February 8, 2008

On the Strategic Use of Representative Democracy in International Agreements*

Abstract

We consider as endogenous the choice of the delegation rule in an international agreement between two countries. We study three potential types of delegation: strong, weak or no delegation, the latter case corresponding to direct democracy. We show that populations decide to bind themselves by delegating the national policy decision-making to a "powerful conservative representative", in order to improve their bargaining position. These non-cooperative behaviors of countries when they decide on their delegation rule induce negative political externalities between countries, which cancel the gains achieved by the internalization of economic externalities in the case of political integration. We then examine the consequences of ratification by referendum. We conclude that a Pareto improvement of the international agreement would be to incorporate an *ex post* referendum.

Keywords: Delegation; International Agreements; Nash Bargaining Solution; Direct Democracy; Representative Democracy; Ratification; Referendum.

JEL classification: D72; H77.

Grégoire Rota Graziosi
CERDI-CNRS, Université d'Auvergne
65 boulevard François Mitterrand
63000 Clermont Ferrand, France
Email: gregoire.rota_graziosi@u-clermont1.fr

*I would like to thank Hubert Kempf for his precious help, as well as seminar participants at Paris-1-Panthéon-Sorbonne (January 2006), at the European Public Choice Society Conference (April 2006, Turku, Finland) and at the European Economic Association Conference (August 2006, Vienna, Austria). The ACI "Local Public Goods, Politics and Economics" provided valuable financial support.

1 Introduction

The aim of this paper is to analyze the nature of political democratic regimes and international cooperation through negotiated agreements. We determine which democratic regime (direct or representative) populations choose at the equilibrium when an international agreement is considered, what consequences these choices induce for this agreement and how a ratification through national referendum improves its efficiency. To address this issue, we set-up a two-country model where a public good produced in one country generates spillovers and thus legitimates transfers from the other country. We then assume that the cost of providing the public good is not shared between participant countries. An illustrative example of such public good is biodiversity. Only very few sites throughout the world contain the majority of species. These "hotspots" are located in certain less developed countries, while biodiversity conservation is highly valued by rich countries. Finally the Convention on Biological Diversity and Biosafety Protocol requires developed countries to provide financial assistance to poor countries. But, as Barrett (2003) emphasizes, the amount of money remains low.[1]

Here constituencies choose their political regime while taking into consideration the fact that their elected political authority has to bargain with its counterparts over international matters. Our framework assumptions point unambiguously to the efficiency of an international agreement which organizes the provision of the public good and the payement of a transfer. However, we observe that at the equilibrium populations decide to bind themselves by delegating the national policy decision to a "powerful conservative representative",[2] in order to improve their bargaining position. These non cooperative behaviors induce negative political externalities and cancel any gain of cooperation in terms of aggregate welfare. Moreover we show that the international agreement involves some redistribution among countries, which is harmful for the cooperation itself. We then emphasize that a ratification procedure through an *ex post* referendum in each country is Pareto-improving by limiting strategic delegation. Our paper therefore provides a rationalization for the use of referenda to ratify international agreements.

Our issue is formally related to the literature of the functioning of a federation. Oates'

[1] Other examples of this kind of public goods are policies on defense, reductions of some pollution (acid rain, pollution of international rivers), immigration policy or more broadly any international public good produced in a specific site and having no close substitute in the other country.

[2] A "powerful conservative representative" corresponds to a representative who remains in place whatever the outcome of international negotiations and who is less favourable to public spending than the median national voter. This notion will be formalized through the proposed framework in the following sections.

decentralization theorem presented an initial trade-off between the benefits of the centralization of policy-making and the costs of policy uniformity. Alesina and Spolaore (1997) or Bolton and Roland (1997) have renewed this analysis by adopting a positive approach and emphasizing the political process of integration or secession. These authors highlight a second kind of trade-off between the costs of heterogeneity in large populations and the benefits to large countries or unions of countries in providing public goods or in increasing private incomes. These works and some of their developments, such as Gradstein (2004) or Goyal and Staal (2004), only consider direct democracy. Going beyond this limit, Lockwood (2002) and Besley and Coate (2003) develop models of representative democracy where centralization does not systematically involve policy uniformity. Besley and Coate (2003) establish that over-provision of public goods appears as a result of strategic delegation by jurisdictions, while in contrast, Dur and Roelfsema (2005) or Lorz and Willmann (2005) reach the opposite conclusion. Finally, Laussel (2002) shows how entry in an election race entails an abstention effect which reinforces the strategic delegation and yields extremism.

Besides the literature on federalism, several papers have shown the impact of political regime over the cooperative outcome in different contexts. Focusing on International Environmental Agreements, Buchholz, Haupt, and Peters (2005) establish that each individual delegates to a representative who is less "eco-friendly" than (s)he is. For Laussel and Riezman (2005) representative democracy leads to a more aggressive trade policy, i.e. more protectionist than direct democracy. This comparison between direct and representative democracy is further pursued by Redoano and Scharf (2004), who study the process of institutional change itself and determine which political regime is the more favorable to centralization. These authors conclude that representative democracy is better able to ensure centralization and is also Pareto superior.

In this paper, we go beyond these comparisons of political regimes as we endogenize the strategic choice of political regime by constituencies which anticipate cooperation, i.e. bargaining over a set of policy instruments.[3] We establish that direct democray is not sustainable: it is not a Perfect Subgame Nash Equilibrium (PSNE) even if it may be Pareto-superior. The remainder of the paper is divided into five sections. Section 2 presents the set-up of the model. In Section 3 we define two benchmarks for our analysis: the social optimum and the isolationist equilibrium. In Section 4, we consider an international agreement, which involves cooperation among policy

[3] Redoano and Sharf (2004, p. 811) suggest we should "think of an initial constitutional stage" where voters in each region would choose between direct and representative democracy.

decision-makers. Section 5 proposes a ratification procedure which is Pareto-improving. Section 6 concludes by discussing the limits and the impact of our findings.

2 Preliminaries

In this section we develop our model inspired by Gradstein (2004) who presents a highly stylised (and tractable) representation of inter-jurisdictional policy spillovers. We consider two countries (1 and 2), which might cooperate through an international agreement. There is no cross-border mobility. We assume that country 1 with a population normalized to 1 provides a public good, in quantity g, which generates externalities for the other country, namely country 2, of size d ($d \lesseqgtr 1$). Country 2 pay a transfer, denoted by T, to country 1 in order to increase the production of the public good. This asymmetry between the two countries which contrasts with most of articles on this issue, enables us to distinguish the countries beyond their respective size, income distribution or distribution of the individual preferences for the public good.

The inhabitants of country 1 are assumed to have preferences that can be described by the following utility function:

$$U\left(g,T;a_i\right)=a_i g-c\left(g\right)+dT, \tag{1}$$

where a_i is the appreciation of the public good for inhabitant i in country 1, and $c\left(g\right)$ is the cost of producing the public good in quantity g. For tractability, we consider a quadratic form of this cost: $c\left(g\right)=\frac{g^2}{2}$. In country 2, we assume:

$$V\left(g,T;b_j\right)=b_j g-T, \tag{2}$$

where b_j is the the appreciation of the public good for inhabitant j in country 2. The paramaters a_i and b_j formalize the heterogenity of the countries' populations. By assumption, they are distributed over $[\underline{a},\overline{a}]$ and $\left[\underline{b},\overline{b}\right]$ with respect to the density functions $h_1\left(.\right)$ and $h_2\left(.\right)$. Let A and B denote the mean values of a_i and b_j, whereas A_m and B_m will denote their median values. We pose the following conditions on the preference distributions:

Condition 1 *:* $A_m-dB_m\geqslant\underline{a}$.

Condition 2 *:* $A_m<A+dB$.

As we shall see, the first condition ensures that the application of the Median Voter Theorem (MVT) always involves an interior solution for the choice of the representatives in the set $[\underline{a}, \overline{a}]$. The second condition induces an under-provision of the public good at the decentralized equilibrium. It is respected if we link these preference distributions to the national income distributions as in Bolton and Roland (1997) ($A_m < A$).

3 Benchmarks

Before studying the political integration we consider two benchmarks: a normative one which consists in the maximization of a social welfare function and an isolationist situation which corresponds to the non-cooperative equilibrium.

3.1 Social optimum

In order to contrast the positive predictions against a normative benchmark, we consider the Benthamite social welfare function, defined by:

$$W(g) = \int_{\underline{a}}^{\overline{a}} U(g, T; a_i) h_1(a_i) da_i + d \int_{\underline{b}}^{\overline{b}} V(g, T; b_i) h_2(b_i) db_i = g\left(A + dB - \frac{g}{2}\right)$$

The maximization of this welfare function with respect to g gives the following optimal levels of public good and welfare, respectively denoted by g^{opt} and W^{opt}:

$$g^{opt} = A + dB, \quad W^{opt} \equiv W\left(g^{opt}\right) = \frac{1}{2}(A + dB)^2. \tag{3}$$

At the Pareto optimum, the level of transfer is undetermined, since by definition of the welfare function compensatory transfers are allowed between countries.

3.2 Isolationism

As a second benchmark, we consider the situation when the countries do not cooperate. We describe this situation, called isolationism, by the following three-stage game, denoted by Γ:

- *Stage 1, political regime:* the population of each country simultaneously chooses the kind of political regime it wants: direct or representative democracy. By assumption every political regime has the same cost.

