varrho, \mathring{\varrho}) \in \Gamma$, $\varrho = \mathring{\varrho}$.*

Definition 2 *Welfare losses are larger in the case of excessive (insufficient) autonomy if, for every $(\varrho, \mathring{\varrho}) \in \Gamma$, $\varrho > \mathring{\varrho}$ ($\varrho < \mathring{\varrho}$).*

Definition 1 states that voters are equally damaged by excessive and insufficient autonomy when the welfare function is symmetric about the value $\hat{a}^B$. In other words, the

[11] Note that the asymmetric losses caused by insufficient or excessive autonomy levels naturally emerge in our set-up (see: Figure 1). For example, when $k(a)$ is linear (solid line in the lower orthant), because of concavity of w^c and linearity of w^d, we have stronger losses moving to the right of $\hat{a}^B$ than to the left. When $k(a)$ is sufficiently concave (dashed line), the result is reversed.

welfare loss is the same when the level of autonomy is larger $(\hat{a}^B + \mathring{\varrho})$ or smaller $(\hat{a}^B - \varrho)$ than the optimal one, provided that the deviations from $\hat{a}^B$ have the same size, $\varrho = \mathring{\varrho}$. Definition 2 concerns situations where the welfare function is no more symmetric about the value $\hat{a}^B$. Welfare losses are larger with excessive autonomy if $\varrho > \mathring{\varrho}$, and, therefore: $U_j\left(\hat{a}^B + \varrho\right) < U_j\left(\hat{a}^B - \varrho\right)$. On the contrary, welfare losses are larger in the case of insufficient autonomy if $\varrho < \mathring{\varrho}$, and, therefore: $U_j\left(\hat{a}^B + \varrho\right) > U_j\left(\hat{a}^B - \varrho\right)$. Generally speaking, a possible explanation of the welfare function asymmetry can be found in the preferences of the representative voter. Ideological reasons, such as nationalism or secessionism, may differently affect voters utility in the case of insufficient and excessive autonomy. An alternative explanation of this asymmetry is presented in Section 6, where we directly link welfare function to the intensity of inter-jurisdictional spillovers.

Proposition 5 *If Assumptions 1 and 2 hold, $\lambda/\psi = \Lambda/\Psi$ and welfare losses are larger with excessive (insufficient) autonomy, then a federal (unitary) system is preferable if:*
i) Λ/Ψ is sufficiently close to 0*; or*
ii) for any $(\varrho, \mathring{\varrho}) \in \Gamma$ there exists $y(\varrho) = \mathring{\varrho}$ such that either $y\prime(\varrho) > 1$ or $y\prime(\varrho) < 1$.

Proposition 5 says that if there are larger losses with excessive autonomy then a federal system is more efficient than a unitary one, and vice versa in the case of insufficient autonomy. Condition *i*) requires that $\hat{a}^U$ and $\hat{a}^F$ are in the neighbourhood of a^B. That is, the distortions arising from the election are sufficiently small. This situation is more likely to occur when parties are weakly self-interested, the electoral outcome is sufficiently certain, and representative voters' preferences are strongly polarized towards a given level of autonomy. Condition *ii*) does not limit the domain of $\hat{a}^U$ and $\hat{a}^F$, i.e. the result holds for any Λ/Ψ, and imposes a slightly more stringent restriction on the shape of U_j, by requiring that the asymmetry in U_j derives from an axis-symmetric contraction or expansion around $\hat{a}^B$.[12] This assumption is consistent with the fact that voters can assign larger welfare losses to deviations in one direction to the other. As mentioned above, this can be generally explained by ideological reasons or formally derived from differences in the inter-jurisdictional spillovers.

The intuition of this key result is as follows. In choosing their electoral programmes, parties face a trade-off between having a greater political power and a greater likelihood of winning the elections. When a departure from the optimal autonomy level induces a large potential welfare loss, the parties stay close to the optimal autonomy level in order not to greatly reduce their probability of winning the elections, while if the same departure only induces a small potential welfare loss, they prefer to focus on increasing the political

[12]For example, Condition *ii*) is satisfied, when $\mathring{\varrho} = A\varrho^B$ with $A, B \geq 1$, or with $0 \leq A, B \leq 1$.

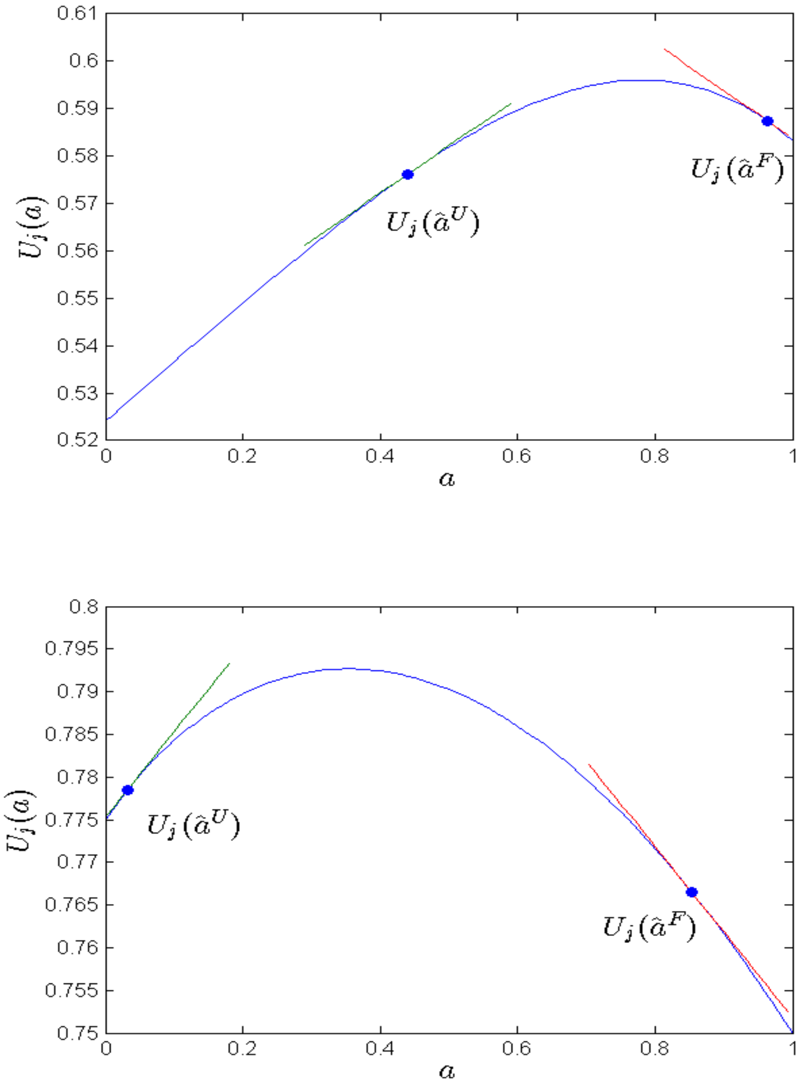

Figure 2: Equilibrium welfare under unitary and federal government systems with asymmetric welfare functions. $C(g) = g^2$ and $k(a) = \kappa a^{\frac{1-\theta}{\theta}}$. Parameters $\kappa = 1$; $\sigma^2 = \frac{1}{2}$; $\frac{\lambda}{\psi} = \frac{1}{10}$; $\theta = \frac{5}{12}$ (upper panel), $\frac{9}{12}$ (lower panel).

power. In other words, when the self-interested behaviour of parties (i.e. deviations from the optimal level of autonomy) is much more tolerated by voters, we expect to observe a larger deviation from the optimal allocation, and a higher welfare loss. In order to minimize the inefficiency, it is better to choose the government system that is potentially more damaging, henceforth inducing a smaller departure from the optimum.

Figure 2 illustrates the result. The upper panel describes the situation where there are larger welfare losses with excessive autonomy, and the lower panel the situation with insufficient autonomy. Using Propositions 3 and 4 and the assumption $\lambda/\psi = \Lambda/\Psi$, from Proposition 5, we obtain that the autonomy levels in the two government systems satisfy: $\left|U_j'\left(\hat{a}^U\right)\right| = \left|U_j'\left(\hat{a}^F\right)\right|$. When departing from the optimal level induces a fast decrease in welfare, the slope of the welfare function soon becomes steep and the equilibrium solution is reached with a smaller bias and a lower welfare loss.

