

# Gradualism and uncertainty in international union formation: The European Community's first enlargement

Nikitas Konstantinidis

Received: 18 March 2008 / Accepted: 30 July 2008 /  
Published online: 24 September 2008  
© Springer Science + Business Media, LLC 2008

**Abstract** This paper introduces a new theoretical framework of international unions *qua* coalitions of countries adopting a common policy and common supranational institutions. I make use of a three-country spatial bargaining game of coalition formation, in order to examine the endogenous strategic considerations in the creation and enlargement of international unions. Why would we observe a gradualist approach in the formation of the grand coalition even if the latter is assumed to be weakly efficient? I propose *asymmetric information* about the benefits of integration as a mechanism that can generate *gradual* union formation in equilibrium. As it turns out, it may well be in the ‘core’ countries’ interest to delay the accession of a third, ‘peripheral’ country in order to (1) stack the institutional make-up of the initial union in their favor and (2) signal their high resolve to wait out the expansion of their bilateral subunion. A related case from the European experience provides an interesting illustration.

**Keywords** Gradualism · International unions · Coalition formation · Supranational bargaining · EC enlargement · Charles de Gaulle

**JEL Classification** F53 · C72 · D82

## 1 Introduction

This paper introduces a formal game-theoretic framework elucidating various aspects of the ‘variable geometry’ and the dynamics of international union formation

---

N. Konstantinidis (✉)

Political Economy Program, Robertson Hall, Woodrow Wilson School, Princeton University,  
Princeton, NJ 08544-1013, USA  
e-mail: nkonstan@Princeton.EDU  
URL: <http://www.princeton.edu/~nkonstan>

N. Konstantinidis  
IBEI, Elisabets 10, 08001 Barcelona, Spain

with a focus on European integration. I wish to study questions about the coalitional dynamics of the creation and piece-meal widening of a union. In a non-cooperative game of intergovernmental bargaining where the *grand* union is Pareto efficient, why would we ever observe delay in its formation? For example, why did formerly eligible countries like the UK, Denmark, Austria, and Sweden join the European Economic Community (currently known as the European Union) at a much later time? Naturally, I am particularly interested in endogenous strategic considerations in the creation of a union by a core of countries rather than exogenous changes in the geopolitical and geoeconomic environment, which rendered new democracies like Greece, Spain, and Portugal or former communist countries eligible candidate-members at some later stage of the union formation process. Countries such as Greece, Spain, and Portugal and the former communist Central and Eastern European countries were invited to join the existing Union soon after they became politically (i.e., democratic) and/or economically (i.e., liberalized market economies) eligible. These latter cases may be plausibly modeled by postulating exogenous (mainly geopolitical) constraints to entry that cease to bind at some point in time.

Starting from the empirical observation of gradualism and piece-meal expansion of international unions and regional *blocs* (including to various degrees ASEAN and MERCOSUR) in the absence of clearly perceived shocks to the global (or even regional) geopolitical and economic system, I propose *strategic delay* in the intergovernmental bargaining process as a theoretical explanation. International cooperation and regional integration agreements are usually initiated by a core of ‘natural partners’, who wish to reap the immediate gains from cooperation. Hence, even without the formalization of arbitrary eligibility criteria, other aspirant members are at first effectively excluded from the ‘enacting’ coalition only to join later at less favorable institutional and policy terms. So even when the inclusion of a wider collectivity of states has always been Pareto efficient, the formation of the enlarged union is strategically delayed by the founding signatories for reasons to be explained.

To tackle some of the above questions, I make use of a game-theoretic model theorizing about the strategic interaction among states in the realm of international cooperation. The rise of supranationalism does not necessarily imply the decline of the nation-state, which is why I use the latter as the main unit of analysis. For the most part, I treat countries as unitary actors and international unions as coalitions among states. Particularly prevalent within the *realist* tradition in international relations theory, the unitary actor assumption treats democratically elected governments as representative agents seeking to maximize the welfare of the ‘average’ citizen or just some other aggregated national objective.<sup>1</sup> In the same vein, I talk about states, not ruling parties, thus abstracting away from micropolitical considerations of vote maximization; my macrorealist perspective is essentially predicated on the assumption of domestic political consensus with respect to a country’s core

<sup>1</sup> Adding special interests and preference heterogeneity within countries gives rise to the possibility of strategic delegation and cross-country popular alliances and is more conducive to the *liberal intergovernmentalist* approach to regional integration and national preference formation (Moravcsik 1997). This nuanced analytical perspective is much more pervasive among economists in the political economics literature on regional integration (see for example Brou and Ruta 2006).

national interests in the pursuit of international cooperation.<sup>2</sup> In the bulk of the analysis, I choose to subsume these important micro-level questions of national preference formation within exogenous assumptions. In essence, this is a theory of ‘grand bargains’ among states rather than an institutionalist account of the workings and policy-making functions of supranational bodies. The formal nature of the approach renders its results generalizable to other cases of *gradual* coalition formation among countries in the pursuit of international cooperation.