- *Stage 2, national representation:* if a country opts for representative democracy, then a vote fixes the identity of the representative. Otherwise (direct democracy) this stage vanishes and national policies are determined directly through referendum.

- *Stage 3, policy stage:* in each country a national representative or a vote fixes policy non cooperatively.

Our approach is original in considering as endogenous the constitutional design, or more precisely the democratic regime which will determine the identity and the "power" of the national representative.

Applying backward induction we determine the national policies before we turn to the delegation issue. Let a_R and b_R be the preference of the respective political representative in country 1 and in country 2. The national policies are given by the following system:[4]

$$\begin{cases} g^{iso}(a_R, b_R) \equiv \underset{g \geqslant 0}{\arg\max} \{U(g, T; a_R)\} \\ T^{iso}(a_R, b_R) \equiv \underset{T \geqslant 0}{\arg\max} \{V(g, T; b_R)\} \end{cases} \Leftrightarrow \begin{cases} g^{iso}(a_R, b_R) = a_R \\ T^{iso}(a_R, b_R) = 0 \end{cases} \tag{4}$$

We now consider the second stage of the game, i.e. the choice of the country's representative. The representative of country 1 is the unique relevant political decision-maker. Each inhabitant of this country 1, characterized by a_i, votes for her preferred representative, denoted by $a_R^{iso}(a_i)$. Following Besley and Coate (1997) we define the set of available strategies for a voter in country 1 by $[\underline{a}, \overline{a}]$. Therefore, we have:

$$a_R^{iso}(a_i) \equiv \underset{a_R \in [\underline{a}, \overline{a}]}{\arg\max} \left\{U\left(g^{iso}(a_R, b_R), T^{iso}(a_R, b_R); a_i\right)\right\} \Leftrightarrow a_R^{iso}(a_i) = a_i.$$

The unidimensionality of the policy space ($[\underline{a}, \overline{a}]$) and the single-peakness of the utility function allows us to apply the MVT. It yields: $a_R^{iso} \equiv a_R^{iso}(A_m) = A_m$ and then $g^{iso} \equiv g^{iso}(A_m, b_R) = A_m$. Under isolationism, we observe that there is no strategic delegation: the representative of country 1 corresponds to the median voter of this country. The first stage of the game is then trivial since both political regimes are strictly equivalent. The individual utilities levels in both

[4] The superscript *iso* stands for isolationist equilibrium values.

countries and the aggregate welfare are respectively given by:

$$\begin{aligned} U^{iso}(a_i) &\equiv U\left(g^{iso}, T^{iso}; a_i\right) = A_m\left(a_i - \frac{A_m}{2}\right), \\ V^{iso}(b_i) &\equiv V\left(g^{iso}, T^{iso}; a_i\right) = A_m b_i, \\ W^{iso} &\equiv U^{iso}(A) + dV^{iso}(B) = A_m\left(A - \frac{A_m}{2} + dB\right). \end{aligned}$$

In the isolationist situation country 2 freerides country 1 since equilibrium transfer is nil. Under Condition (2), we note that decentralization involves an under-provision of the public good due to the presence of inter-jurisdictional spillovers. Political integration will allow countries to internalize some externalities and to bridge part of the gap between the Pareto optimum and decentralized equilibrium. However, we will show in the next sections that the detail of political decision-making is not neutral and affects the degree to which the gap is bridged.

4 International Agreement

In order to address the bargaining on international agreement, we modify the preceding game in two ways. First national policies are negotiated. This change brings us to a second difference from game Γ which is to distinguish two kinds of representative democracy: strong or weak, depending on whether the elected authority remains in place (strong delegation) or is relieved if negotiations fail (weak delegation). The new game, denoted by Γ', is then:

- *Stage 1', political regime:* the population of each country simultaneously chooses the kind of political regime it wants or, in other terms, the delegation rule for international agreement. We consider two types of representative democracy: strong and weak denoted by *sd* and *wd*. We denote by *dd* the case of direct democracy. Each player then has three distinct pure strategies available, denoted by x (resp. y) for country 1 (resp. 2), with $(x, y) \in \{sd, wd, dd\}^2$.
- *Stage 2, national representation:* similar to game Γ.
- *Stage 3', policy stage:* countries determine their national policies cooperatively. More formally, we capture international agreement through a general Nash Bargaining Solution (NBS).[5]

[5] By not restricting our analysis to the symmetric NBS we will consider the effect of the bargaining power on

The case of weak delegation (wd) involves distinguishing the identity of the representative depending on the outcome of the international negotiations. We denote by $a_{R'}$ the preference of the political representative in country 1 if negotiations fail. In the case of strong delegation the representative remains in place irrespective of the success or failure of negotiations, which gives: $a_{R'} = a_R$; in the case of weak delegation, if negotiation fails, a new round of elections produces a new representative, who will then correspond to the median voter,[6] which gives: $a_{R'} = A_m$; finally, under direct democracy the representative is always the median voter since the vote is by assumption sincere, which gives: $a_{R'} = A_m \, (= a_R)$.

We are now able to define the threatpoint of the NBS for the representatives who negotiate (a_R and b_R). In case of disagreement, each country chooses its own policy which yields: $g^{nc} = a_{R'}$ and $T^{nc} = 0$. The utility levels are then given by: $U^{nc}(a_{R'}; a_R) \equiv U(g^{nc}, T^{nc}; a_R) = (a_R - a_{R'}/2)\, a_{R'}$ and $V^{nc}(a_{R'}; b_R) \equiv V(g^{nc}, T^{nc}; b_R) = b_R a_{R'}$. We observe that the threatpoint of the NBS depends on the delegation rule choosen at the first stage of the game. This feature is crucial for our analysis since it will allow country 1 to strengthen its threat of an under-provision of the public good.

4.1 Equilibrium

We now determine the equilibrium. Applying backward induction we focus on stage 3. Let (g^c, T^c) denote the NBS of our problem:[7]

$$(g^c, T^c) \equiv \underset{(g,T)\in\mathbf{R}^2_+}{\arg\max} \left\{ \left[U(g, T; a_R) - U^{nc}(a_{R'}; a_R)\right]^{\alpha} \left[V(g, T; b_R) - V^{nc}(a_{R'}; b_R)\right]^{1-\alpha} \right\}, \qquad (5)$$

where α is the relative bargaining power of country 1. The First Order Conditions (FOCs) yield:[8]

$$\begin{cases} g^c = a_R + d b_R \\ T^c = \frac{a_R - a_{R'} + d b_R}{2d} \left((1-\alpha)(a_{R'} - a_R) + (1+\alpha)\, d b_R\right) \end{cases} \qquad (6)$$

We turn to the second stage of the game: the choice of the national representatives. Let $a^c_R(x, y; a_i)$ ($b^c_R(x, y; b_j)$) denote the preferred representative for an inhabitant a_i (b_j) in coun-

the equilibrium's values of the public good, the transfer and the utility levels.

[6] If negotiations fail, countries behave non-cooperatively and the new election designates the median voter as under isolationism.

[7] Notice that we do not consider the degenerate NBS problem, where $U(g, T; a_R) = U^{nc}(a_{R'}; a_R)$ or $V(g, T; b_R) = V^{nc}(a_{R'}; b_R)$.

[8] The details of the proofs are available upon request.

try 1 (2) when the delegation rule x (y) has been chosen at the first stage in country 1 (2). By definition, under direct democracy, there is no election of the political representative. Thus we only consider strong and weak delegation: $(x, y) \in \{sd, wd\}^2$. Note that $\frac{da_{R'}}{da_R} = 1$ in the case of strong delegation and $\frac{da_{R'}}{da_R} = 0$ for weak delegation. We have:[9]

$$\begin{cases} a_R^c\left(x, y; a_i\right) \equiv \underset{a_R \in [\underline{a}, \overline{a}]}{\arg\max} \left\{U\left(g^c, T^c; a_i\right)\right\} \\ b_R^c\left(x, y; b_j\right) \equiv \underset{b_R \in [\underline{b}, \overline{b}]}{\arg\max} \left\{V\left(g^c, T^c; b_j\right)\right\} \end{cases}$$

Strict concavity of the utility functions involves the single-peakness of the preferences, while the citizen-candidate model we use imposes the unidimensionality of the policy space. We then apply the MVT to the FOCs of the preceding programs. For strong delegation and weak delegation in country 1 and whatever the choice of country 2 may be ($\forall y \in \{sd, wd\}\,,$), we obtain the following identities of the representatives:

$$\begin{cases} a_R^c\left(sd, y; A_m\right) = A_m - \frac{1}{1+\alpha} dB_m \\ b_R^c\left(sd, y; B_m\right) = \frac{1}{1+\alpha} B_m \end{cases} \quad \text{and} \quad \begin{cases} a_R^c\left(wd, y; A_m\right) = A_m - \frac{1-\alpha}{2} dB_m \\ b_R^c\left(wd, y; B_m\right) = \frac{2-\alpha}{2} B_m \end{cases} \tag{7}$$

The first stage of the game concerns the choice of the delegation rule, which corresponds to a simultaneous move for each player. The normal form of the game is given by:[10]

Insert Table 1.

We deduce the following PROPOSITION:

Proposition 1 *Under Condition* (1)*, the resolution of the game* Γ' *yields*
(i) the Subgame Perfect Nash Equilibrium is unique and involves strong delegation in both countries, (sd, sd)*;*
(ii) in each country, the median voter delegates to a representative more conservative than herself:

$$a_R^c\left(sd, sd; A_m\right) = A_m - \frac{dB_m}{1+\alpha} \quad \textit{and} \quad b_R^c\left(sd, sd; B_m\right) = \frac{B_m}{1+\alpha};$$

[9] $U\left(g^c, T^c; a_i\right)$ and $V\left(g^c, T^c; b_i\right)$ are strictly concave with respect to a_R and b_R.