The recent literature has amply documented that the choice of a government system is driven by many factors such as initial conditions (Galeotti, 2000; Rao, 2006) and historical accidents (Spahn, 2006; Breton, 2006), but it is rarely motivated by economic reasons (Ahmad et al., 2008) and by a forward-looking strategy of parties (Crémer and Palfrey, 2000; Spahn, 2006). These arguments and the fact that government systems usually exhibit persistency and path dependence (David, 1985) may provide an explanation for countries' welfare differences: those countries which have chosen the wrong government system in the past, or although initially selecting the right one, have faced structural changes such that the past choice is no longer the proper one, have, all other things equal, lower economic welfare than those which have implemented the proper institutional choice.

6 Inter-jurisdictional spillovers

In this section, we focus on the role of inter-jurisdictional spillovers in determining the preferred government system. In order to provide more striking results we need to specify the cost function, and how inter-jurisdictional spillovers vary in accordance with a. Let $C(g) = g^2$ and $k(a, \theta) = \kappa a^{\frac{1-\theta}{\theta}}$, with $\kappa \in (0, 1]$, $\theta \in (0, 1]$. A high value of θ means that most of the public goods generate strong inter-jurisdictional spillover (k is concave), and a low value of θ means that most of the public goods produce weak inter-jurisdictional spillovers (see also: fn. 11). Let $\frac{\Lambda}{\Psi} = \frac{\lambda}{\psi} > 0$ and $\Lambda, \Psi, \lambda, \psi > 0$. To guarantee that there exists an internal solution, we further assume that:

$$4\frac{\lambda}{\psi} < \sigma^2 < \kappa^2 - 4\frac{\lambda}{\psi}. \tag{21}$$

Espression (21) says that, in order to have an internal solution, we need that maximal inter-jurisdictional spillovers (κ) are sufficiently large, and that the self-interested behaviour of politicians (λ) and the uncertainty about elections ($\frac{1}{\psi}$) are sufficiently small. Using (8) and (9), we find that the optimal public goods provision at the local level is $g_{aj}^{d} = \gamma_{aj}/2$, and at the central level is $g_{a}^{c} = (1 + k(a, \theta))/2$. Replacing previous expressions in (10), taking the expectations and knowing that $E\left(\gamma_{aj}^{2}\right) = \left(1 + \sigma^{2}\right)$ and $E\left(\gamma_{aj}\gamma_{a,-j}\right) = 1$, we obtain:

$$U_{j}(\hat{a}) = \int_{0}^{\hat{a}} \left(\frac{1}{4}\left(1 + \sigma^{2}\right) + \frac{1}{2}k(a, \theta)\right) da + \frac{1}{4}\int_{\hat{a}}^{1} (1 + k(a, \theta))^{2}\, da + I. \tag{22}$$

From Lemma 1, we know that the expected utility function is concave in $\hat{a}$. Taking the first derivative and equating it to ξ (explanation will follow), we get: $\frac{1}{4}\left(1 + \sigma^{2}\right) + \frac{1}{2}k(\hat{a}, \theta) - \frac{1}{4}(1 + k(\hat{a}, \theta))^{2} = \xi$, where $\xi = 0$ in the first-best solution (benchmark solution), $\xi = \frac{\lambda}{\psi}$ in the unitary case, and $\xi = -\frac{\lambda}{\psi}$ in the federal case. Solving the second-order equation in $k(a, \theta)$, taking the positive solution and using $k(a, \theta) = \kappa a^{\frac{1-\theta}{\theta}}$, we get: $\hat{a} = \left(\frac{1}{\kappa}\sqrt{\sigma^{2} + 4\xi}\right)^{\frac{1}{\theta}}$. Therefore:

$$\hat{a} = \begin{cases} \hat{a}^{B} = \left(\frac{\sigma}{\kappa}\right)^{\frac{\theta}{1-\theta}} & \text{in the benchmark case;} \\ \hat{a}^{U} = \left(\frac{1}{\kappa}\sqrt{\sigma^{2} - 4\frac{\lambda}{\psi}}\right)^{\frac{\theta}{1-\theta}} & \text{in the unitary system;} \\ \hat{a}^{F} = \left(\frac{1}{\kappa}\sqrt{\sigma^{2} + 4\frac{\lambda}{\psi}}\right)^{\frac{\theta}{1-\theta}} & \text{in the federal system.} \end{cases} \tag{23}$$

Thanks to (21), the three solutions are internal and satisfy: $0 < \hat{a}^{U} < \hat{a}^{B} < \hat{a}^{F} < 1$. Using a Taylor expansion, we can approximate $U_{j}\left(\hat{a}^{U}\right)$ and $U_{j}\left(\hat{a}^{F}\right)$ around $U_{j}\left(\hat{a}^{B}\right)$:

$$U_{j}\left(\hat{a}^{U}\right) \simeq U_{j}\left(\hat{a}^{B}\right) + \frac{U_{j}''\left(\hat{a}^{B}\right)}{2}\left(\hat{a}^{U} - \hat{a}^{B}\right)^{2}, \tag{24}$$

$$U_{j}\left(\hat{a}^{F}\right) \simeq U_{j}\left(\hat{a}^{B}\right) + \frac{U_{j}''\left(\hat{a}^{B}\right)}{2}\left(\hat{a}^{F} - \hat{a}^{B}\right)^{2}. \tag{25}$$

Therefore, the unitary system is preferred when $\left|\hat{a}^{F} - \hat{a}^{B}\right| < \left|\hat{a}^{U} - \hat{a}^{B}\right|$, and the federal system when $\left|\hat{a}^{F} - \hat{a}^{B}\right| > \left|\hat{a}^{U} - \hat{a}^{B}\right|$. Note that, if we reuse the Taylor expansion, we can approximate $\hat{a}^{U}$ and $\hat{a}^{F}$ around $\hat{a}^{B}$, obtaining the following result:

$$\left|\hat{a}^{F} - \hat{a}^{B}\right| - \left|\hat{a}^{B} - \hat{a}^{U}\right| \simeq \frac{1}{4\sigma^{4}} \frac{\theta}{(1-\theta)^{2}} (3\theta - 2) \left(\frac{1}{\kappa}\sigma\right)^{\frac{1}{\theta}} \left(4\frac{\lambda}{\psi}\right)^{2}.$$

Thus, the benchmark case is closer to the unitary case when there are strong inter-jurisditional spillovers $\theta > \frac{2}{3}$, and to the federal case when they are weak $\theta < \frac{2}{3}$.

We interpret this result in the following way. Assume, for example, that there are weak inter-jurisdictional spillovers for the major share of the public goods. In this case, it is preferable to assign most of the competencies (power) to the local governments, leaving only the strictly necessary ones at central level. Since most of the power has been already decentralized, a further allocation of competencies from the central government to the local governments would induce a large reduction of welfare: the only public goods under the control of central government are those which provide significant returns from coordination. On the contrary, the re-allocation of competencies from the local governments to the central government provokes a lower reduction of welfare: previously the local governments retained some competencies, which have only small benefits with respect to centralization, and that therefore can be easily transferred to the local government without a significant welfare loss.

We therefore find a smaller departure from the optimal allocation of power in the electoral programmes of local parties in a federal system than in those of central parties in a unitary one.

The following Proposition summarizes and extends previous arguments when $\hat{a}^F$ and $\hat{a}^U$ are not necessarily in the neighbourhood of $\hat{a}^B$.

Proposition 6 *Under (21), when $\theta < \frac{2}{3}$ ($\theta > \frac{2}{3}$), the federal (unitary) system is second-best solution; when $\theta = \frac{2}{3}$, the unitary and the federal systems provide the same level of welfare.*

Proposition 6 offers a new explanation of the general preference for a federal (unitary) system when jurisdictions are weakly (strongly) interconnected (Oates, 1972; Janeba and Wilson, 2011; Bloch and Zenginobuz, 2012). The usual argument to identify the better government system relies on the combination of two considerations.[13] First, when spillovers are weak, it is better to have a high level of autonomy, because the local government chooses better policies since it takes into account jurisdictional differences. Second, the federal system induces a higher autonomy level than the unitary one (in our framework, Propositions 3 and 4). Summing up these two points, it is usually argued that when there are weak spillovers a federal system should be preferred to a unitary one. In our set-up we reach the same conclusion, but with different arguments. Indeed, the reason why the federal system is preferred by voters when there are weak spillovers is because

[13] We develop the argument in favour of the federal system, but a similar argument can be made in favour of the unitary one.