I offer one particular mechanism that can generate *strategic delay* in the formation of the *grand* union: *asymmetric information* over the synergistic benefits generated by unions of which a country is a member. According to the proposed theory, uncertainty over the exogenous coalition surplus may result in *strategic delay* through a semi-separating signaling equilibrium, whereby a positive support of country types choose to initially participate in a smaller union in order to shape the terms of enlargement to their benefit. In other words, I interpret the bandwagoning phenomenon in union formation and expansion as a war-of-attrition game, whereby states place themselves temporally on the coalition-building process in such a way as to signal their resolve in waiting out the formation of the Pareto efficient *grand* union. In equilibrium, the proposal order affects the order of entry to the coalition, which in turn is a strong predictor of surplus allocation.

The next section provides a review of the breadth and scope of the political science literature on international cooperation with a particular emphasis on European integration as well as an account of the political economy literature on international unions. The subsequent section presents a spatial bargaining model of union formation examining *uncertainty* as a cause of gradualism in union formation.<sup>3</sup> I first discuss the game with complete information and then proceed to demonstrate why *asymmetric information* is a prerequisite for gradualism in coalition-formation and the mechanism through which that occurs. I also study the efficiency properties of the proposed equilibrium. By manner of empirical justification, I further go on to show how the theory applies to the case of the first enlargement of the European Community (EC), focusing in particular on the French–German–British triptych of actors and the bargaining dynamics between them with respect to British accession to the EC. The concluding section summarizes the implications of the model and looks at some possible extensions.

---

<sup>2</sup> My case study on the EEC’s first enlargement in the 1960s and early 1970s analyzes for example the various policy shifts that took place despite the continuous dominant presence of Gaullists in power (initially General Charles de Gaulle himself followed by his ideological heir and successor, Georges Pompidou). Of course, the assumption of continuity in economic and political integration policy across partisan lines is just an analytical simplification, not an empirical iron law. Gruber (2000), however, provides a theoretical explanation for the scant evidence of radical policy shifts with respect to decisions to accede to and/or secede from international regimes by ideologically distinct governments.

<sup>3</sup> See Downs and Rocke (1995) for a comprehensive analysis of the role of uncertainty in international cooperation.

## 2 Related Literature

This paper draws from a variety of related work on both positive and normative aspects of international union formation and policy centralization and relates to diverse strands of literature in both economics and political science. It falls within the general field of comparative political economy with a substantive application to regional integration.

The political science literature on international cooperation has been dominated by international relations theorists of various traditions. The early debate on the theoretical and empirical relevance of supranational institutions was instigated by the neoliberal school of thought, giving rise to a vast body of work collectively dubbed as *regime theory* (e.g., Keohane 1984; Milner 1997). Their focus on the ‘mutual gains’ rationale for international cooperation in an anarchic world came as a rebuttal to the Waltzian realist mantra of power politics and national interests, which deemed the emergence of supranational institutions as epiphenomenal to the existing balance of power and essentially inconsequential within the system of international relations. However, the acceleration of the European integration project and the proliferation of regional and global institutions in the 1980s shifted the focus of the neorealist critique from the Pareto efficiency of international regimes (Krasner 1983) onto the ‘relative gains’ of their participants and the enforceability<sup>4</sup> of those decentralized agreements (Grieco 1990; Mearsheimer 1994/1995).

Sharing a common rational choice methodological perspective, the emergence of a neoliberal–neorealist consensus on the importance, causes, and effects of supranationalism has gradually given way to the analysis of distributional considerations in the evolution and design of those institutions. Gruber (2000), for example, views international regimes not simply as *incomplete contracts* or *focal points* in the selection of multiple equilibria of coordination, but essentially as the manifestation of ‘go-it-alone’ power by rational ‘enacting’ governments seeking to restrict the choice set of domestic opponents and ‘peripheral’ states.<sup>5</sup>

Within the same rational choice tradition, formal international relations theorists have ventured of late to model the intricacies of the process and evolution of multilateralism, previously dubbed as too intractable a problem to be solved analytically. Downs et al.’s (1998) simple majority-voting game seeks to explain the rationale for *sequentialism* (or else *gradualism*) in the construction of international agreements and their expansion within the context of liberalization of international norms and standards. Our unanimity-driven model, on the other hand, arrives at a similar result without the postulate of exogenous shifts in preferences and the ‘gains from trade’. Gilligan (2004) shifts the discussion to the breadth vs. depth debate

---

<sup>4</sup> On federalism or supranationalism as a compliance problem see Bednar (2007).