[10] If one or both countries chooses to decide on its policy through a referendum at the first stage of the game, it seems unrealistic to imagine a negotiation between the population and the possible political representative in the other country. Thus we assume that there is a "fictitious" delegation in which the representative corresponds exactly to the median voter. This hypothesis is equivalent to that of sincere voting. Like Laussel and Riezman (2005), we might also assume that only the candidates motivated by winning the elections commit themselves to the ideal policy of the median voter.

(iii) the equilibrium policies are:

$$g^c(sd, sd) = A_m \quad \text{and} \quad T^c(sd, sd) = \frac{dB_m^2}{2(1+\alpha)}.$$

Proof. see TABLE 1. For country 1, the strategy wd and ref are strictly dominated by the strategy sd. The best response to sd is sd for country 2. ■

At the equilibrium, both countries choose a "powerful conservative representative": voters opt for strong delegation (sd, sd) and they strategically delegate to representatives who are more adverse to public spending than they are. PROPOSITION 1 is then a refinement of the paradox of weakness emphasized by Schelling (1960).[11] Each country looks to improve its negotiation position by two means: the identity of the political representative and the kind of delegation which transfers the domestic power to this political representative. The first point is a direct application of the Schelling conjecture: populations strategically choose a representative less in favor of the public good, and thus more reluctant to centralize, in order to improve their bargaining position. The second element, the choice of a strong delegation rule, constitutes the originality of our analysis. By choosing this rule, populations bind themselves in order to improve their initial bargaining situation. The conservative representative is all the more powerful since she remains in place if negotiations fail. Strong delegation is a credible strategic commitment,[12] which modifies the threat of disagreement and finally the bargaining results.

We note that direct democracy is not an equilibrium even though it may be Pareto superior. Indeed we have: $U^c(dd, dd) > U^c(sd, sd)$ for $\alpha > \widetilde{\alpha} = (\sqrt{5} - 1)/2$, $V^c(dd, dd) > V^c(sd, sd)$ for every $\alpha \in [0, 1]$ and $W^c(dd, dd) > W^c(sd, sd)$ for $A^m \leqslant A$ and $B^m \leqslant B$.[13] Comparing direct and representative democracy as Redoano and Scharf (2004) did is inadequate to conclude whether centralization or political integration actually occurs or not. Indeed, the most favorable regime to centralization might not be a PSNE.

A final remark concerns the effect of a change in the bargaining power of country 1 (α): the more powerful country 1 is in international negotiations, the lower the utility level of its median

[11] See Schelling, 1960, pp. 22:

> "the power to constrain an adversary may depend on the power to bind oneself; that, in bargaining, weakness is often strength, freedom may be freedom to capitulate, and to burn bridges behind one may suffice to undo an opponent."

[12] The credibility of our equilibrium corresponds to its subgame perfection.

[13] The preceding conditions are only sufficient and not necessary. We have:

$$W^c(ref, ref) - W^c(sd, sd) = \frac{dB_m}{2}\left[2(A - A_m) + d(2B - B_m)\right]$$

voter ($\partial U^c(sd,sd)/\partial\alpha < 0$, $\partial V^c(sd,sd)/\partial\alpha > 0$). Since identity of representatives are in fact strategic substitutes,[14] an increase in α would induce the median voter in country 1 to delegate to a representative closer to her (him), while in country 2 the representative becomes more conservative ($\partial a_R^c(.)/\partial\alpha > 0$, $\partial b_R^c(.)/\partial\alpha < 0$). Under strong delegation these moves reduce the threat of country 1 to underprovide the public good[15] and then allow country 2 to reduce its transfer ($\partial T^c(sd,sd)/\partial\alpha < 0$). This result might be interpreted as an exploitation of the powerful by the weak.

4.2 International Agreement versus Isolationism

The equilibria under *isolationism* (game Γ) and *international agreement* (game Γ') can be compared in terms of their predictions for policies as well as their implications for welfare. We obtain the following COROLLARY:

Corollary 1 *Under Condition (2), the international agreement leads to the same aggregate welfare as under isolationism ($W^c(sd,sd) = W^{iso}$), but it clearly involves some redistribution from country 2 to country 1: $U^c(sd,sd) > U^{iso}(A_m)$ and $V^c(sd,sd) < V^{iso}(B_m)$.*

The first result highlights the absence of any gain in international negotiation at the aggregate level. Moreover, not only does political integration not allow the countries to reach the Pareto optimum, but it also does not increase the aggregate welfare with respect to the isolationist situation. The sovereignty of the jurisdictions involves a non-cooperative behavior at the initial stage. These non-cooperative behaviors "pollute" the following steps of the game, in particular the cooperative one where the policies are negotiated.

By comparing the threatpoint of the NBS to the isolationist outcome, we observe that the provision of public good is lower in the first case ($g^{nc}(sd,sd) < g^{iso}$). This threat which results from the strategy of delegation in country 1 leads country 2 to pay a positive transfer at the equilibrium, although the equilibrium level of public good remains unchanged with respect to the isolationist equilibrium ($g^c(sd,sd) = g^{iso}$). Political integration involves some redistribution

[14] For $\frac{da_{R'}}{da_R} \in \{0,1\}$ (strong or weak delegation):

$$\begin{aligned} \frac{\partial^2 U(g^c,T^c;a_R)}{\partial a_R \partial b_R} &= -d\left[1-\alpha\left(1-\frac{da_{R'}}{da_R}\right)\right] < 0 \\ \frac{\partial^2 V(g^c,T^c;b_R)}{\partial a_R \partial b_R} &= -\alpha\left(1-\frac{da_{R'}}{da_R}\right) \leqslant 0 \end{aligned}$$

[15] With our quadratic specification, these effects compensate perfectly and the level of public good remains constant with respect to α ($\partial g^c(.)/\partial\alpha = 0$).

among both countries. Country 1 improves its situation at the expense of country 2 by threatening the latter to reduce the level of public good through the identity and the power of its political representative. From another point of view, we may also affirm that the international agreement prevents country 2 from freeriding. Remember that this country corresponds to the developed countries in the context of biodiversity conservation.

COROLLARY 1 shed light on the failure of the internalization of externalities through bargaining. This result relies on our formalization of political integration. By using an NBS, we highlight the incentives of policy-makers to misrepresent their policy preferences. In particular, we emphasize the influence of the threatpoint, which is usually neglected in the literature on political integration in spite of its effect.[16] For instance, Redoano and Scharf (2004) observe that under centralization of policy decision, the region which values the public good the most, is unable to counter freeriding of the low-preference region in centralization. Here we show how the delegation process, through the treathpoint, provides an instrument to reduce these opportunistic behaviors.

5 Ratification Requirement

The preceding Section emphasizes a redistribution issue among countries which would prevent country 2 from playing the game Γ', or equivalently, to consider political integration. In order to go beyond this deadlock, we consider a ratification requirement and study a third game denoted by Γ''. We construct this game by adding a fourth stage to the game Γ':

- *Stage 4, ratification stage:* each country organizes a referendum to ratify the international agreement. In case of rejection by the majority in one country, no international agreement occurs and the negotiations are considered as failed.

Our approach here is more normative since we do not examine the tactical commitments available through ratification requirements. Everything happens as if a clause has been introduced into the international treaty and imposes a national referendum as the sole ratification procedure. The last stage of the game then involves the satisfaction of each national median voter with respect to the isolationist situation. Proceeding by backward induction we consider

[16] A notable exception is the work of Buchholz, Haupt, and Peters (2005) who assume an NBS with side payments. These authors show that political integration might increase ecological damage in comparison with the *status quo* due to the strategic delegation.

the maximization program (5) which becomes constrained. Let $\left(g^c_{Rat}(.), T^c_{Rat}(.)\right)$ denote the equilibrium values:

$$\left(g^c_{Rat}(.), T^c_{Rat}(.)\right) \equiv \underset{(g,T)\in\Psi}{\arg\max} \left\{ \left[U(g,T;a_R) - U^{nc}(a_{R'};a_R)\right]^{\alpha} \left[V(g,T;b_R) - V^{nc}(a_{R'};b_R)\right]^{1-\alpha} \right\} \tag{8}$$

$$s.t. \quad \begin{cases} U(g,T;A_m) \geqslant U^{iso}(A_m) \\ V(g,T;B_m) \geqslant V^{iso}(B_m) \end{cases}$$

After resolving steps 3 (national policy) and 2 (national representation) of the game, we deduce the normal form of the reduced game corresponding to Γ''. Let denote by $U^c_{Rat}(x,y)$ and $V^c_{Rat}(x,y)$ the equilibrium utility levels depending on the delegation rules x and y, we have:

Insert Table 2

We obtain the following PROPOSITION:

Proposition 2 *Under Condition* (1)*, the resolution of the game* Γ'' *yields*
(i) the SPNE is unique and involves weak delegation in country 1 *and strong delegation in country* 2*,* (wd, sd)*;*
(ii) the equilibrium policy-makers are :

$$a^c_{RRat}(wd,sd;A_m) = A_m - \frac{1-\alpha}{2} dB_m \quad and \quad b^c_{RRat}(wd,sd;B_m) = \frac{2-\alpha}{2} B_m;$$