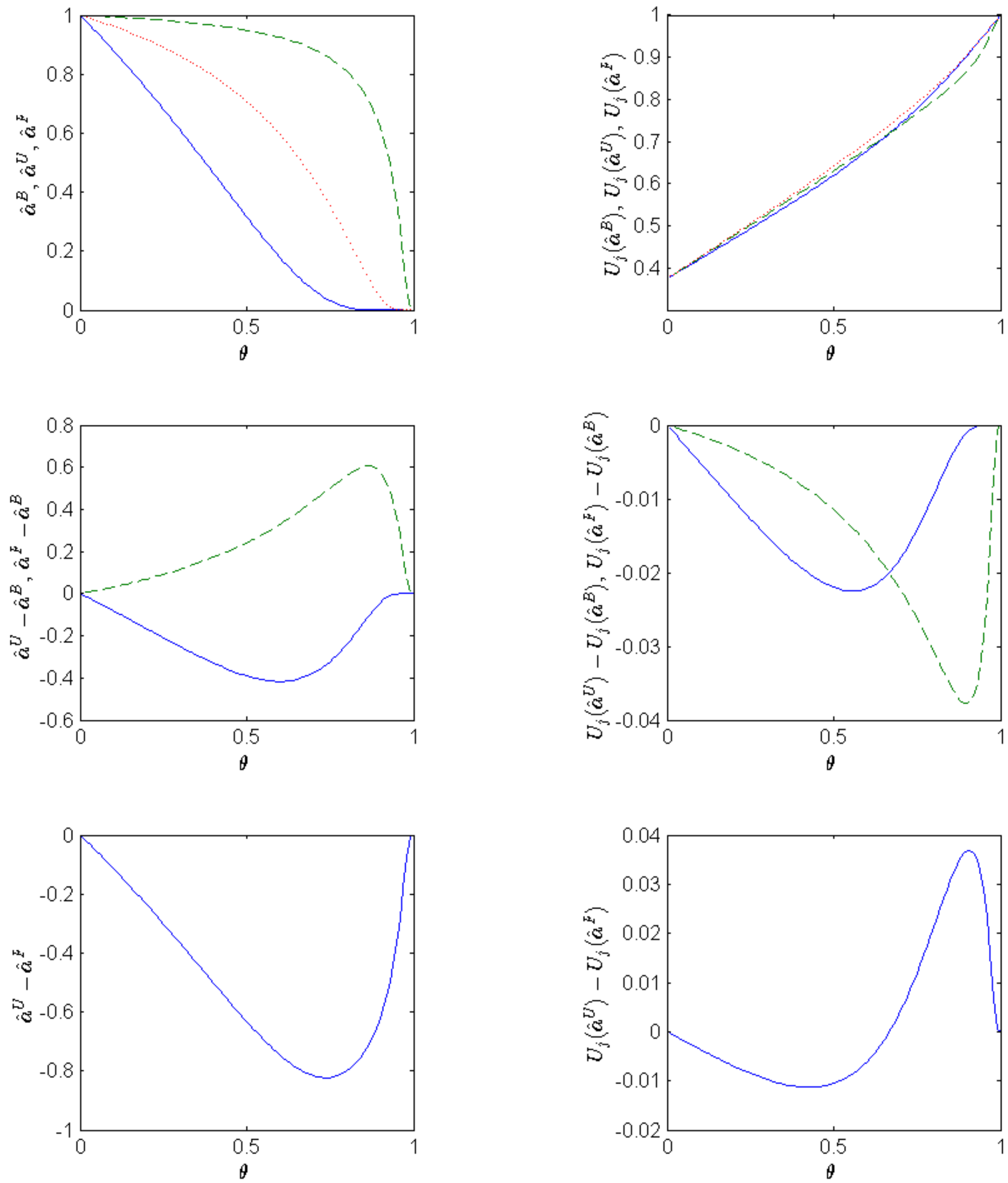

Figure 3: Welfare function, U_j, and level of autonomy, $\hat{a}$, as a function of inter-jurisdictional spillovers, θ. The dotted line refers to the benchmark case, dashed line to the federal case and the continuous line to the unitary case. $C(g) = g^2$ and $k(a, \theta) = \kappa a^{\frac{1-\theta}{\theta}}$. Parameters $\kappa = 1$; $\sigma^2 = \frac{1}{2}$; $\frac{\lambda}{\psi} = \frac{\Lambda}{\Psi} = \frac{1}{10}$.

potential welfare losses are larger (and not smaller) with excessive autonomy (see: Section 5), and this fact limits the ability of local parties to deviate from the optimal level.

We further develop the analysis by providing a graphic representation of the results (Figure 3). The figure shows, in the left panels, the three equilibrium autonomy levels, $\hat{a}^B$, $\hat{a}^F$, and $\hat{a}^U$, and their differences as a function of θ; and, in the right panels, the corresponding welfare levels and differences. Note that there is a link between the left and right panels of the figure: the closer $\hat{a}^U$ and $\hat{a}^F$ to $\hat{a}^B$ the higher the corresponding welfare levels. The dotted line refers to the benchmark case, the dashed line to the federal case, and the continuous line to the unitary case.

The welfare loss due to the departure from the first-best solution is not uniform over θ. Indeed, it decreases near the corners. In other words, when inter-jurisdictional spillovers are very strong/weak, the choice of the government system is irrelevant. This stems from the fact that, in extreme cases, the politicians find themselves forced to choose an autonomy level very close to the optimal one, otherwise they will almost surely lose the elections. Instead, larger welfare losses occur when the θ assumes intermediate values in the unitary case, and medium to high values in the federal case. Figure 3 also shows that the welfare loss with a wrong government system is not symmetric. The worst outcome in the unitary case yields lower welfare losses than (the worst outcome) in the federal one. The maximum welfare loss caused by excessive autonomy is smaller than that caused by insufficient autonomy. Therefore, unitary systems are, on average, less distortionary. If the choice of the government system is taken under a veil of ignorance on the level of θ, a unitary system is the preferred choice.

7 Central and local parties

In the previous sections, we have analysed how welfare asymmetries lead us to identify the preferred government system, provided that the electoral bias, i.e. the self-interested behaviour of central and local parties (Λ and λ) and the uncertainty about central and local elections ($\frac{1}{\Psi}$ and $\frac{1}{\psi}$), is the same on both government levels ($\frac{\Lambda}{\Psi} = \frac{\lambda}{\psi}$). We now continue the analysis considering a different electoral bias between local and central levels. We first study the case where welfare losses are identical under excessive and insufficient autonomy, and we discuss the main links to the recent literature. Then, we analyse a more general case where there is an asymmetric welfare function and a different electoral bias.

Proposition 7 *Assume that welfare losses are identical in the case of excessive and insufficient autonomy, when $\lambda/\psi > \Lambda/\Psi$ ($\lambda/\psi < \Lambda/\Psi$) then a unitary (federal) system is preferable.*

Proposition 7 provides a simple result: the government system which assigns the choice of the autonomy level to the government level with the lower electoral bias is the preferred one. Although the result is largely expected, a more in-depth analysis can be useful because of its interesting practical implications. Usually, two groups of factors can be directly or indirectly related to electoral bias: lobbies, corruption and closeness to the people, on the one hand; and freedom of the press and media pressure, on the other. The former factor is usually associated with an increase of the self-interested behaviour of the parties, while the latter is often linked to faster changes in public opinion, and, therefore, to electoral uncertainty. As far as the first factor is concerned, most of the theoretical works suggest that the local parties exhibit a stronger self-interested behaviour because the greater closeness of people to local governments makes the latter more easily captured by lobbies (Redoano, 2010).[14] Other theoretical works show that local governments and the central government are differently affected by the two factors: lobby pressure is usually greater at the local level and electoral uncertainty at the central level. The former implication has already been discussed above, while the latter is clarified by, for example, Bardhan and Mookherjee (2000), who demonstrate that electoral uncertainty is larger at the central level.

Previous discussion suggests that it is an empirical issue which is the preferred government system (Tabellini, 2000) or the preferred degree of autonomy (Shleifer and Vishny, 1993). However, also from an empirical viewpoint, there is not a unanimous conclusion. For example, using a cross-country analysis, Treisman (2000) shows that federal systems generate more corruption, while Fisman and Gatti (2002) and Lessmann and Markwardt (2010) do not find significant differences in the two government systems.

Propositions 6 and 7 each offer a distinctive rationale for identifying the better government system. The former bases its explanation on the intensity of spillovers, and the latter on the potentially distortionary effects associated with the self-interest behaviour of the parties and electoral uncertainty at the local and the central level (electoral bias).