<sup>5</sup> Gruber (2000) is critical of the neoliberal–neorealist consensus on the Pareto efficiency of international regimes and is mostly interested in the winners vs. losers dimension of international cooperation, arguing that it is often the case (citing NAFTA and the European Monetary System as his primary examples) that some late signatory countries to such regimes are better off in an autarchic *status quo ante* of no cooperation than their current state of wider integration. However, the *status quo ante* has been removed from their choice set by the *fait accompli* of partial integration, thereby rendering the costs of joining an existing international cooperation agreement lower than those of staying out.

arguing that such a trade-off in the shape and form of multilateral institutions is a direct result of uniformity in the design of common policies.

While the emphasis among international relations scholars has been primarily on the rationale for international cooperation agreements, their enforceability, and their overall effect on the international system, this paper seeks to shift the focus to the dynamics of union formation and the evolution of their institutional design. Given the wide consensus among political scientists and economists on the existence of ‘mutual gains’ in international cooperation—effectively amounting to a Pareto-improving response to international policy spillovers and externalities intrinsic within a globalized environment of interdependence—the general arguments of the paper concentrate on the strategic calculus of surplus distribution in union formation taking efficiency considerations for granted.

Economists have also strongly contributed to the interdisciplinary field of regional integration by drawing on well-established theories in the fields of public and international economics. Highly influential in these models has been the public economics literature on *fiscal federalism* and *decentralization* (e.g., Oates 1972, 1999; Besley and Coate 2003; Hafer and Landa 2007; Persson and Tabellini 1996; Cremer and Palfrey 2000), which examines the welfare and distribution effects of government structure on the provision of public goods. Widely applicable to the phenomenon of political integration is also the *theory of clubs* and *overlapping jurisdictions* (Casella 1992a; Casella and Feinstein 2002), which models the interaction between markets as sets of rules for the exchange of private goods and institutions as organizations for the provision of public goods.

The above theoretical bodies of work have recently spawned a fast-growing literature on the political economy of federalism, regional integration, and international unions (see for example Alesina et al. 2001, 2005; Alesina and Spolaore 2003; Baldwin 1999; Bolton and Roland 1997; Bordignon and Brusco 2006; Brou and Ruta 2006; Ellingsen 1998; Harstad 2007), which mainly consists of game-theoretic models studying the economic incentives of integration and/or secession (on secession see Bordignon and Brusco 2001) as well as the strategic determinants of country and union size, often yielding normative conclusions on constitutional design.<sup>6</sup> These papers take a non-generic approach to the specification of union benefits by modeling an international union as an efficient central provider of public goods, characterized by *economies of scale* and *spillovers* across union (and non-union) members, as in Alesina et al. (2001, 2005). This modeling approach has been amply applied to explain the coalitional dynamics of European monetary integration and currency unions (Alesina and Grilli 1993; Casella 1992b) and to design the optimal membership rules for the European Monetary Union (EMU)—whether through rigid membership criteria or *gradual* expansion through flexible rules of integration (see Fratianni 1998; Pisani-Ferry 1995). On the basis of the theoretical results of these models, a number of economists have ventured to contribute to the ongoing debate about the institutional structure of the European

---

<sup>6</sup> For a brief survey of economic theories of (dis)integration see Ruta (2005).

Union (see for example Alesina and Wacziarg 1999; Alesina and Perotti 2004; Jacquemin and Sapir 1995).<sup>7</sup>

In contrast to the emphasis of Alesina et al. (2001, 2005) on the stability and size of equilibrium unions in light of public good spillovers, the focus of this paper is on the bargaining dynamics of the coalition-formation game and the strategic incentives inherent in negotiating the creation of an international union of countries. Harstad (2007) actually addresses the trade-off between strategic delay in the process of political centralization and the cost of policy uniformity using a similar signaling mechanism to the one below, albeit within a two-region framework, which implies that he does not consider the possibility of endogenous enlargement. Aghion et al. (2007) employ an analogous dynamic bargaining framework in the context of international free trade agreements to model the choice of a ‘leading country’ between sequential and multilateral negotiations and how it depends on the structure of trade and protection. Finally, the idea of *gradual* and *partial* coalition-formation in international cooperation is also characteristic of Bordignon and Brusco’s (2006) paper on the efficiency and distributional effects of subunion formation within the context of a policy centralization game with shifting ‘gains from trade’. They study the role of commitment and costly transfers with respect to the optimality of ‘enhanced cooperation’ as a formalized rule of flexible integration within a federation, whilst the focus of this paper is on *asymmetric information* as a cause of *gradual* union formation.

In what follows, I analyze the case of a three-country regional setting through a non-cooperative spatial bargaining model, in order to gauge the extent to which the dynamic process of union formation can be explained by endogenous strategic factors such as *uncertainty*. For reasons of analytical tractability, I choose to examine the simple three-country case, so as to allow for the possibility of subcoalitions and endogenous enlargement in the formation of the *grand* union.