(iii) the equilibrium policies are:

$$g^c_{Rat}(wd,sd) = A_m + \frac{1}{2} dB_m \quad and \quad T^c_{Rat}(wd,sd) = \frac{3-\alpha}{8} dB^2_m.$$

(iv) under Condition (2) *the equilibrium aggregate welfare increases with respect to the isolationsist situation:* $W^c_{Rat}(wd,sd) > W^{iso}$*.*

Proof. see TABLE 2. For country 1 the strategies *sd* and *ref* are strictly dominated by the strategy *wd*. If country 1 plays *wd*, the best response of country 2 is *sd*. ■

Ratification through national referendum limits harmful delegation. A comparison between the equilibriums of the games Γ and Γ'' shows that the international agreement under referendum ratification improves the utility levels of the median voters of each country with respect to isolationism: $U^{iso} < U^c_{Rat}(wd,sd)$ and $V^{iso} < V^c_{Rat}(wd,sd)$. Opposing the equilibrium payoffs in games Γ' and Γ'' allows us to highlight the fact that this ratification procedure reduces the threat of misrepresentation. In both countries, the strategic bias, i.e. the distance between the representative and the median voter decreases: $A_m > a^c_{RRat}(wd, sd) > a^c_R(sd, sd)$ and $B_m > b^c_{RRat}(wd, sd) > b^c_R(sd, sd)$. Moreover, country 1 opts for weak delegation which reduces the power of its representative in negotiations. This change involves an increase in the level of public good ($g^c_{Rat}(wd, sd) > g^c(sd, sd)$) and a decrease in the transfer paid by country 2 ($T^c_{Rat}(wd, sd) < T^c(sd, sd)$). The resulting aggregate welfare remains lower than the Pareto optimum, but it is better than without the ratification requirement (game Γ'), or equivalently, in the isolationist case (game Γ) as long as Condition (2) is respected.[17]

On one hand our analysis differs from Redoano and Scharf (2004), who highlight the fact that delegation can make centralization possible in situations where a referendum would not support it.[18] On the other hand our second Proposition is close to Gradstein (2004) who establishes that political integration can be superior to decentralization if an election round follows.

6 Conclusion

In this paper we study an international agreement in a two-country model with heterogenous policy preferences and international spillovers. At the equilibrium, inhabitants always prefer representative democracy, more accurately, strong delegation, in order to ensure a strategic pre-commitment and to reinforce their bargaining power. Moreover, they delegate strategically to representatives who are less keen on public goods than the country's median voter. These behaviors imply an inefficient international agreement. Indeed, the choice of delegation rule and the identity of the political representatives generate negative political externalities between countries, which cancel the internalization of economic externalities through negotiations. Our main result completes the analysis of Redoano and Scharf (2004), who conclude that centralization is more likely to occur under representative democracy than under direct democracy. We establish here that the second kind of regime is never chosen at the equilibrium. We also reinforce the conclusions of Buchholz, Haupt, and Peters (2005) by considering the choice of delegation rule as endogenous. Finally, we highlight the beneficial role of an *ex post* referendum, imposed as the

[17] We can also note that the median voter of country 1 is worse off with the ratification constraint than without, while the median voter of country 2 enjoys an improvement of her utility: $\forall \alpha \in]0, 1[$, $U^c_{Rat}(wd, sd) < U^c(sd, sd)$ and $V^c_{Rat}(wd, sd) > V^c(sd, sd)$. We may then deduce that the ratification procedure, particularly a referendum, might be a strategic tool in the same way as the delegation rule. Here we have adopted a normative viewpoint.

[18] These authors assume a polar situation where the heterogeneity of the population is represented by a couple of real values. Therefore the strategic delegation is very restricted since there are only two possible types of representatives. Moreover, political integration does not involve any bargaining among representatives.

unique ratification requirement. This rule leads to a welfare-enhancing agreement by restraining the strategic delegations.

Our analysis focused on agreements which specify the provision of a public good in a country or a group of countries and the payments from the other participative countries. An immediate development would be to assume two public goods, one in each country. However, Gradstein (2004) develops a similar framework to study the European Union. In this context, our analysis suggests one possible explanation of the "democratic deficit". By "democratic deficit", we mean the "policy drift" between the will of national majorities and the policies implemented at the international level. Proposition 1 established that national populations prefer to delegate to a representative whose preferences are different from the median voter, even when they are able to support their position through direct democracy. The "democratic deficit" here results from a deliberate and strategic choice of voters.

References

Alesina, A., and E. Spolaore (1997): "On the Number and Size of Nations," *The Quarterly Journal of Economics*, 112(4), 1027–1056.

Barrett, S. (2003): *Environment and Statecraft: The Strategy of Environmental Treaty-Making.* Oxford University Press, Oxford, UK.

Besley, T., and S. Coate (1997): "An Economic Model of Representative Democracy," *The Quarterly Journal of Economics*, 112(1), 85–114.

——— (2003): "Centralized versus Decentralized Provision of Local Public Goods: a Political Economy Approach," *Journal of Public Economics*, 87(12), 2611–2637.

Bolton, P., and G. Roland (1997): "The Break-Up of Nations: A Political Economy Analysis," *The Quarterly Journal of Economics*, 112(4), 1057–1090.

Buchholz, W., A. Haupt, and W. Peters (2005): "International Environmental Agreements and Strategic Voting," *Scandinavian Journal of Economics*, 107(1), 175–195.

Dur, R. A., and H. J. Roelfsema (2005): "Why does Centralisation fail to internalise Policy Externalities?," *Public Choice*, 122(3-4), 395–416.

Goyal, S., and K. Staal (2004): "The Political Economy of Regionalism," *European Economic Review*, 48(3), 563–593.

Gradstein, M. (2004): "Political Bargaining in a Federation: Buchanan Meets Coase," *European Economic Review*, 48(5), 983–999.

Laussel, D. (2002): "Delegation effects in representative democracies: do they foster extremism?," *Journal of Public Economics*, 85(2), 191–205.

Laussel, D., and R. Riezman (2005): "The Sources of Protectionist Drift in Representative Democracies," *European Economic Review*, 49(7), 1855–1876.

Lockwood, B. (2002): "Distributive Politics and the Cost of Centralization," *Review of Economic Studies*, 69(2), 313–337.

Lorz, O., and G. Willmann (2005): "On the Endogenous Allocation of Decision Powers in Federal Structures," *Journal of Urban Economics*, 57(2), 242–257.

Redoano, M., and K. A. Scharf (2004): "The Political Economy of Policy Centralization: Direct versus Representative Democracy," *Journal of Public Economics*, 88(3-4), 799–817.

Schelling, T. (1960): *The Strategy of Conflict*. Harvard University Press, Cambridge, MA.

A Appendix

		Country 2	
		sd	dd
Country 1	sd	$U^c(sd,sd) = \frac{1}{2}A_m^2 + \frac{1}{2(1+\alpha)}d^2B_m^2,$ $V^c(sd,sd) = A_mB_m - \frac{1}{2(1+\alpha)}dB_m^2.$	$U^c(sd,dd) = \frac{1}{2}A_m^2 + \frac{1+\alpha}{2}d^2B_m^2,$ $V^c(sd,dd) = A_mB_m - \frac{1+\alpha}{2}dB_m^2.$
	wd	$U^c(wd,sd) = \frac{1}{2}A_m^2 + \frac{2-\alpha}{8}d^2B_m^2,$ $V^c(wd,sd) = A_mB_m + \frac{1+\alpha}{8}dB_m^2.$	$U^c(wd,dd) = \frac{1}{2}A_m^2 + \frac{1}{2(2-\alpha)}d^2B_m^2,$ $V^c(wd,dd) = A_mB_m + \frac{1-\alpha}{2(2-\alpha)^2}dB_m^2.$
	dd	$U^c(dd,sd) = \frac{1}{2}A_m^2 + \frac{\alpha}{2(1+\alpha)^2}d^2B_m^2,$ $V^c(dd,sd) = A_mB_m - \frac{1}{2(1+\alpha)}dB_m^2.$	$U^c(dd,dd) = \frac{1}{2}A_m^2 + \frac{\alpha}{2}d^2B_m^2,$ $V^c(dd,dd) = A_mB_m + \frac{1-\alpha}{2}dB_m^2.$

Table 1: Normal form of the Negotiation game.

		Country 2	
		sd	dd
Country 1	sd	$U^c_{Rat}(sd,sd) = \frac{1}{2}A_m^2,$ $V^c_{Rat}(sd,sd) = A_mB_m.$	$U^c_{Rat}(sd,dd) = \frac{1}{2}A_m^2,$ $V^c_{Rat}(sd,dd) = A_mB_m.$
	wd	$U^c_{Rat}(wd,sd) = \frac{1}{2}A_m^2 + \frac{2-\alpha}{8}d^2B_m^2,$ $V^c_{Rat}(wd,sd) = A_mB_m + \frac{1+\alpha}{8}dB_m^2.$	$U^c_{Rat}(wd,dd) = \frac{1}{2}A_m^2 + \frac{1}{2(2-\alpha)}d^2B_m^2,$ $V^c_{Rat}(wd,dd) = A_mB_m + \frac{1-\alpha}{2(2-\alpha)^2}dB_m^2.$
	dd	$U^c_{Rat}(dd,sd) = \frac{1}{2}A_m^2,$ $V^c_{Rat}(dd,sd) = A_mB_m.$	$U^c_{Rat}(dd,dd) = \frac{1}{2}A_m^2 + \frac{\alpha}{2}d^2B_m^2,$ $V^c_{Rat}(dd,dd) = A_mB_m + \frac{1-\alpha}{2}dB_m^2.$

Table 2: Normal form of the Negotiation game with an *ex post* referendum.