In Figure 4, we combine these two aspects using the same assumptions of Section 6. The horizontal axis is the intensity of inter-jurisdictional spillovers. The vertical axis displays the electoral bias intensity at the central level over the total, under the assumption

[14] Tanzi (1996) also shows that the self-interested behaviour of local parties is amplified in poor countries. Persson et al. (1997) show that there are two different outcomes. The first is in line with previous findings, although the explanation is different and is based on the common-pool problem. But he second reaches the opposite conclusion, showing that the self-interested behaviour of national parties is stronger because of their higher prestige.

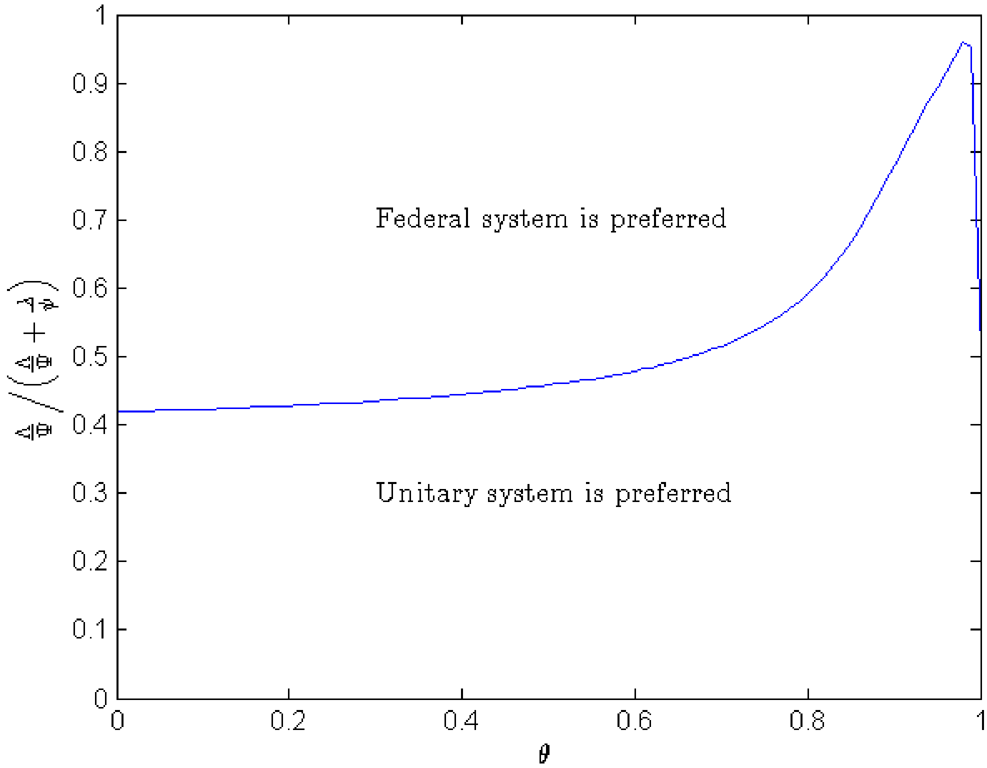

Figure 4: The impact on the preferred government system of the allocation of bias between central and local government and inter-jurisdictional spillovers. Parameters $\kappa = 1$; $\sigma^2 = \frac{1}{2}$; $\frac{\Lambda}{\Psi} + \frac{\lambda}{\psi} = \frac{1}{10}$.

that the total bias is constant; in this example: $\frac{\Lambda}{\Psi} + \frac{\lambda}{\psi} = \frac{1}{10}$. The curve identifies the couples for which voters are indifferent in the two government systems. Above the curve, the federal system is preferred, and below the curve the unitary one. As expected, when $\theta = \frac{2}{3}$, indifference between the two government systems occurs at: $\frac{\Lambda}{\Psi} = \frac{\lambda}{\psi} = \frac{1}{20}$. Visual inspection of this graph, which uses a wide range of values suggests that the shape and position of the indifference curve are weakly affected by the overall electoral bias level. For example, moving from $\frac{1}{10}$ to $\frac{1}{40}$, the intercept increases from 0.435 to 0.484, and the top of the curve reduces from 0.941 to 0.763.

Figure 4 clearly shows that, in order to identify the preferred government system, it is necessary not only to consider the electoral bias but also the intensity of inter-jurisdictional spillovers. That is, for a given electoral bias at the central level, say 55 per cent, the federal system is preferred when there are low-to-medium spillovers, and the unitary one is preferred for high spillovers.

Strong asymmetries emerge when we compare federal and unitary systems. If the level of electoral bias is below 43 per cent (of the total), then a unitary system is always preferred, regardless of the level of inter-jurisdictional spillovers. However, in order to prefer the federal system for any level of inter-jurisdictional spillovers, the electoral bias in the unitary system should be larger than 94 per cent.

8 Conclusions

In this paper, we use a political economy framework to investigate power allocation and welfare in two different government systems: unitary and federal. We show that, all things being equal, countries choosing the wrong government system have lower economic welfare than those which have implemented the proper institutional choice. Another key result of the paper links the preferred government system to inter-jurisdictional spillovers. When they are high, a unitary system is better since it is more effective in preventing parties veering too far from the optimal power allocation, as voters are less inclined to tolerate deviations of central parties than those of local ones. We also show that the federal system may perform very poorly in bad cases, while the unitary system is less inefficient.

The paper can be extended in many directions. For example, it would be interesting to investigate the properties of mixed government systems (Kessler et al., 2009) or the impact of super-majorities or weighted votes on welfare (Lockwood, 2004; Di Giannatale and Passarelli, 2013; Grazzini and Petretto, 2014). Moreover, a dynamic perspective may be applied to analyse how institutional design may affect economic growth (Oates, 1993).

As a final remark, it is worth linking our findings to the recent economic debate. Feld et al. (2007) pointed out that the empirical analysis often reaches opposing conclusions to the theoretical models; and Blume and Voigt (2011) argued that this lack of correspondence often arises from the confusion by using decentralization and federalism as synonyms. Our model shows that welfare implications depend not only on the choice of the government system and/or of how allocate power, but also on jurisdictional linkages, political preferences, and the unpredictability of elections. We leave the in-depth analysis of these factors and the empirical validation of our findings to future research.

References

[1] Aghion, P., Alesina, A. and Trebbi, F., 2004. "Endogenous Political Institutions" The Quarterly Journal of Economics, vol. 119(2), pages 565-611, May.

[2] Ahmad, E., Brosio, G. and Tanzi, V., 2008. "Local Service Provision in Selected OECD Countries: Do Decentralized Operations Work Better?" IMF Working Papers, pages 1-35, March.

[3] Alesina, A., Angeloni, I. and Etro, F., 2005. "International Unions" The American Economic Review, vol. 95(3), pages 602-615, June.

[4] Bardhan, P. and Mookherjee, D., 2000. "Capture and Governance at Local and National Levels" American Economic Review, vol. 90, pages 135-139.

[5] Besley, T. and Coate, S., 2003. "Centralized Versus Decentralized Provision Of Local Public Goods: A Political Economy Approach" Journal of Public Economics, vol. 87(12), pages 2611-2637, December.

[6] Billingsley, P., 1995, Probability and Measure, Johan Wiley & Sons, New York.

[7] Bloch, F. and Zenginobuz, U., 2012. "Oates' Decentralization Theorem with Household Mobility" Working Papers hal-00657823, HAL.

[8] Blume, L. and Voigt, S., 2011. "The economic effects of federalism and decentralization: a cross-country assessment" Public Choice, vol. 151, pages 229-254.

[9] Brennan, D. and Buchanan, J., 1980. The Power to Tax: Analytical Foundations of a Fiscal Consitution. Cambridge University Press, Cambridge.

[10] Breton, A., 2006. Modelling vertical competition. In E. Ahmad and G. Brosio (eds.), Handbook of Fiscal Federalism. Chapter 3. Edited by Cheltenham, Edward Elgar.

[11] Breton, A. and Scott, A., 1978. The Economic Constitution of Federal States. University of Toronto Press, Toronto.

[12] Caplan, B., 2001a. "When is two better than one? How federalism mitigates and intensifies imperfect political competition" Journal of Public Economics, vol. 80(1), pages 99-119, April.

[13] Caplan, B., 2001b. "Has Leviathan Been Bound? A Theory of Imperfectly Constrained Government with Evidence from the States" Southern Economic Journal, vol. 67(4), pages 825-847, April.

[14] Cheikbossian, G., 2008. "Rent-seeking, spillovers and the benefits of decentralization" Journal of Urban Economics, vol. 63(1), pages 217-228, January.

[15] Crémer, J. and Palfrey, T.R., 1996. "In or Out? Centralization by Majority Vote" European Economic Review, vol. 40(1), pages 43-60, January.