### 3 The Model

#### 3.1 Basic Framework

My modeling approach consists of a combination of simple unidimensional spatial analysis with a non-cooperative game of coalition formation in the context of policy centralization within an international union. The spatial approach is based on the interpretation of an international union as a commitment device to centralize policy across countries and is better suited to analyze the bargaining aspects of union formation and/or policy coordination compared to the public goods approach used in the political economics literature. It also provides for a parsimonious formalization of the *liberal intergovernmentalist* (Moravcsik 1998) theory on the ‘grand bargains’ of EU treaties. The coalition formation approach, on the other hand, is permeated by the notion of international unions as coalitions of countries adopting common

<sup>7</sup> The Center for Economic Policy Research issues yearly reports on various issues of European integration providing the opportunity for political economists to contribute to policy-making debates relating to the European Union (see for example Dewatripont et al. 1995 and Berglöf et al. 2003).

supranational policies and institutional structures. Since endogenizing the choice of policy areas to be centralized at the supranational level is not within the scope of this paper, the use of a unidimensional policy variable is meant to capture the institutional shape and ideological character of an international union.

The underlying cooperative game of coalition formation prescribes the set of players (or countries)  $N$  as well as the *value* or *worth*  $y_c \in \mathfrak{R}$  of each non-empty coalition of countries  $c \subseteq N$ . The source of these exogenous ‘functional synergies’ is not explicitly modeled but is implicitly linked to the emerging economic and political interdependence among countries in the era of globalization. These orthogonal and indivisible benefits may thus be construed as pure public goods supplied supranationally and enjoyed only by citizens of union member-states. Their level of provision is assumed to be strictly dependent on the size and membership composition of the union, but not its ideological character. This orthogonality assumption implies that the unidimensional policy variable serves the role of an imperfect (because of insufficient dimensionality) and potentially distortionary transfer mechanism that reflects relative proposal and bargaining power in the coalition-formation process.<sup>8</sup> In light of absolute deviation Euclidean preferences, any shift in the common policy between the bliss points of the proposed union’s members, i.e., within the Pareto set of the coalition, will result in a utility increase for at least one country at the expense of one or more of the others. This in a way mirrors a system of intergovernmental transfers in international negotiations for the pursuit of ‘common gains’.

The model also makes use of a typical assumption in the coalition-formation literature, whereby the *grand* coalition is weakly efficient. This contingency is what I refer to as *weak superadditivity* in the structure of the underlying cooperative game, i.e., the total *worth* of the *grand* union is greater or equal to the sum of the *worths* of any constituent subunions, or—prosaically put—the whole is better than the sum of its parts. Otherwise, it would be trivial to explain why the *grand* coalition doesn’t form or even impossible to explain its *gradual* formation.<sup>9</sup> This assumption is formally defined as:

**Assumption 1** (weak superadditivity)  $y_N \geq \sum_{c \in \pi} y_c(\pi) \quad \forall \pi, c \in \pi$ , where  $\pi$  denotes any partition of the set of players  $N$ .

In what follows, I present a multilateral non-cooperative bargaining game of coalition formation for the simple case of  $N=3$  countries. Within this parsimonious framework, I examine the implications of private information over the size of the exogenous benefits for the potential of *strategic delay* in the formation of the *grand* coalition (i.e., whether coalition formation is *gradual* or *immediate* and whether the *grand* coalition actual forms or not). The general modeling framework consists of

<sup>8</sup> As in Padoan (1997), a political union is basically viewed as an ‘economic club’ that yields excludable and indivisible benefits, given that any such synergistic relationship between sovereign nation-states is embedded within a broader environment of economic interdependence through trade and the exchange of people, ideas, and factors of production (which is the standard view of the engine of integration in post-WWII Western Europe).

<sup>9</sup> Unlike earlier models in coalition theory, recent non-axiomatic work on coalition formation with externalities does not necessarily predict the emergence of the *grand* coalition (unlike the cooperative solution concepts of the *core* and the *Shapley value*).

absolute deviation Euclidean utility functions and a dynamic bargaining protocol with equal recognition probabilities, thus shying away from risk aversion and inequality in agenda-setting power as possible sources of *strategic delay*. The use of a random recognition protocol is just an abstraction for more institutionalized enlargement negotiations, whereby an aspirant member may first have to receive official candidate status before embarking upon negotiations over the exact terms of accession subject to the unanimous approval of existing members. It would be interesting to endogenize the identity of the proposer through a pre-play agenda-setting game for the purpose of comparing the evolutionary dynamics of different regional *blocs*, but that is not essential for our results. I generally restrict my attention to bargaining equilibria that exhibit gradualism in the formation of the *grand* union of all three countries. As it turns out, gradualism by dint of *strategic delay* may only come about once either of the two ideologically extreme countries gets to propose first.