B Extended Appendix (for reviewers and not for publication)

B.1 Nash Bargaining Solution

We define $a_R^c(x,y;a_i)$ and $b_R^c(x,y;b_i)$ the ideal representative for individual of type a_i in country 1 and b_i in country 2, in case of the delegation rules (x,y) in both countries. Similarly, we denote by $g^c(x,y)$, $T^c(x,y)$ and $W^c(x,y)$ the equilibrium values of the quantity of public good, the transfer and the aggregate welfare for a couple of strategies (x,y), $(x,y) \in \{sd, wd, ref\}^2$. Note that the utility functions $U(g^c(.),T^c(.);a_i)$ and $V(g^c(.),T^c(.);b_j)$ are strictly concave with respect to a_R and b_R. Since $\frac{\partial a_{R'}}{\partial a_R} \in \{0,1\}$ and $\frac{\partial a_{R'}^2}{\partial^2 a_R} = 0$, we have:

$$\begin{aligned}
\frac{\partial^2 U(g^c(x,y),T^c(x,y);a_i)}{\partial a_R^2} &= -2+\alpha+(1-\alpha)\left(2-\frac{da_{R'}}{da_R}\right)\frac{da_{R'}}{da_R} < 0, \\
\frac{\partial^2 V(g^c(x,y),T^c(x,y);b_j)}{\partial b_R^2} &= -d(1+\alpha) < 0.
\end{aligned}$$

We have to consider the nine possible cases. We present the resolution of the game where both countries choose *sd*. The other developments are similar.[19]

1. For (sd, sd), we have : $a_{R'} = a_R$ and $b_{R'} = b_R$. The system (6) becomes:

$$\begin{cases} g^c(a_R,b_R) = a_R + db_R \\ T^c(a_R,a_R,b_R) = \frac{(1+\alpha)}{2} db_R^2 \end{cases}$$

Substituting these expressions in the utility functions, one can determine the optimal policies chosen by the representatives in the case of strong delegation. These choices are solutions of the following system:

$$\begin{cases} a_R^c(sd,sd;a_i) \equiv \arg\max\limits_{a_R \in [\underline{a},\overline{a}]} \{U(g^c(a_R,b_R),T^c(a_R,a_R,b_R);a_i)\} \\ b_R^c(sd,sd;b_j) \equiv \arg\max\limits_{b_R \in [\underline{b},\overline{b}]} \{V(g^c(a_R,b_R),T^c(a_R,a_R,b_R);b_j)\} \end{cases}$$

which yields:

$$\begin{cases} a_R^c(sd,sd;a_i) = a_i - \frac{1}{1+\alpha} db_j \\ b_R^c(sd,sd;b_j) = \frac{b_j}{1+\alpha} \end{cases} \quad (9)$$

Applying the MVT to (9) involves:

$$\begin{cases} a_R^c(sd,sd;A_m) = A_m - \frac{1}{1+\alpha} dB_m \\ b_R^c(sd,sd;B_m) = \frac{B_m}{1+\alpha} \end{cases} \quad (10)$$

We deduce that

$$\begin{cases} g^c(sd,sd) = A_m \\ T^c(sd,sd) = \frac{1}{2(1+\alpha)} dB_m^2 \end{cases}$$

and

$$\begin{aligned}
U^c(sd,sd) &\equiv U(g^c(sd,sd),T^c(sd,sd);A_m) = \frac{1}{2}A_m^2 + \frac{1}{2(1+\alpha)} d^2 B_m^2, \\
V^c(sd,sd) &\equiv V(g^c(sd,sd),T^c(sd,sd);B_m) = A_m B_m - \frac{1}{2(1+\alpha)} dB_m^2, \\
W^c(sd,sd) &= A_m\left(A - \frac{A_m}{2} + dB\right).
\end{aligned}$$

2. For (sd, wd), country 1 chooses strong delegation while country 2 prefers weak delegation, we have the same results as for (sd, sd). Indeed, weak delegation in country 2 does not affect behavior in country 1, since in the case of disagreement, every inhabitant in country 2 prefers to pay no transfer to country 1.

[19] To avoid computational errors, we have checked our calculations with *Mathematica*. The details are available upon request.

3. For (sd, dd), $a_{R'} = a_R$ and $b_{R'} = b_R = B_m$, the system (6) becomes:

$$\begin{cases} g^c(a_R, B_m) = a_R + dB_m \\ T^c(a_R, a_R, B_m) = \frac{1+\alpha}{2} dB_m^2 \end{cases}$$

We determine the optimal choice of the representative in country 1, solution of:

$$a_R^c(sd, dd; a_i) \equiv \arg\max_{a_R \in [\underline{a}, \overline{a}]} \{U(g^c(a_R, B_m), T^c(a_R, a_R, B_m); a_i)\} \tag{11}$$

The FOC of (11) involves $a_R^c(sd, dd; a_i) = a_i - dB_m$, which gives after applying the MVT: $a_R^c(sd, dd; A_m) = A_m - dB_m$. We obtain

$$\begin{cases} g^c(sd, dd) = A_m \\ T^c(sd, dd) = \frac{(1+\alpha)}{2} dB_m^2 \end{cases}$$

and

$$\begin{aligned} U^c(sd, dd) &= \frac{1}{2}A_m^2 + \frac{1+\alpha}{2} d^2 B_m^2, \\ V^c(sd, dd) &= A_m B_m - \frac{1+\alpha}{2} dB_m^2, \\ W^c(sd, dd) &= A_m\left(A - \frac{A_m}{2} + dB\right). \end{aligned}$$

4. For (wd, sd), $a_{R'} = A_m$, the system (6) becomes:

$$\begin{cases} g^c(a_R, b_R) = a_R + db_R \\ T^c(a_R, A_m, b_R) = \frac{1}{2d}(a_R - A_m + db_R)((1-\alpha)(A_m - a_R) + (1+\alpha) db_R) \end{cases}$$

The optimal choices of the political representatives in this case induce:

$$\begin{cases} a_R^c(wd, sd; a_i) \equiv \arg\max_{a_R \in [\underline{a}, \overline{a}]} \{U(g^c(a_R, b_R), T^c(A_m, a_{R'}, b_R); a_i)\} \\ b_R^c(wd, sd; b_j) \equiv \arg\max_{b_R \in [\underline{b}, \overline{b}]} \{V(g^c(a_R, b_R), T^c(A_m, a_{R'}, b_R); b_j)\} \end{cases}$$

which yields:

$$\begin{cases} a_R^c(wd, sd; a_i) = \frac{1}{2}(a_i - db_i + A_m + \alpha(a_i + db_j - A_m)) \\ b_R^c(wd, sd; b_j) = b_j - \frac{\alpha}{2d}(a_i + db_j - A_m) \end{cases}$$

Applying the MVT involves

$$\begin{cases} a_R^c(wd, sd; A_m) = A_m - \frac{1-\alpha}{2} dB_m \\ b_R^c(wd, sd; B_m) = \frac{2-\alpha}{2} B_m \end{cases}$$

We obtain

$$\begin{cases} g^c(wd, sd) = A_m + \frac{1}{2} dB_m \\ T^c(wd, sd) = \frac{3-\alpha}{8} dB_m^2 \end{cases}$$

and

$$\begin{aligned} U^c(wd, sd) &= \frac{1}{2}A_m^2 + \frac{2-\alpha}{8} d^2 B_m^2, \\ V^c(wd, sd) &= A_m B_m + \frac{1+\alpha}{8} dB_m^2, \\ W^c(wd, sd) &= \frac{1}{8}(2A_m + dB_m)[4A - 2A_m + d(4B - B_m)]. \end{aligned}$$

5. For (wd, wd), the results are identical.

6. For (wd, dd), $a_{R'} = A_m$ and $b_R = B_m$, the system (6) becomes:

$$\begin{cases} g^c(a_R, B_m) = a_R + dB_m \\ T^c(a_R, A_m, B_m) = \frac{1}{2d}(a_R - A_m + dB_m)((1-\alpha)(A_m - a_R) + (1+\alpha) dB_m) \end{cases}$$

The optimal choice of the representatives in country 1 is given by:

$$a_R^c(wd, dd; a_i) \equiv \underset{a_R \in [\underline{a}, \overline{a}]}{\arg\max} \{U(g^c(a_R, B_m), T^c(A_m, a_{R'}, B_m); a_i)\}$$

which yields: $a_R^c(wd, dd; a_i) = \frac{1}{2-\alpha}[a_i - (1-\alpha)(A_m - dB_m)]$. Applying the MVT involves:

$$a_R^c(wd, dd; A_m) = A_m - \frac{1-\alpha}{2-\alpha} dB_m.$$

We deduce that:

$$\begin{cases} g^c(wd, dd) = A_m + \frac{1}{2-\alpha} dB_m \\ T^c(wd, dd) = \frac{(3-\alpha)}{8(2-\alpha)^2} dB_m^2 \end{cases}$$

and

$$\begin{aligned} U^c(wd, dd) &= \frac{1}{2}A_m^2 + \frac{1}{2(2-\alpha)} d^2 B_m^2, \\ V^c(wd, dd) &= A_m B_m + \frac{1-\alpha}{2(2-\alpha)^2} dB_m^2, \\ W^c(wd, dd) &= \frac{((2-\alpha)A_m + dB_m)[(2-\alpha)(2A - A_m) + d(2(2-\alpha)B - B_m)]}{2(2-\alpha)^2}. \end{aligned}$$