[16] Crémer, J. and Palfrey, T.R., 2000. "Federal Mandates by Popular Demand" Journal of Political Economy, vol. 108(5), pages 905-927, October.

[17] David, P.A, 1985. "Clio and the Economics of QWERTY" American Economic Review, vol. 75(2), pages 332-337, May.

[18] Di Giannatale, P. and Passarelli, F., 2013. "Voting chances instead of voting weights" Mathematical Social Sciences, vol. 65(3), pages 164-173.

[19] Dixit, A.K. and Londregan, J., 1998. "Ideology, Tactics, and Efficiency in Redistributive Politics" The Quarterly Journal of Economics, vol. 113(2), pages 497-529, May.

[20] Feidler, J. and Staal, K., 2012. "Centralized and decentralized provision of public goods" Economics of Governance, vol. 13(1), pages 73-93, March.

[21] Feld, L.P., Zimmermann, H. and Döring, T, 2007. "Federalism, Decentralization, and Economic Growth" Public Economics and Public Choice, pages 103-133.

[22] Fisman, R. and Gatti, R., 2002. "Decentralization and corruption: evidence across countries" Journal of Public Economics, vol. 83(3), pages 325-345, March.

[23] Galeotti, G., 2000. Founding fathers versus rotten kids: a positive approach to constitutional politics. In G. Galeotti, P. Salmon and R. Wintrobe (eds.), Competition and Structure. The Political Economy of Collective Decisions: Essays in Honor of Albert Breton. Chapter 20. Cambridge and New York: Cambridge University Press.

[24] Grazzini, L. and Petretto, A., 2014. "Federalism with Bicameralism" German Economic Review, pages 1-23, March.

[25] Hayo, B. and Voigt, S., 2013. "Endogenous constitutions: Politics and politicians matter, economic outcomes don't" Journal of Economic Behavior & Organization, vol. 88(C), pages 47-61.

[26] Janeba, E. and Wilson, J.D., 2011. "Optimal fiscal federalism in the presence of tax competition" Journal of Public Economics, vol. 95(11), pages 1302-1311.

[27] Kessler, A., Lüelfesmann, C. and Myers, G.M., 2009. "The Architecture of Federations: Constitutions, Bargaining, and Moral Hazard" CEPR Discussion Papers 7244, C.E.P.R. Discussion Papers.

[28] Lessmann, C. and Markwardt, G., 2010. "One Size Fits All? Decentralization, Corruption, and the Monitoring of Bureaucrats" World Development, vol. 38(4), pages 631-646, April.

[29] Lijphart, A., 2012. Patterns of Democracy. Yale University Press, New Haven and London.

[30] Lockwood, B., 2000. "The Assignment of Powers in Federal and Unitary States" The Warwick Economics Research Paper Series (TWERPS) 569, University of Warwick, Department of Economics.

[31] Lockwood, B., 2002. "Distributive Politics and the Costs of Centralization" Review of Economic Studies, vol. 69(2), pages 313-337, April.

[32] Lockwood, B., 2004. "Decentralization via Federal and Unitary Referenda" Journal of Public Economic Theory, vol. 6(1), pages 79-108, February.

[33] Lockwood, B., 2005. "Fiscal decentralization: a political economy perspective" Working Paper, University of Warwick, Department of Economics.

[34] Lorz, O. and Willmann, G., 2005. "On the endogenous allocation of decision powers in federal structures" Journal of Urban Economics, vol. 57(2), pages 242-257, March.

[35] Lorz, O. and Willmann, G., 2013. "Size versus scope: on the trade-off facing economic unions" International Tax and Public Finance, vol. 20(2), pages 247-267, April.

[36] Musgrave, R.A., 1959. The theory of public finance: a study in public economy. Edited by McGraw-Hill, New York.

[37] Oates, W.E., 1972. Fiscal Federalism. Edited by Harcourt-Brace, New York.

[38] Oates, W.E., 1993. "Fiscal Decentralization and Economic Development" National Tax Journal, vol.46, pages 237-243.

[39] Panizza, U., 1999. "On the determinants of fiscal centralization: Theory and evidence" Journal of Public Economics, vol. 74(1), pages 97-139, October.

[40] Persson, T., Roland, G. and Tabellini, G., 1997. "Separation of Powers and Political Accountability" The Quarterly Journal of Economics, vol. 112(4), pages 1163-1202, November.

[41] Persson, T. and Tabellini, G., 2002. Political Economics: Explaining Economic Policy. Chapter 3. Edited by MIT Press.

[42] Rao, G., 2006. Fiscal federalism in planned economies. In E. Ahmad and G. Brosio (eds.), Handbook of Fiscal Federalism. Chapter 9. Published by Cheltenham, Edward Elgar.

[43] Redoano, M. and Scharf, K.A., 2004. "The political economy of policy centralization: direct versus representative democracy" Journal of Public Economics, vol. 88(3-4), pages 799-817, March.

[44] Redoano, M., 2010. "Does Centralization Affect the Number and Size of Lobbies?" Journal of Public Economic Theory, vol. 12(3), pages 407-435, June.

[45] Rota Graziosi, G., 2009. "On the Strategic Use of Representative Democracy in International Agreements" Journal of Public Economic Theory, vol. 11(2), pages 281-296.

[46] Rothschild, M. and Stiglitz, J.E., 1971. "Increasing Risk II: Its Economic Consequences", Journal of Economic Theory, vol. 3, pages 66-84.

[47] Shleifer, A. and Vishny, R.W., 1993. "Corruption" The Quarterly Journal of Economics, vol. 108(3), pages 599-617, August.

[48] Spahn, P.B., 2006. Contract federalism. In E. Ahmad and G. Brosio (eds.), Handbook of Fiscal Federalism. Chapter 7. Published by Cheltenham, Edward Elgar.

[49] Strand, J., 2012. "Low-level versus high-level equilibrium in public utility services" Journal of Public Economics, vol. 96(1), pages 163-172.

[50] Tabellini, G., 2000. "Constitutional Determinants of Government Spending" CESifo Working Paper Series 265, CESifo Group Munich.

[51] Tanzi, V., 1996. Fiscal federalism and decentralization: A review of some efficiency and macroeconomic aspects. In M. Bruno and B. Pleskovic (eds.), Annual World Bank Conference on Development Economics, pages 295-316. Washington, DC: World Bank.

[52] Treisman, D., 2000. "The causes of corruption: a cross-national study" Journal of Public Economics, vol. 76(3), pages 399-457, June.

[53] Watts, R.L., 1998. "Federalism, federal political systems, and federations" Annual Review of Political Science, vol. 1, pages 117-137, June.

[54] Wrede, M., 2004. "Small States, Large Unitary States and Federations," Public Choice, vol. 119(1-2), pages 219-240, April.

Appendix A - Proofs

Proof of Lemma 1. Substituting (3) in (10) and taking the second-order derivative, we obtain:

$$U_j''(\hat{a}) = E\left(\gamma_{\hat{a}j} g^d_{\hat{a},-j} k'\right) - E\left(\gamma_{\hat{a}j} \frac{\partial g^c_{\hat{a}}}{\partial k} k'(1+k) + \gamma_{\hat{a}j} g^c_{\hat{a}} k' - C'\left(g^c_{\hat{a}}\right) \frac{\partial g^c_{\hat{a}}}{\partial k} k'\right),$$

where expectations are taken over $\gamma_{\hat{a}j}$ and $\gamma_{\hat{a},-j}$. Remembering that $Corr\left(\gamma_{\hat{a}j}, \gamma_{\hat{a},-j}\right) = 0$, using (9) and collecting similar terms, after simple computations, we obtain:

$$U_j''(\hat{a}) = E\left(\gamma_{\hat{a}j}\right)\left(E\left(g^d_{\hat{a},-j}\right) - g^c_{\hat{a}}\right) k' + \left(1 - E\left(\gamma_{\hat{a}j}\right)\right)(1+k) \frac{\partial g^c_{\hat{a}}}{\partial k} k'.$$