Using the above parsimonious three-country framework, I wish to demonstrate how the existence of private information over the synergistic benefits of any given coalition can hamper political compromise at an earlier date thus giving rise to inefficiencies in the bargaining process. So long as any country may not correctly anticipate the synergies inherent in a political union wherefrom it is excluded, equilibrium delay from the point of view of the ‘peripheral’ country essentially arises as a cost of extracting information about one’s marginal contribution to the *grand* coalition. Along these lines, any unilateral demand over the ideological make-up of a proposed union is construed as a credible signal of a country’s bargaining leverage, namely how much it stands to gain and contribute in terms of added value from cooperating with any subset of its potential coalition partners. The dynamic interaction between France, Germany, and the UK in the early days of the EEC constitutes the interesting case at hand and will be analyzed later on.

Let  $i = A, B, C$  denote countries as unitary actors bargaining over the creation of a union at specific policy terms  $x_c \in X$ , where  $c$  represents any non-empty union subset of  $N = \{A, B, C\}$ , i.e.,  $c \in 2^N \setminus \emptyset$ , and  $X \subseteq \mathfrak{R}$  denotes the single policy (or ideological) dimension over which bargaining takes place. To avoid confusion, I denote *coalition structures*  $\pi_t \in \Pi$  at time  $t=0,1$  by  $c(x_c) | \{Nc\}$ ,  $|c| \geq 2$ , where  $c$  denotes a bilateral or trilateral union with common policy  $x_c$ , otherwise  $A|B|C$  denotes the fully autarchic coalition structure. I introduce preference heterogeneity over policy by assuming distinct country-specific ideal points on the real line  $m_A < m_B < m_C \in X$ , where the letter  $m$  denotes the bliss point of the median voter in each country. As long as any subset of countries agree to coordinate on a common policy  $x_c \in X$ , then each member of that coalition (or political union) reaps the common coalition benefits  $y_c \geq 0$ ,  $c \in 2^N \setminus \emptyset$ . Note that the assumption of *weak superadditivity* guarantees that  $y_c \leq y_{c'}$ ,  $c \subseteq c' \in 2^N \setminus \emptyset$ , i.e., the *grand* coalition  $ABC$  is weakly efficient and any bilateral union of countries may not generate strictly higher ‘gains from trade’. Moreover, I do not allow for any *policy externalities* across countries; hence, autarchy yields no exogenous benefits per se, i.e.,  $y_i = 0$ ,  $y_i = A, B, C$ .<sup>10</sup> It remains the case that in a

<sup>10</sup> Since countries are assumed to be symmetric with respect to economic and population size, this is a reasonable assumption. Making the autarchic reservation utility different across players would not affect the structure of the equilibrium, just their relative bargaining leverage and, in consequence, their final payoffs.



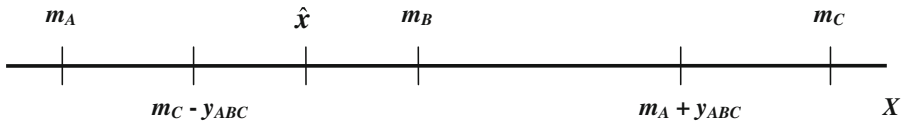


Fig. 1 Spatial configuration of equilibrium policy proposals

subgame-perfect Nash equilibrium *autarchic* policy coincides with the median ideal point in each country.<sup>11</sup> To keep things simple, for any country  $i$  belonging to a coalition  $c$  with common policy  $x_c$  utility takes the linear absolute deviation form  $u_t(m_i; x_c; y_c) = -|x_c - m_i| + y_c, i \in c, i = A, B, C$ <sup>12</sup> in each period  $t = 0, 1$ .

I further postulate the following set of technical assumptions with respect to the spatial configuration of median ideal points and the *worth* of the *grand* coalition  $ABC$  (see Fig. 1):

**Assumption 2**

$$m_C \geq m_B + (m_B - m_A)$$

This assumption essentially implies that preference-wise the moderate country  $B$  is closer to  $A$  than  $C$  and, thus,  $A$  and  $B$  are the ‘natural partners’ in any pairwise coalition, while  $C$  is spatially ‘peripheral’ to the other two. By ‘natural partners’ I wish to denote a subset of countries that lie closer in terms of institutional structures, historical traditions, and economic fundamentals, which all together make for enhanced ideological contiguity in terms of policy preferences.