7. For (dd, sd), $a_{R'} = a_R = A_m$ and $b_{R'} = b_R$, the system (6) becomes:

$$\begin{cases} g^c(a_R, b_R) = A_m + db_R \\ T^c(a_R, a_R, b_R) = \frac{1+\alpha}{2d} db_R^2 \end{cases}$$

One can determine the optimal choice of the representative in country 2. It yields:

$$b_R^c(dd, sd; b_i) \equiv \underset{b_R \in [\underline{b}, \overline{b}]}{\arg\max} \{V(g^c(A_m, b_R), T^c(A_m, A_m, b_R); b_i)\}$$

The FOC gives: $b_R^c(dd, sd; b_i) = \frac{b_i}{1+\alpha}$, which involves by applying the MVT $b_R^c(dd, sd; B_m) = \frac{B_m}{1+\alpha}$. We deduce that

$$\begin{cases} g^c(dd, sd) = A_m + \frac{1}{1+\alpha} dB_m \\ T^c(dd, sd) = \frac{1}{2(1+\alpha)} dB_m^2 \end{cases}$$

and

$$\begin{aligned} U^c(dd, sd) &= \frac{1}{2}A_m^2 + \frac{\alpha}{2(1+\alpha)^2} d^2 B_m^2, \\ V^c(dd, sd) &= A_m B_m - \frac{1}{2(1+\alpha)} dB_m^2, \\ W^c(dd, sd) &= -\frac{1}{2(1+\alpha)^2}[(1+\alpha)A_m + dB_m][(1+\alpha)(A_m - 2A) + d(B_m - 2(1+\alpha)B)]. \end{aligned}$$

8. For (dd, wd), the results are identical.

9. For (dd, dd), $a_{R'} = a_R = A_m$ and $b_{R'} = b_R = B_m$, the system (6) becomes:

$$\begin{cases} g^c(A_m, B_m) = A_m + dB_m \\ T^c(A_m, A_m, B_m) = \frac{1+\alpha}{2} dB_m^2 \end{cases}$$

We deduce that

$$\begin{aligned} U^c(dd, dd) &= \frac{1}{2}A_m^2 + \frac{\alpha}{2} d^2 B_m^2, \\ V^c(dd, dd) &= A_m B_m + \frac{1-\alpha}{2} dB_m^2, \\ W^c(dd, dd) &= \frac{1}{2}(A_m + dB_m)[(2A - A_m) + d(2B - B_m)]. \end{aligned}$$

We the obtain Table 1.

For country 1, the strategy wd and dd are strictly dominated by the strategy sd. Indeed, we have:

$$\begin{aligned}
U^c(sd,sd) - U^c(wd,sd) &= \frac{2-\alpha+\alpha^2}{8(1+\alpha)} d^2 B_m^2 > 0, \\
U^c(sd,sd) - U^c(dd,sd) &= U^N(sd,dd) - U^N(dd,dd) = \frac{1}{2(1+\alpha)^2} d^2 B_m^2 > 0, \\
U^c(sd,dd) - U^c(wd,dd) &= \frac{1+\alpha-\alpha^2}{2(2-\alpha)} d^2 B_m^2 > 0.
\end{aligned}$$

If country 1 plays sd, country 2 will play sd since $V^c(sd,sd) > V^c(sd,dd)$.[20] We can conclude that (sd,sd) is the unique Perfect Nash Equilibrium of the game.

Using the preceding results yields:

$$\begin{aligned}
W^c(dd,dd) - W^c(sd,sd) &= W^c(dd,dd) - W^c(sd,dd) \\
&= \frac{1}{2} dB_m \left[2(A - A_m) + d(2B - B_m)\right], \\
W^c(dd,dd) - W^c(dd,sd) &= \frac{\alpha}{2(1+\alpha)^2} dB_m \left[2(1+\alpha)(A - A_m) + d(2(1+\alpha)B - (2+\alpha)B_m)\right], \\
W^c(dd,dd) - W^c(wd,sd) &= \frac{1}{8} dB_m \left[4(A - A_m) + d(4B - 3B_m)\right], \\
W^c(dd,dd) - W^c(wd,dd) &= \frac{1-\alpha}{2(2-\alpha)^2} dB_m \left[2(2-\alpha)(A - A_m) + d(2(2-\alpha)B - (3-\alpha)B_m)\right].
\end{aligned}$$

We deduce that $\forall\alpha \in [0,1]$, if $A_m \leqslant A$ and $B_m \leqslant \frac{4}{3}B$, we have:

$$W^c(dd,dd) = \max\left\{W^c(sd,sd), W^c(dd,sd), W^c(wd,sd), W^c(wd,dd)\right\}.$$

B.2 *Ex post* Referendum

We denote by $U^c_{Rat}(x,y)$, $V^c_{Rat}(x,y)$ and $W^c_{Rat}(x,y)$ the equilibrium values of the utilities and the aggregate welfare for the couple strategies (x,y). As in Appendix A.1, we restrict our presentation to the resolution of the game where both countries choose sd. The other developments are available upon request. For (sd,sd), we have: $a_{R'} = a_R$ and $b_{R'} = b_R$. From maximization program (8) we deduce the following Lagrangian function:

$$\begin{aligned}
L(g,T;\lambda,\mu) = &\left[U(g,T;a_R) - U^{nc}(a_{R'};a_R)\right]^{\alpha} \left[V(g,T;b_R) - V^{nc}(a_{R'};b_R)\right]^{1-\alpha} \\
&-\lambda\left[U(g,T;A_m) - U^{iso}(A_m)\right] - \mu\left[V(g,T;B_m) - V^{iso}(B_m)\right]
\end{aligned} \tag{12}$$

where λ and μ are the Lagrange multipliers. For each couple of delegation rule (x,y) we have to consider four cases depending on the values of the Lagrange multipliers.

1. For (sd,sd) we have:

   - If $\lambda = 0$ and $\mu = 0$, we obtain the same solution as in the non-constrained case and we reject this case since: $V^c(sd,sd) = A_m B_m - \frac{1}{2(1+\alpha)} dB_m^2 < V^{iso}(B_m)$ (see Table 3).
   - If $\lambda \neq 0$ and $\mu \neq 0$, we are back to the decentralized equilibrium: $U^c_{Rat}(sd,sd) = U^{iso}(A_m)$ and $V^c_{Rat}(sd,sd) = V^{iso}(B_m)$.
   - If $\lambda \neq 0$ and $\mu = 0$, the first constraint is binding: $U(g^c_{Rat}, T^c_{Rat}; A_m) = U^{iso}(A_m)$. The maximization of (12) yields:

   $$\begin{cases}
   g^c_{Rat}(sd,sd) = a_R + db_R + \frac{d\lambda(A_m - a_R)}{1-\alpha} \\
   T^c_{Rat}(sd,sd) = \frac{d[\lambda(a_R - A_m) - (1-\alpha)b_R][(A_m - a_R)\lambda(-1+\alpha+d\lambda) - (1-\alpha)b_R(1+\alpha-d\lambda)]}{2(1-\alpha)^2(1-d\lambda)}
   \end{cases}$$

   Substituting these expressions of T and g, we determine the identity of the representatives:

   $$\begin{cases}
   a^c_{RRat}(sd,sd) = \frac{A_m\left[(1-\lambda d)\alpha^3 - (1-2\lambda d)\alpha^2 + (1-\lambda d)^2\alpha + (1-\lambda d)^3\right] - dB_m(1-\alpha)^2(1-d\lambda)^2}{\alpha^3(1-\lambda d) - \alpha^2(1-2\lambda d) + \alpha(1-\lambda d)^2 + (1-\lambda d)^3} \\
   b^c_{RRat}(sd,sd) = \frac{B_m(1-\lambda d)(1-\alpha-\lambda d)[1-\alpha-\lambda d(2-\alpha)]}{\alpha^3(1-\lambda d) - \alpha^2(1-2\lambda d) + \alpha(1-\lambda d)^2 + (1-\lambda d)^3}
   \end{cases}$$

[20] Note that: $\frac{1}{2(1+\alpha)} < \frac{1}{2} < \frac{1+\alpha}{2}$, $\forall\alpha \in [0,1]$.

Using the expressions of g^c_{Rat},T^c_{Rat}, a^c_{RRat} and b^c_{RRat}, we determine for which value of $\lambda\,(>0)$, the other constraint, formally $U\left(g^c_{Rat},T^c_{Rat};A_m\right)=U^{iso}\left(A_m\right)$, is respected. Except for $\lambda=\frac{1}{d}$, which yields an indeterminate form of T and thus of $U\,(.)$ and $V\,(.)$, there are only two possible solutions for λ: $\lambda_1=\frac{1-\alpha}{d}$ and $\lambda_2=\frac{1-\alpha}{d(2-\alpha)}$. Considering λ_1, both constraints are binding, since $g^c_{Rat}\left(sd,sd\right)=A_m$ and $T^c_{Rat}\left(sd,sd\right)=0$. The utility values are then equivalent to those at the decentralized equilibrium. For $\lambda_2=\frac{1-\alpha}{d(2-\alpha)}$, we observe that $g^c_{Rat}\left(sd,sd\right)=A_m-\left(1-\alpha\right)dB_m<g^{iso}$ and $T^c_{Rat}\left(sd,sd\right)>0$, it is then obvious that the median voter of country 2 would be better under separation and reject political integration.