Since $E\left(\gamma_{\hat{a}j}\right) = 1$, the last term is null and, noting that $k' > 0$, the sign of $U_j''(\hat{a})$ coincides with the sign of $E\left(g^d_{\hat{a},-j}\right) - g^c_{\hat{a}}$. Thus, the expected utility function is concave if $E\left(g^d_{\hat{a},-j}\right) < g^c_{\hat{a}}$. We show that this certainly holds if $C''' \geq 0$. Using (8) and (9), we obtain: $E\left(C'\left(g^d_{\hat{a},-j}\right)\right) = E\left(\gamma_{\hat{a}j}\right) = 1 < 1 + k = C'\left(g^c_{\hat{a}}\right)$. Since $C'' > 0$ and $C''' \geq 0$, by Jensen inequality, we obtain $C'\left(E\left(g^d_{\hat{a},-j}\right)\right) \leq E\left(C'\left(g^d_{\hat{a},-j}\right)\right) < C'\left(g^c_{\hat{a}}\right)$, and, finally, since $C' > 0$, we have: $E\left(g^d_{\hat{a},-j}\right) < g^c_{\hat{a}}$. ■

Proof of Proposition 1. Since $U_j = U_{-j}$, we focus on the maximization of U_j. Assumption 2 ensures the uniqueness and differentiability of the equation. Thanks to Assumption 1, the solution is internal, and the maximum is attained when the first-order derivative is null. ■

Proof of Proposition 2. Note that w^c and w^d may be written as:

$$w^c(k) = E\left(\max_g \gamma g(1+k) - C(g)\right) \tag{A.1}$$

$$w^d(k) = E\left(\max_g \gamma(g + k\bar{g}) - C(g)\right), \tag{A.2}$$

where $\bar{g} = E\left(g^d_{a,-j}\right)$ is the average level of the good in the other jurisdiction. Let $g^c_*(k) = \arg\max_g \{\gamma g(1+k) - C(g)\}$, and $g^d_*(\gamma) = \arg\max_g \{\gamma g - C(g)\}$. Using previous notation, equations (A.1) and (A.2) may be re-written as: $w^c = E\left(\gamma g^c_*(k)(1+k) - C\left(g^c_*(k)\right)\right)$ and $w^d = E\left(\gamma g^d_*(\gamma) - C\left(g^d_*(\gamma)\right)\right) + k\bar{g}$. Hence: $\frac{\partial w^c}{\partial k} = E\left(\gamma g^c_*\right) = g^c_*(k)$, where the first equality come from (9). Moreover, we have: $\frac{\partial w^d}{\partial k} = \bar{g}$. From the proof of Proposition (1), we know that $g^c_*(k) > \bar{g}$.

Part a). Using the previous notation, the equilibrium can be written in terms of k.

Let us define the solution of the equilibrium k^B, i.e. $w^c\left(k^B\right) = w^d\left(k^B\right)$. This means that $k^B = k\left(\hat{a}^B, \theta\right)$. The thesis directly follows from the fact that k is increasing in $\hat{a}$, and θ by assumptions.

Part b). From (9) we know that $\partial g_*^c / \partial \gamma = 0$, so that $\partial w^c / \partial \sigma = 0$. Because C is quadratic, from (8), we know that g_*^d is linear and increasing in γ. Let $H(\gamma) = \gamma g_*^d(\gamma) - C\left(g_*^d(\gamma)\right)$. Taking the first and second derivatives, we obtain: $H'(\gamma) = g_*^d(\gamma) > 0$ and $H''(\gamma) = 1/C''(g) > 0$. Hence, $H(\gamma)$ is quadratic, increasing and convex in γ, i.e. the coefficient of the γ^2 term is positive. Therefore, $\partial E(H(\gamma))/\partial\sigma > 0$. In addition, because of linearity of $g_{a,-j}^d$ in γ_{-aj}, $d\bar{g}/d\sigma = 0$, and hence: $\partial w^d / \partial \sigma > 0$.

Summing up the previous results, $w^d(k)$ is increasing in k and σ, while $w^c(k)$ is increasing in k but is not affected by σ. If we assume that the solution is internal and unique, the optimal level of autonomy $\hat{a}^B$ is reached when k^B is such that $w^d\left(k^B, \sigma\right) = w^c\left(k^B, \sigma\right)$. Since a rise in σ implies a positive shift in w^d but no effects in w^c, the level of inter-jurisdictional spillovers k^B associated with the optimal level of autonomy $\hat{a}^B$ increases, implying a larger value of $\hat{a}^B$. ■

Proof of Proposition 3. Replacing (13) in (14) and computing the first-order derivative, we obtain:

$$-\Lambda\left(\frac{1}{2} + \frac{1}{2}\Psi\left(U_j(\hat{a}_i) - U_j(\hat{a}_{-i})\right)\right) + \frac{1}{2}\Psi\frac{dU_j(\hat{a}_i)}{d\hat{a}_i}\left(1 - \Lambda(\hat{a}_i - \hat{a}_{-i})\right) = 0.$$

Imposing symmetry $\hat{a}_i = \hat{a}_{-i}$, from strict concavity of U_j, we obtain the result. ■

Proof of Proposition 4. Replacing (18) in (19) and computing the first-order derivative, we obtain:

$$\frac{1}{2}\lambda\left(\frac{1}{2} + \frac{1}{2}\psi\left[U_j\left(\frac{1}{2}\hat{a}_{ij} + \frac{1}{2}\hat{a}_{-j}^W\right) - U_j\left(\frac{1}{2}\hat{a}_{-i,j} + \frac{1}{2}\hat{a}_{-j}^W\right)\right]\right) +$$

$$+\frac{1}{4}\psi U_j'(\hat{a}_{ij})\left(1 + \frac{1}{2}\lambda(\hat{a}_i - \hat{a}_{-i})\right) = 0.$$

Imposing symmetry $\hat{a}_{ij} = \hat{a}_{i,-j} = \hat{a}_{-j}^W$, from strict concavity of U_j we obtain the result.
■

Proof of Proposition 5.

The proof is presented assuming welfare losses are larger with insufficient autonomy.

Part a). Definition 2 is equivalent to $U_j\left(\hat{a}^B - \varrho\right) > U_j\left(\hat{a}^B + \varrho\right)$, $\forall\varrho$. For λ/ψ small enough, we can approximate $U_j\left(\hat{a}^B - \varrho_1\right) \approx U_j\left(\hat{a}^B\right) + \frac{1}{2}U_{j-}''\left(\hat{a}^B\right)\varrho_1^2$ and $U_j\left(\hat{a}^B + \varrho_2\right) \approx U_j\left(\hat{a}^B\right) + \frac{1}{2}U_{j+}''\left(\hat{a}^B\right)\varrho_2^2$, since $U_j'\left(\hat{a}^B\right) = 0$. Since $U_j\left(\hat{a}^B - \varrho_1\right) > U_j\left(\hat{a}^B + \varrho_1\right)$, therefore, $\frac{1}{2}U_{j-}''\left(\hat{a}^B\right)\varrho_1^2 > \frac{1}{2}U_{j+}''\left(\hat{a}^B\right)\varrho_1^2$. In other terms: $U_{j-}''\left(\hat{a}^B\right) > U_{j+}''\left(\hat{a}^B\right)$. Say $U_{j-}''\left(\hat{a}^B\right) =$

$-\phi_1$ and $U''_{j+}\left(\hat{a}^B\right) = -\phi_2$ with $\phi_2 > \phi_1 > 0$. Assume now that $U'_j\left(\hat{a}^B - \varrho_1\right) = \lambda/\psi = -U'_j\left(\hat{a}^B + \varrho_2\right)$ or $\phi_1\varrho_1 = \phi_2\varrho_2$. Let $\tilde{\varrho}_2$ be the value wihch equalizes the utility in the two cases: $-\frac{1}{2}\phi_1\varrho_1^2 = -\frac{1}{2}\phi_2\tilde{\varrho}_2^2$. Because the utility function is convex, (in absolute value) the slope is decreasing towards $\hat{a}^B$, i.e. for lower values ϱ_2. Now we have to show that the slope in $\tilde{\varrho}_2$ is larger than in ϱ_1, and, therefore, that $\tilde{\varrho}_2 > \varrho_2$. By contradiction, assume that $\tilde{\varrho}_2 \leq \varrho_2$. From the condition on the slope, we have: $\phi_1\varrho_1 \geq \phi_2\tilde{\varrho}_2$. Using: $\phi_1\varrho_1^2 = \phi_2\tilde{\varrho}_2^2$, solving for $\varrho_1 = \sqrt{\frac{\phi_2}{\phi_1}}\tilde{\varrho}_2$ and replacing in the previous inequality, we obtain: $\phi_1\sqrt{\frac{\phi_2}{\phi_1}}\tilde{\varrho}_2 \geq \phi_2\tilde{\varrho}_2$ or: $\phi_1 \geq \phi_2$ which contradicts.