**Assumption 3**

$$m_C - m_A \geq y_{ABC} \geq m_C - \frac{m_A + m_B}{2} \left( > \frac{2}{3} (m_C - m_A) \right)$$

According to assumption 3, the exogenous benefits associated with the *grand* union  $ABC$  are at such a level as to allow for meaningful policy negotiations over the ideological direction of the trilateral union but not high enough to make everyone’s participation constraints trivially binding. In light of the linear structure of the model,

<sup>11</sup> Under autarchy, each country retains full sovereignty over the determination of its own domestic policy by democratic means. Assuming single-peakedness, the ideal point of the median country representative  $m_i$  is the only *Condorcet winner* in any pairwise election (or referendum). The *no-policy-externalities* assumption implies that the autarchic, ‘go-it-alone’ payoff of a country does not depend on the entire coalition structure, i.e., whether the other two players coalesce or not. See Maskin (2004) for an axiomatic extension of the *Shapley value* solution concept allowing for *coalition externalities* and *partition functions*.

<sup>12</sup> Note that, by using this simple linear functional form, I effectively impose risk-neutrality; therefore, risk aversion will not factor into the results. The common exogenous benefit of integration  $y_c$  enters utility as an additive component that is orthogonal to the ideological policy dimension. This essentially represents the *economies of scale* property of policy centralization within international unions.

the imposed bounds serve the purpose of ruling out corner solutions and simplifying the structure of the equilibrium.<sup>13</sup>

As for the bargaining protocol of the game, there are two successive rounds of negotiations, whereby each country representative is randomly recognized with equal probability in each period regardless of the pre-existing coalition structure.<sup>14</sup> Extending the dynamic form of the game to an infinite-horizon bargaining framework would add little insight for our purposes, at the expense of equilibrium multiplicity and analytical complexity. Once recognized at time  $t=0,1$ , the representative of country  $i$  makes an *unconditional* common policy proposal  $x_t^i \in X$ , for a given coalition structure  $\pi_{t-1}$ , which may be accepted or rejected by the other countries  $j, k \neq i$ . This is just a technical assumption—typical in the coalition theory literature—which does not allow for the outright exclusion of a player from a proposed coalition.<sup>15</sup> Then, payoffs accrue at the end of *each* period  $t$  as a function of the resulting coalition structure  $\pi_t$ .

Let  $x_t^i(\pi_t | y_{ij}, y_{ik}) : \Pi \rightarrow X$  and  $\alpha_t^j(x_t^i | \pi_t, y_{ij}, y_{jk}) : X \rightarrow \{0, 1\}$ ,  $i \neq j \neq k$  denote the pure *proposal* and *acceptance* strategies respectively for each player-type and time period. An international union with common policy  $x_t^i \in X$  will form at time  $t$  only between the proposing country and those who accept that proposal. Existing union policy may only be renegotiated subject to the unanimous approval of all its members. I further assume excessively high fixed costs of union disintegration, effectively implying that once created, a union partnership may not dissolve or equivalently that a union member may not unilaterally withdraw.<sup>16</sup> An existing union may only amend its common policy and expand its membership to the excluded country as long as all of its participating members agree to the proposed policy terms.<sup>17</sup> The postulate of veto power in enlargement negotiations is avowedly

<sup>13</sup> For example, the lower bound on the set of permissible values for  $y_{ABC}$  implies that extreme countries  $A$  and  $C$  may only form a bilateral union with moderate country  $B$  at their ideal points in a subgame-perfect Nash equilibrium, thus ruling out subgames of partial coalition-formation at intermediate policy positions.

<sup>14</sup> I, hence, abstract away from the possibility that participation in a union at an earlier time enhances one's proposal prerogative at subsequent enlargement negotiations. This could make for an alternative rationale for gradualism, by incentivizing early participation through enhanced bargaining leverage in subsequent enlargement negotiations. Roberts (1999) presents a related dynamic model of clubs with endogenous membership.

<sup>15</sup> Note that this term should not be mistaken as *unconditional* on the past history of play (which will certainly not be the case in the proposed equilibrium). What it basically means is that a proposal may not be extended only to a specific proposed coalition, but should be available to all players. This is just a technical assumption that simplifies the proposal strategy set to  $X$ , rather than  $2^N \times X$  as in the case of offers *conditional* on the proposed coalition  $c \in 2^N \setminus \emptyset$ . With *conditional* offers, proposals get implemented *if and only if* all parties of the proposed coalition concur. Gomes (1999) teases out the implications of allowing both *conditional* and *unconditional* offers for efficiency and concludes that inefficient delay can be significant in superadditive games with small discount factors, once *unconditional* offers have been ruled out.

<sup>16</sup> This assumption is closely related to Seidmann and Winter's (1998) concept of *irreversibility* of coalition agreements, which once agreed upon become immediately enforced allowing the contracting parties to reap the related payoffs. Their implicit assumption is that the fixed costs of divesting a coalition agreement are so high that the latter becomes an enforceable outside option in the bargaining process.

<sup>17</sup> This assumption essentially reflects the unanimity requirement for EU enlargement and treaty amendment.

one of the driving forces of the main result. Finally, the future is discounted at a common rate  $\delta \in [0, 1]$ .