- If $\lambda=0$ and $\mu\neq0$, the second constraint is binding: $V\left(g^c_{Rat},T^c_{Rat};B_m\right)=V^{iso}\left(B_m\right)$, which involves: $T^c_{Rat}=B_m\left(g-A_m\right)$ or equivalently $g^c_{Rat}\left(sd,sd\right)=A_m+\frac{T^c_{Rat}}{B_m}$. The maximization of (12) involves:

$$\begin{cases} g^c_{Rat}\left(sd,sd\right)=\frac{(1-\alpha)(a_R+db_R)+\mu(a_R+dB_m)}{1-\alpha+\mu} \\ T^c_{Rat}\left(sd,sd\right)=\frac{d[b_R(1-\alpha)+\mu B_m][(1+\alpha)b_R+\mu B_m]}{2(1+\mu)(1-\alpha+\mu)} \end{cases}$$

After substitution, we deduce the identity of the country's representative:

$$\begin{cases} a^c_{RRat}\left(sd,sd\right)=A_m-\frac{dB_m}{1+\alpha} \\ b^c_{RRat}\left(sd,sd\right)=\frac{[1-\alpha(1+\mu)]B_m}{1-\alpha^2} \end{cases}$$

Using these expressions, we determine the value of μ which ensures that $V\left(g^c_{Rat},T^c_{Rat};B_m\right)=V^{iso}\left(B_m\right)$. It yields: $\mu=\frac{1-\alpha}{2\alpha-1}$. Note that $\mu>0$ involves that $\alpha>\frac{1}{2}$. Under this assumption, the equilibrium levels of the utility function correspond to the decentralized one.
To resume the case (sd,sd), we obtain systematically the values of decentralized equilibrium, with: $U^c_{Rat}\left(sd,sd\right)=\frac{A_m^2}{2}$ and $V^c_{Rat}\left(sd,sd\right)=A_mB_m$.

2. For (sd,wd), the results are identical.

3. For (sd,dd), $a_{R'}=a_R$ and $b_{R'}=b_R=B_m$, we still have four cases.

- If $\lambda=0$ and $\mu=0$, we obtain the same solution as in the non-constrained case (see table 3). Note that: $U^N\left(wd,sd\right)>U^{iso}\left(A_m\right)$ and $V^N\left(wd,sd\right)>V^{iso}\left(B_m\right)$.
- If $\lambda\neq0$ and $\mu\neq0$, we are back to the decentralized equilibrium.
- If $\lambda\neq0$ and $\mu=0$, the maximization of (12) yields:

$$\begin{cases} g^c_{Rat}\left(sd,dd\right)=a_R+db_R+\frac{d\lambda(A_m-a_R)}{1-\alpha} \\ T^c_{Rat}\left(sd,dd\right)=\frac{d[\lambda(a_R-A_m)-(1-\alpha)B_m][(A_m-a_R)\lambda(-1+\alpha+d\lambda)-(1-\alpha)B_m(1+\alpha-d\lambda)]}{2(1-\alpha)^2(1-d\lambda)} \end{cases}$$

Substituting these expressions of $T^c_{Rat}\left(sd,dd\right)$ and $g^c_{Rat}\left(sd,dd\right)$, we determine the identity of country 1's representative:

$$a^c_{RRat}\left(sd,dd\right)=A_m+dB_m\left(\frac{\alpha}{1-\alpha-d\lambda}-\frac{1}{1-\alpha-\left(2-\alpha\right)d\lambda}\right)$$

Using the expressions of g^c_{Rat}, T^c_{Rat}, a^c_{RRat} and b^c_{RRat}, we observe that no positive real value of $\lambda\,(>0)$ allows us to respect the constraint: $U\left(g^c_{Rat},T^c_{Rat};A_m\right)=U^{iso}\left(A_m\right)$. This case must be rejected.
- If $\lambda=0$ and $\mu\neq0$, the second constraint is binding: $V\left(g^c_{Rat},T^c_{Rat};B_m\right)=V^{iso}\left(B_m\right)$, which involves: $T^c_{Rat}\left(sd,dd\right)=B_m\left(g^c_{Rat}-A_m\right)$, or equivalently, $g^c_{Rat}\left(sd,dd\right)=A_m+\frac{T}{B_m}$. The maximization of (12) involves:

$$\begin{cases} g^c_{Rat}\left(sd,dd\right)=a^c_{RRat}+dB_m \\ T^c_{Rat}\left(sd,dd\right)=\frac{(1+\alpha+\mu)}{2(1+\mu)}dB_m^2 \end{cases}$$

After substitution, we deduce the identity of country 1's representative:

$$a^c_{RRat}\left(sd,dd\right)=A_m-dB_m.$$

Using these expressions, the unique value of μ is negative, equal to $-1-\alpha$ and this case is not relevant.
Thus, we deduce that: $U^c_{Rat}\left(sd,sd\right)=\frac{A_m^2}{2}$ and $V^c_{Rat}\left(sd,sd\right)=A_mB_m$.

4. For (wd, sd), $a_{R'} = A_m$,

   - If $\lambda = 0$ and $\mu = 0$, we obtain the same solution as in the non-constrained case:

$$\begin{cases} U^c_{Rat}(wd, sd) = U^c(wd, sd) = \frac{1}{2}A_m^2 + \frac{2-\alpha}{8}d^2B_m^2 > U^{iso}(A_m) \\ V^c_{Rat}(wd, sd) = V^c(wd, sd) = A_mB_m + \frac{1+\alpha}{8}dB_m^2 > V^{iso}(B_m) \end{cases}$$

   - If $\lambda \neq 0$ and $\mu \neq 0$, we are back to the decentralized equilibrium.
   - If $\lambda \neq 0$ and $\mu = 0$, the maximization of (12) yields:

$$\begin{cases} g^c_{Rat}(wd, sd) = a_R + dB_m + \frac{d\lambda(A_m-a_R)}{1-\alpha} \\ T^c_{Rat}(wd, sd) = \frac{[(1-\alpha)(A_m-a_R-db_R)+d\lambda(a_R-A_m)]\left[(1-\alpha)^2(a_R-A_m)+\left(1-\alpha^2\right)db_R-(1-\alpha)d^2\lambda b_R+d^2\lambda^2(A_m-a_R)\right]}{2d(1-\alpha)^2(1-d\lambda)} \end{cases}$$

     Substituting these expressions of T and g, we determine the identity of the country 1's representative:

$$\begin{cases} a^c_{RRat}(wd, sd) = A_m - \frac{(1-\alpha)^2(1-d\lambda)}{(2-\lambda d)(1-\alpha-\lambda d)}dB_m \\ b^c_{RRat}(wd, sd) = \frac{B_m(1-\lambda d)(2-\alpha)}{2-\lambda d} \end{cases}$$

     Using the expressions of g^c_{Rat},T^c_{Rat}, a^c_{RRat} and b^c_{RRat}, we observe that the unique solution of $U(g^c_{Rat}, T^c_{Rat}; A_m) = U^{iso}(A_m)$ yields $\lambda = \frac{1}{d}$, which involves an indeterminate form for T.
   - If $\lambda = 0$ and $\mu \neq 0$, the second constraint is binding: $V(g^c_{Rat}, T^c_{Rat}; B_m) = V^{iso}(B_m)$. The maximization of (12) involves:

$$\begin{cases} g^c_{Rat}(wd, sd) = \frac{(1-\alpha)(a_R+db_R)+\mu(a_R+dB_m)}{1-\alpha+\mu} \\ T^c_{Rat}(wd, sd) = \frac{[(1-\alpha)(a_R-A_m+db_R)+\mu(a_R-A_m+dB_m)]((1-\alpha+\mu)(A_m-a_R)+d[(1+\alpha)b_R+\mu B_m])}{2d(1-\alpha+\mu)(1+\mu)} \end{cases}$$

     After substitution, we deduce the identity of the representatives:

$$\begin{cases} a^c_{RRat}(wd, sd) = A_m + \frac{1+\mu-2\alpha+\alpha^2-(1-\alpha)\alpha\mu}{\alpha(2+\mu)-2(1+\mu)}dB_m \\ b^c_{RRat}(wd, sd) = \frac{\alpha\mu(3+\mu)-\alpha^2(1+\mu)+3\alpha-2(1+\mu)}{\alpha(2+\mu)-2(1+\mu)}B_m \end{cases}$$

     Using these expressions, the solution of $V(g^c_{Rat}, T^c_{Rat}; B_m) = V^{iso}(B_m)$ in μ is negative.
     Thus, we deduce that: $U^c_{Rat}(wd, sd) = \frac{1}{2}A_m^2 + \frac{2-\alpha}{8}d^2B_m^2$ and $V^c_{Rat}(wd, sd) = A_mB_m + \frac{1+\alpha}{8}dB_m^2$.