Part b). Note that since $\mathring{\varrho} = y(\varrho)$, and the initial assumption, we have: $y'(\varrho) > 1$. Using Definition 2, we have:

$$U_j\left(\hat{a}^B - \varrho\right) = U_j\left(\hat{a}^B + y(\varrho)\right).$$

Taking the first-order derivative with respect to ϱ, it emerges that:

$$-U'_j\left(\hat{a}^B - \varrho\right) = U'_j\left(\hat{a}^B + y(\varrho)\right) y'(\varrho).$$

Since $y'(\varrho) > 1$, and noting that $U'_j\left(\hat{a}^B - \varrho\right) > 0$ and $U'_j\left(\hat{a}^B + y(\varrho)\right) < 0$, we obtain:

$$U'_j\left(\hat{a}^B - \varrho\right) > -U'_j\left(\hat{a}^B + y(\varrho)\right). \tag{A.3}$$

Let now choose ϱ such that $U'_j\left(\hat{a}^B - \varrho\right) = \Lambda/\Psi = U'_j\left(\hat{a}^U\right)$, and h such that $U'_j\left(\hat{a}^B + h\right) = -\lambda/\psi = U'_j\left(\hat{a}^F\right)$. Since $\Lambda/\Psi = \lambda/\psi$ by assumption, it follows that $U'_j\left(\hat{a}^B - \varrho\right) = U'_j\left(\hat{a}^B + h\right)$. Using this last result, the inequality in (A.3) and Assumption 2, we obtain $h > y(\varrho)$, and, from this, the thesis. ■

Proof of Proposition 6. From (22) and (23), after some computations, we obtain:

$$\begin{aligned}\Delta U_j &= EU_j\left(\hat{a}^U\right) - EU_j\left(\hat{a}^F\right) = \sigma^2\left[\left(\frac{\sigma^2}{\kappa^2} - 4\frac{\lambda}{\psi\kappa^2}\right)^{\frac{\theta}{2(1-\theta)}} - \left(\frac{\sigma^2}{\kappa^2} + 4\frac{\lambda}{\psi\kappa^2}\right)^{\frac{\theta}{2(1-\theta)}}\right] + \\ &\quad -\frac{\theta}{2-\theta}\kappa^2\left[\left(\frac{\sigma^2}{\kappa^2} - 4\frac{\lambda}{\psi\kappa^2}\right)^{\frac{2-\theta}{2(1-\theta)}} - \left(\frac{\sigma^2}{\kappa^2} + 4\frac{\lambda}{\psi\kappa^2}\right)^{\frac{2-\theta}{2(1-\theta)}}\right].\end{aligned}$$

Rearranging the terms, it emerges $\Delta U_j > 0$ when $(X - Y)^\Theta (X + \Theta Y) - (X + Y)^\Theta (X - \Theta Y) > 0$, where $X = \frac{\sigma^2}{\kappa^2}$, $Y = 4\frac{\lambda}{\psi\kappa^2}$, and $\Theta = \frac{2(1-\theta)}{\theta}$, with $0 < Y < X < 1 - Y < 1$ and $\Theta \in [0, \infty)$. Replacing $x = X/Y > 1$, we obtain: $(x - 1)^\Theta (x + \Theta) - (x + 1)^\Theta (x - \Theta) > 0$.

When $x \leq \Theta$, it follows that: $\Delta U_j > 0$. When $x > \Theta$, we have $\Delta U_j > 0$, if:

$$\Upsilon = \left(\frac{x-1}{x+1}\right)^{\Theta} \frac{x+\Theta}{x-\Theta} > 1.$$

Note that $\Delta U_j = 0$ (i.e. $\Upsilon = 1$) if $\Theta = 0$ or $\Theta = 1$ (i.e. $\theta = \frac{2}{3}$). Moreover, since Υ and $\frac{d\Upsilon}{d\Theta}$ are continuous in $\Theta \in (0, x)$; $\frac{d\Upsilon}{d\Theta} = 0$ only in $\Theta = \sqrt{2\left(\ln \frac{x-1}{x+1}\right)^{-1} x + x^2} \in (0,1)$; and $\frac{d\Upsilon(0)}{d\Theta} = \frac{2}{x} + \left(\ln \frac{x-1}{x+1}\right) < 0$, then, for $\Theta \in (0,1)$, $\Delta U_j < 0$; for $\Theta = 1$, $\Delta U_j = 0$; and for $\Theta \in (1, x)$, $\Delta U_j > 0$.

Therefore for $\theta \in \left(0, \frac{2}{3}\right)$ we have $\Delta U_j < 0$; for $\theta = \frac{2}{3}$, $\Delta U_j = 0$; and for $\theta \in \left(\frac{2}{3}, 1\right)$, we have $\Delta U_j > 0$. ■

Proof of Proposition 7. Let $f_1(\varrho) = U_j\left(\hat{a}^B - \varrho\right)$ and $f_2\left(\mathring{\varrho}\right) = U_j\left(\hat{a}^B + \mathring{\varrho}\right)$, with $\varrho, \mathring{\varrho} > 0$. By assumption, we know that if welfare losses are identical in the case of excessive and insufficient autonomy, then $f_1(\varrho) = f_2(\varrho), \forall \varrho$, and, therefore, $f_1'(\varrho) = f_2'(\varrho)$. Noting that $f_1'(\varrho_1) = -\Lambda/\Psi > -\lambda/\psi = f_2'(\varrho_2) = f_1'(\varrho_2)$ and $f_1''(\varrho) < 0$, then $\varrho_1 < \varrho_2$. Since $f_1'(\varrho) < 0$, we have $f_1(\varrho_1) > f_1(\varrho_2) = f_2(\varrho_2)$, i.e. the federal system is preferable. Similar reasoning applies to the other case. ■

Appendix B - Extensions

In this Appendix we generalize some of the previous results in three directions. In subsection B.1, we consider the possibility that parties can offer an electoral programme that is not of the form $A_i \equiv [0, \hat{a}_i)$. In subsection B.2, we show that the model easily accommodates the case of $n \geq 2$ jurisdictions. Finally, in subsection B.3, we show that results do not change if the bargaining power of the two jurisdictions is not equal.

B.1 Endogenous allocation of power

In the baseline model we have assumed that any electoral programme with a degree of autonomy $\hat{a}_i$ splits the set of public goods indexed by $a_i \in \Omega \equiv [0,1)$ into two intervals: $A_i \equiv [0, \hat{a}_i)$ indicating the public goods provided at the local level, and $A_i^c \equiv \Omega \backslash A_i \equiv [\hat{a}_i, 1)$ being those provided at the central level. We now assume that an electoral programme A is a Borelian set in $[0,1)$, i.e. $A = \bigcup_{h=1}^{\infty} [r_h, s_h)$, where $[r_h, s_h)$ are disjoint intervals with $0 \leq r_h < s_h \leq 1$ and $r_h < r_k$, $h < k$ (Billingsley, 1995). Let $\mu\{A\} = \sum_{h=1}^{\infty} \mu\{[r_h, s_h)\}$ be a measure of probability, which defines the autonomy level for any programme $\hat{A}_i$: $\hat{a}_i = \mu\left\{\hat{A}_i\right\}$.

Proposition 8 *The (almost everywhere) optimal strategy for the political party i is to choose a programme of the form $A_i \equiv [0, \hat{a}_i)$.*

Proof We start by showing that, for any level of autonomy $\hat{a}_i$, the probability of winning the elections, Π_i, is maximized when $\hat{A}_i = [0, \hat{a}_i)$. Let $\bar{U}_{-i}$ be the welfare coming from a generic proposal $\tilde{A}_{-i} \in \mathcal{A}$. Then, from (13), it follows that, in order to maximize Π_i, party i has to maximize welfare. According to the new formulation of a programme, we can define the welfare function (10) as:

$$\tilde{U}_j (A_i) = \int_{A_i} w^d (a) \, da + \int_{A_i^c} w^c (a) \, da + I,$$

where $w^d (a) = E \left(u_{aj} \left(g^d_{aj}, g^d_{a,-j} \right) - C \left(g^d_{aj} \right) \right)$ and $w^c (a) = E \left(u_{aj} \left(g^c_{aj}, g^c_{a,-j} \right) - C \left(g^c_{aj} \right) \right)$.