To recapitulate the structure of the game, I present its timing in bullet form:

- Initially, countries  $A$ ,  $B$ , and  $C$  reside in an autarchic *status quo ante*  $\pi_{-1} = \{A|B|C\}$ , i.e., policy is set at the domestic level and no coalitions have formed.
- At time  $t=0$  (first period), the representative of each country gets randomly recognized as a proposer with equal probability. Once recognized, the agenda-setting country  $i$  will make an *unconditional* policy proposal  $x_0^i \in X$ , which the remaining two players simultaneously choose whether to accept or reject.
- A first-period coalition  $c_0 \subseteq \{A, B, C\}$  forms between the proposer and whoever else chooses to accept the initial proposal and utility payoffs accrue at the end of the period as a function of the *interim* coalition structure  $\pi_0$ . If both reject, then policy is set at the domestic level and players receive their autarchic payoffs.
- At time  $t=1$  (second period), a proposer  $j$  is once again randomly recognized regardless of the coalition structure and the identity of the previous proposer  $i$  and makes a policy proposal  $x_1^j \in X$  to the other two players.
- An existing two-country coalition carried over from the first period may not dissolve and it may only expand at the proposed policy  $x_1^j \in X$  with the unanimous consent of its existing members. The members of an existing coalition  $c_0$  may also jointly agree to amend their common supranational policy  $x_{c_0} \in X$  at time  $t=1$ .
- Finally, the players reap the benefits pertaining to the coalition structure ensuing at the end of the second period and the game ends.

Before I proceed to describe the equilibrium, I introduce the final piece of notation:

**Notation** Let  $\langle \pi_{t=0}, \pi_{t=1} \rangle$ , where  $\forall c \in \pi_{t=0}$  and  $c' \in \pi_{t=1}$ ,  $c' \not\subset c$  (no-union-dissolution assumption), denote a two-period coalition-formation path.

**Definition 1** A coalition-formation path  $\langle \pi_{t=0}, \pi_{t=1} \rangle$  is called *comprehensive* if and only if the *grand* coalition  $ABC$  forms in either period 0 or 1, otherwise the coalition-formation process is *partial*.

**Definition 2** A coalition-formation path  $\langle \pi_{t=0}, \pi_{t=1} \rangle$  is called *gradual* if and only if there exist  $c \in \pi_{t=0}$ ,  $c' \in \pi_{t=1}$  such that  $c \subset c'$ . Otherwise, it is *immediate*.<sup>18</sup>

### 3.2 Two-Period Bargaining Game with Complete Information

I first look at the baseline version of the model with complete information about the *value* of each and every potential coalition. In a simple application of backwards induction to this two-stage bargaining game, the main result of this section is that a *gradual* coalition-formation path will never arise in any *subgame-perfect Nash*

<sup>18</sup> Seidmann and Winter (1998) offer an excellent theoretical account of the concepts of immediate vs. gradual as well as partial vs. comprehensive coalition formation.

*equilibrium* for any value of the discount factor  $\delta$  and the set of coalition benefits  $y_c$ ,  $\forall c \in 2^N \setminus \emptyset$ . Perfect foresight, intertemporal discounting, and uniform benefits of integration do not allow for equilibria of *gradual* integration and union enlargement, whereby an initially excluded candidate-country is invited to join an existing two-country union in the second period of the model.

**Proposition 1** *For any  $\delta \in [0, 1]$  and  $y_c \geq 0$ ,  $\forall c \in 2^N \setminus \emptyset$  subject to assumptions 1, 2, and 3 there is no gradual coalition-formation path in any subgame-perfect Nash equilibrium of the above game with complete information.<sup>19</sup>*

*Proof* See Appendix.

Coalition formation can only be *immediate*; whether is *partial* or *comprehensive* depends on who gets to propose, the specific coalition values, and the discount factor. The basic intuition is that if either of the extreme countries  $A$  or  $C$  happens to face an initial policy proposal, which it finds unacceptable (vis-à-vis the autarchic *status quo*) regardless of the other's response, then the interval of unanimously acceptable policy proposals in the second period will effectively disappear given the *interim* coalition structure and utility payoffs. If either  $A$  or  $C$  are initially recognized at  $t=0$ , then, depending on  $\delta$  and  $y_c \geq 0$ ,  $\forall c \in 2^N \setminus \emptyset$ , they will either make an accommodating policy proposal such that the other two are just willing to join the *grand* union  $ABC$  or a more extreme proposal (which following assumption 3 will coincide with their ideal point) such that only moderate country  $B$  is just willing to accept regardless of the third country's decision.<sup>20</sup> For all high types  $y_{AB}$ ,  $y_{BC} \in (2y_{ABC} - (m_C - m_A), y_{ABC}]$ , either of the extreme countries  $A$  or  $C$  would be better off in a *partial* coalition-formation outcome  $AB(m_A)|C$  or  $A|BC(m_C)$  respectively, since that would be preferable even to a *grand* coalition with common policy as close as possible to their ideal points, i.e.,  $ABC(m_C - y_{ABC})$  or  $ABC(m_A + y_{ABC})$  respectively. Otherwise, for low types, there always exists a Pareto superior common policy such that *immediate grand* coalition formation is unanimously preferred to the proposer's optimal *gradual* coalition-formation path. On the other hand, if  $B$  gets to propose first, then the *grand* coalition will always form *immediately*.