5. For (wd, wd), the results are identical.

6. For (wd, dd), we have:

   - If $\lambda = 0$ and $\mu = 0$, we obtain the same solution than in the non-constrained case:

$$\begin{cases} U^c_{Rat}(wd, dd) = U^c(wd, dd) = \frac{1}{2}A_m^2 + \frac{1}{2(2-\alpha)}d^2B_m^2 > U^{iso}(A_m) \\ V^c_{Rat}(wd, dd) = V^c(wd, dd) = A_mB_m + \frac{1-\alpha}{2(2-\alpha)^2}dB_m^2 > V^{iso}(B_m) \end{cases}$$

   - If $\lambda \neq 0$ and $\mu \neq 0$, we are back to the decentralized equiblibrium.
   - If $\lambda \neq 0$ and $\mu = 0$, the maximization of (12) yields:

$$\begin{cases} g^c_{Rat}(wd, dd) = a_R + dB_m + \frac{d\lambda(A_m-a_R)}{1-\alpha} \\ T^c_{Rat}(wd, dd) = \frac{[(1-\alpha)(A_m-a_R-dB_m)+d\lambda(a_R-A_m)]\left[(1-\alpha)^2(a_R-A_m)+\left(1-\alpha^2\right)dB_m-(1-\alpha)d^2\lambda B_m+d^2\lambda^2(A_m-a_R)\right]}{2d(1-\alpha)^2(1-d\lambda)} \end{cases}$$

     Substituting these expressions of T^c_{Rat} and g^c_{Rat}, we determine the identity of country 1's representative:

$$a^c_{RRat}(wd, dd) = A_m - \frac{(1-\alpha)^2}{(2-\alpha)(1-\alpha-\lambda d)}dB_m.$$

     Using the expressions of g^c_{Rat},T^c_{Rat}, a^c_{RRat} and b^c_{RRat}, we observe that there is no solution of $U(g^c_{Rat}, T^c_{Rat}; A_m) = U^{iso}(A_m)$ in λ.

- If $\lambda = 0$ and $\mu \neq 0$, the second constraint is binding: $V\left(g^c_{Rat}, T^c_{Rat}; B_m\right) = V^{iso}\left(B_m\right)$. The maximization of (12) involves:

$$\left\{ \begin{array}{c} g^c_{Rat}\left(wd, dd\right) = a_R + dB_m \\ T^c_{Rat}\left(wd, dd\right) = \frac{\left(a_R - A_m + db_R\right)\left[\left(1-\alpha+\mu\right)\left(A_m - a_R\right) + \left(1+\alpha+\mu\right)dB_m\right]}{2d\left(1+\mu\right)} \end{array} \right.$$

After substitution, we deduce the identity of the country 1's representatives:

$$a^c_{RRat}\left(wd, dd\right) = A_m - \frac{1-\alpha+\mu}{2\left(1+\mu\right)-\alpha} dB_m.$$

Using these expressions, the solution of μ for $V\left(g^c_{Rat}, T^c_{Rat}; B_m\right) = V^{iso}\left(B_m\right)$ is negative.
Thus, we deduce that: $U^c_{Rat}\left(wd, dd\right) = \frac{1}{2}A^2_m + \frac{1}{2(2-\alpha)}d^2B^2_m$ and $V^c_{Rat}\left(wd, sd\right) = A_mB_m + \frac{1-\alpha}{2(2-\alpha)^2}dB^2_m$.

7. For (dd, sd), $a_{R'} = a_R = A_m$ and $b_{R'} = b_R$, it yields:

- If $\lambda = 0$ and $\mu = 0$, we obtain the same solution as in the non-constrained case:

$$\left\{ \begin{array}{c} U^c_{Rat}\left(dd, sd\right) = U^c\left(dd, sd\right) = \frac{1}{2}A^2_m + \frac{\alpha}{2(1+\alpha)^2}d^2B^2_m > U^{iso}\left(A_m\right) \\ V^c_{Rat}\left(dd, sd\right) = V^c\left(dd, sd\right) = A_mB_m - \frac{1}{2(1+\alpha)}dB^2_m < V^{iso}\left(B_m\right) \end{array} \right.$$

- If $\lambda \neq 0$ and $\mu \neq 0$, we are back to the decentralized equilibrium.
- If $\lambda \neq 0$ and $\mu = 0$, the maximization of (12) yields:

$$\left\{ \begin{array}{c} g^c_{Rat}\left(dd, sd\right) = A_m + db_R \\ T^c_{Rat}\left(dd, sd\right) = \frac{(1+\alpha-d\lambda)}{2(1-d\lambda)}db^2_R \end{array} \right.$$

Substituting these expressions of T and g, we determine the identity of country 2's representative:

$$b^c_{RRat}\left(dd, sd\right) = B_m - \frac{\alpha B_m}{1+\alpha-\lambda d}.$$

Using the expressions of g^c_{Rat}, T^c_{Rat}, a^c_{RRat} and b^c_{RRat}, we observe that there is no solution of $U\left(g^c_{Rat}, T^c_{Rat}; A_m\right) = U^{iso}\left(A_m\right)$ in λ.
- If $\lambda = 0$ and $\mu \neq 0$, the second constraint is binding: $V\left(g^c_{Rat}, T^c_{Rat}; B_m\right) = V^{iso}\left(B_m\right)$. The maximization of (12) involves:

$$\left\{ \begin{array}{c} g^c_{Rat}\left(dd, sd\right) = \frac{(1-\alpha)(A_m+db_R)+\mu(A_m+dB_m)}{1-\alpha+\mu} \\ T^c_{Rat}\left(dd, sd\right) = \frac{d\left[(1-\alpha)b_R+\mu B_m\right]\left[(1+\alpha)b_R+\mu B_m\right]}{2(1-\alpha+\mu)(1+\mu)} \end{array} \right.$$

After substitution, we deduce the identity of country 2's representatives:

$$b^c_{RRat}\left(dd, sd\right) = \frac{1-\alpha-\alpha\mu}{1-\alpha^2}B_m.$$

Using these expressions, the solution of $V\left(g^c_{Rat}, T^c_{Rat}; B_m\right) = V^{iso}\left(B_m\right)$ in μ is negative, equal to $-1+\alpha$.
Thus, we deduce that: $U^c_{Rat}\left(dd, sd\right) = \frac{1}{2}A^2_m$ and $V^c_{Rat}\left(dd, sd\right) = A_mB_m$.

8. For (dd, wd), the results are identical.

9. For (dd, dd), $a_{R'} = a_R = A_m$ and $b_{R'} = b_R = B_m$, it yields:

- If $\lambda = 0$ and $\mu = 0$, we obtain the same solution as in the non-constrained case:

$$\left\{ \begin{array}{c} U^c_{Rat}\left(dd, dd\right) = U^c\left(dd, dd\right) = \frac{1}{2}A^2_m + \frac{\alpha}{2}d^2B^2_m > U^{iso}\left(A_m\right) \\ V^c_{Rat}\left(dd, dd\right) = V^c\left(dd, dd\right) = A_mB_m + \frac{1-\alpha}{2}dB^2_m > V^{iso}\left(B_m\right) \end{array} \right.$$

- If $\lambda \neq 0$ and $\mu \neq 0$, we are back to the decentralized equilibrium.

- If $\lambda \neq 0$ and $\mu = 0$, the maximization of (12) involves:

$$\begin{cases} g^c_{Rat}(dd, dd) = A_m + dB_m \\ T^c_{Rat}(dd, dd) = \frac{(1+\alpha-d\lambda)}{2(1-d\lambda)} dB^2_m \end{cases}$$

Using the expressions of g and T, we observe that there is no solution of $U(g^c_{Rat}, T^c_{Rat}; A_m) = U^{iso}(A_m)$ in λ.

- If $\lambda = 0$ and $\mu \neq 0$, the second constraint is binding : $V(g^c_{Rat}, T^c_{Rat}; B_m) = V^{iso}(B_m)$. The maximization of (12) involves:

$$\begin{cases} g^c_{Rat}(dd, dd) = A_m + dB_m \\ T^c_{Rat}(dd, dd) = \frac{dB^2_m(1+\alpha+\mu)}{2(1+\mu)} \end{cases}$$

Using these expressions, we conclude that there is no positive solution of $V(g^c_{Rat}, T^c_{Rat}; B_m) = V^{iso}(B_m)$ in μ. Thus, we deduce that: $U^c_{Rat}(dd, dd) = \frac{1}{2}A^2_m + \frac{\alpha}{2}d^2B^2_m$ and $V^c_{Rat}(dd, dd) = A_mB_m + \frac{1-\alpha}{2}dB^2_m$.

The equilibrium payments of the nine situations are presented in Table 4. Note that for country 1, the strategies sd and ref are strictly dominated by the strategy wd. Indeed, we have:

$$\begin{aligned} U^c_{Rat}(wd, sd) - U^c_{Rat}(sd, sd) &= U^c_{Rat}(wd, sd) - U^c_{Rat}(dd, sd) = \frac{2-\alpha}{8}d^2B^2_m > 0, \\ U^c_{Rat}(wd, dd) - U^c_{Rat}(sd, dd) &= \frac{1}{2(1+\alpha)^2}d^2B^2_m > 0, \\ U^c_{Rat}(wd, dd) - U^c_{Rat}(dd, dd) &= \frac{(1-\alpha)^2}{2(2-\alpha)}d^2B^2_m > 0. \end{aligned}$$

If country 1 plays wd, country 2 will play sd since $V^c_{Rat}(wd, sd) - V^c_{Rat}(wd, dd) = \frac{\alpha(4-3\alpha+\alpha^2)}{8(2-\alpha)^2}dB^2_m > 0$. We can conclude that (wd, sd) is the unique Nash equilibrium of the game.