Assume by contradiction that there exists an electoral programme $\tilde{A}_i$ with measure $\mu \left\{ \tilde{A}_i \right\} = \hat{a}_i$ and $\mu \left\{ \tilde{A}_i \cap \hat{A}_i^c \right\} > 0$ such that $\tilde{U}_j \left(\tilde{A}_i \right) \geq \tilde{U}_j \left(\hat{A}_i \right)$. Let $\tilde{A}_i = \bigcup_{h=\mathcal{I}(0 \notin \tilde{A}_i)}^{\infty} [r_h, s_h)$, where $\mathcal{I}$ is the indicator function. Let $s_0 = 0$ if $\mathcal{I}(0 \notin \tilde{A}_i) = 1$. Choose a small $\theta > 0$, such that $\theta < \min \{r_1 - s_0, s_1 - r_1\}$, and the modified programme $\ddot{A}_i = \tilde{A}_i \cup [r_1 - \theta, r_1) \setminus [r_1, r_1 + \theta)$. Then, the variation of utility induced by the modified programme $\ddot{A}_i$, $\tilde{U}_j \left(\ddot{A}_i \right) - \tilde{U}_j \left(\tilde{A}_i \right)$, is:

$$\begin{aligned}
&\int_{[r_1, r_1+\theta)} w^c (a) \, da - \int_{[r_1, r_1+\theta)} w^d (a) \, da + \int_{[r_1-\theta, r_1)} w^d (a) \, da - \int_{[r_1-\theta, r_1)} w^c (a) \, da = \\
&= \theta \left[\left(w^c \left(r_1 + \frac{\theta}{2} \right) - w^d \left(r_1 + \frac{\theta}{2} \right) \right) + \left(w^d \left(r_1 - \frac{\theta}{2} \right) - w^c \left(r_1 - \frac{\theta}{2} \right) \right) \right] \\
&= -\theta \left[U_j' \left(r_1 + \frac{\theta}{2} \right) - U_j' \left(r_1 - \frac{\theta}{2} \right) \right] \\
&= -\theta^2 U_j'' \left(r_1 - \theta/2 \right) > 0,
\end{aligned}$$

where the second equality comes from the approximation of each integral using the average value of the integrand function; the third equality comes from the fact that $U_j' (x) = w^d (x) - w^c (x)$; the fourth equality is the definition of the derivative in terms of incremental ratios; and the last inequality comes from the strictly concavity of U_j. Therefore, $\tilde{U}_j \left(\ddot{A}_i \right) > U_j \left(\tilde{A}_i \right)$, which contradicts.

To complete the proof, we show that the equilibrium solution of the electoral competition game cannot admit strategies in which parties voluntary want to lose the elections, i.e. where at least one party wants to reduce their likelihood of winning, and, therefore, it chooses an electoral programme not in the form $\hat{A}_i = [0, \hat{a}_i)$. By contradiction, assume that $\left(\tilde{A}_i, \tilde{A}_{-i} \right)$ is an equilibrium in which at least one party i chooses an electoral programme that for a given level of autonomy $\hat{a}_i$ has a lower probability of winning the election. This is not an equilibrium because, if it prefers its own level of autonomy, it would prefer to increase the probability of winning the election, and if it prefers the level of autonomy of the other party it can choose a programme which has the same level of autonomy of the other party and maximize its probability of winning. Thus, each party would like to choose a programme of the form $\hat{A}_i = [0, \hat{a}_i)$. ■

B.2 n jurisdictions

In the baseline model we have supposed that there are two jurisdictions. We now assume that there are $n \geq 2$ jurisdictions. Equation (3) modifies in:

$$u_{aj} = \gamma_{aj} \left(g_{aj} + \frac{1}{n-1} k(a) \, g_{a,-j} \right), \tag{B.1}$$

with $g_{a,-j} = \sum_{j' \neq j}^{n} g_{aj'}$.

Proposition 9 *Propositions* 3, 4 *and* 5 *continue to hold for* $n \geq 2$.

Proof The unitary system is still described by (13) and therefore Proposition 3 continues to hold.

In the federal system, the winning parties bargaining generalizes as:

$$\hat{a}^F = \frac{\sum_{j=1}^{n} \hat{a}_i^W}{n} = \frac{\hat{a}_j^W + \sum_{-j} \hat{a}_{-j}^W}{n}. \tag{B.2}$$

Hence, (17) becomes:

$$\pi_{ij} = \Pr\left(U_j \left(\frac{1}{n} \hat{a}_{ij} + \frac{1}{n} \sum_{-j} \hat{a}_{-j}^W \right) > U_j \left(\frac{1}{n} \hat{a}_{-ij} + \frac{1}{n} \sum_{-j} \hat{a}_{-j}^W \right) + \delta \right). \tag{B.3}$$

From equation (B.3), we obtain:

$$\pi_{ij} = \frac{1}{2} + \frac{1}{2}\psi \left[U_j \left(\frac{1}{n} \hat{a}_{ij} + \frac{1}{n} \sum_{-j} \hat{a}_{-j}^W \right) - U_j \left(\frac{1}{n} \hat{a}_{-ij} + \frac{1}{n} \sum_{-j} \hat{a}_{-j}^W \right) \right]. \tag{B.4}$$

By noting that voters in j perfectly anticipate that the winning programme in the other local community, and, that, therefore $\hat{a}^F = \frac{1}{n} \left(\pi_{ij} \hat{a}_{ij} + (1 - \pi_{ij}) \, \hat{a}_{-ij} \right) + \frac{1}{n} \sum_{-j} \hat{a}_{-j}^W$, the utility function of party i is:

$$v_{ij} \left(\hat{a}_{ij} | \hat{a}_{-ij} \right) = \left(1 + \frac{1}{2} \lambda \left(\hat{a}_i - \hat{a}_{-i} \right) \right) \pi_{ij} + \frac{1}{2} \lambda \left(\hat{a}_{-j}^W + \hat{a}_{-i} \right), \tag{B.5}$$

where π_{ij} is given by (B.3). Replacing (B.3) in (B.5) and computing the first-order derivative, we obtain:

$$\frac{1}{2}\lambda\left(\frac{1}{2}+\frac{1}{2}\psi\left[U_j\left(\frac{1}{n}\hat{a}_{ij}+\frac{1}{n}\sum_{-j}\hat{a}^W_{-j}\right)-U_j\left(\frac{1}{n}\hat{a}_{-ij}+\frac{1}{n}\sum_{-j}\hat{a}^W_{-j}\right)\right]\right)+$$

$$+\frac{1}{4}\psi U_j'\left(\hat{a}_{ij}\right)\left(1+\frac{1}{2}\lambda\left(\hat{a}_{ij}-\hat{a}_{-ij}\right)\right)=0.$$

Imposing symmetry $\hat{a}_{ij}=\hat{a}_{-ij}=\hat{a}^W_{-j}$, from the concavity of U_j, after some computations, we obtain the Proposition 4. ■

B.3 Asymmetric bargaining power among jurisdictions

In the baseline model we suppose that in a federal system the level of autonomy is chosen at the end of a bargaining phase between the two winning parties, where the parties have the same bargaining power.

Proposition 10 *Proposition* 4 *continues to hold when jurisdictions have asymmetric bargaining power.*

Proof In the federal system, voters living in j perfectly anticipate that the winning programme in jurisdiction $-j$, and that, therefore, $\hat{a}^F=\rho_j\left(\pi_{ij}\hat{a}_{ij}+\left(1-\pi_{ij}\right)\hat{a}_{-ij}\right)+\left(1-\rho_j\right)\hat{a}^W_{-j}$. The utility function of party i is, therefore:

$$v_{ij}\left(\hat{a}_{ij}|\hat{a}_{-ij}\right)=\left(1+\rho_j\lambda\left(\hat{a}_{ij}-\hat{a}_{-ij}\right)\right)\pi_{ij}+\lambda\left(\left(1-\rho_j\right)\hat{a}^W_{-j}+\rho_j\hat{a}_{-ij}\right), \tag{B.6}$$

where π_{ij} is given by (17). Replacing (17) in (B.6) and computing the first-order derivative, we obtain:

$$\rho_j\lambda\left(\frac{1}{2}+\frac{1}{2}\psi\left[U_j\left(\rho_j\hat{a}_{ij}+\left(1-\rho_j\right)\hat{a}^W_{-j}\right)-U_j\left(\rho_j\hat{a}_{-ij}+\left(1-\rho_j\right)\hat{a}^W_{-j}\right)\right]\right)+$$

$$+\frac{1}{2}\psi U_j'\left(\hat{a}_{ij}\right)\rho_j\left(1+\rho_j\lambda\left(\hat{a}_{ij}-\hat{a}_{-ij}\right)\right)=0.$$

Imposing symmetry $\hat{a}_{ij}=\hat{a}_{-ij}=\hat{a}^W_{-j}$, from the concavity of U_j, we obtain Proposition 4. ■