Getting entrenched within a *partial* coalition status quo will never profitably enhance one's bargaining leverage with respect to the excluded country, for the simple reason that the gain from a restricted *grand* coalition Pareto set is less than the cost of delay even for  $\delta$  close to unity. Also note that the fact that  $ABC$  is weakly efficient vis-à-vis the autarchic *status quo* does not imply that it necessarily comes about; unlike traditional coalition-formation models, the inclusion of a spatial bargaining dimension gives rise to the possibility of *partial* coalition formation.

<sup>19</sup> Note that there is multiplicity of *subgame-perfect Nash equilibria* given that there can be more than one Nash equilibria in weakly undominated strategies in several subgames, where (*no*, *no*) and (*yes*, *yes*) can both be mutual best responses to an existing proposal. Credible out-of-equilibrium threats to coordinate on a Pareto suboptimal autarchic outcome may condition the optimal proposal strategy at each stage of the game.

<sup>20</sup> Since proposal strategies only consist of *unconditional* policy proposals not directed to specific coalition partners, it would be trivial to show that coalition  $AC$  may never form in a *partial subgame-perfect Nash equilibrium*, since by necessity moderate country  $B$  would also want to join.

### 3.3 Two-Period Bargaining Game with Asymmetric Information

Having proven that gradualism may never be part of any subgame-perfect equilibrium of the baseline model, I relax the assumption of complete information, in order to show that private information is a precondition for gradualism in the coalition-formation process. I introduce uncertainty in the model by assuming that the representatives of each country only have prior knowledge of the ‘synergistic benefits’ generated by a non-autarchic union in which they participate; in other words, the members of any given union share private information over the *worth* of their coalition.<sup>21</sup> Thus,  $A$  knows the *value* of non-autarchic coalitions  $AB$ ,  $AC$ , and  $ABC$ , and similarly for  $B$  and  $C$ , while symmetric autarchic benefits  $y_A=y_B=y_C=0$  are common knowledge in the game. Otherwise, any excluded partner holds uniform prior beliefs over the common benefits inherent in a union between the other two countries, i.e., for any country  $k \neq i, j$ ,  $y_{ij} \sim U_k[0, y_{ABC}]$ , where  $y_{ABC}$  denotes the total *worth* of the *grand* coalition.<sup>22</sup> For the purposes of the model, I assume independence of *partial* coalition values (even though correlated values would not significantly affect the model). As a result, players cannot deduce the benefits of a bilateral union in which they do not participate from their own private information.

I now focus on the *perfect Bayesian Nash equilibrium* in pure weakly undominated strategies of the sequential bargaining game with incomplete information to demonstrate how gradualism in the formation of the *grand* union  $ABC$  may arise as a consequence of *asymmetric information*. The equilibrium consists of a set of *proposal* and *acceptance* strategies for each player and time-period, namely  $\langle x_t^{i*}(\pi_t | y_{ij}, y_{ik}, \sigma^{i*}(y_{jk})), \alpha_t^{i*}(x_t^{i*} | \pi_t, y_{ij}, y_{ik}, \sigma^{i*}(y_{jk})) \rangle$ ,  $\forall i \neq j, k$  and  $t=0,1$ , and a set of beliefs  $\sigma_t^{i*}(y_{jk} | h_t)$ ,  $\forall i \neq j, k$  given the history  $h_t$  of actions at time  $t$ , such that (1)  $\langle x^*, \alpha^* \rangle$  are *sequentially rational*, i.e., mutual best-responses for any profile of types  $(y_{ij}, y_{ik}, y_{jk})$  and for given equilibrium beliefs (*subgame perfection*), and (2) beliefs  $\sigma_t^{i*}(y_{jk} | h_t)$  are consistent with Bayes’ rule given the history of actions  $h_t$  along the equilibrium path. For reasons of tractability, I focus on equilibria without credible off-the-equilibrium threats to coordinate on Pareto suboptimal *rejection* outcomes.

Each country  $i$ ’s first-period proposal  $x_t^i$  essentially functions as a signal of its type  $y_{ij} \in [0, y_{ABC}]$  in a potential bilateral union agreement with another country  $j$ . A player’s type captures its resolve to wait out the formation of the *grand* union at more favorable policy terms. Taking country  $A$  as an example, high  $y_{AB}$  types will want to credibly signal their strong type by proposing a transitory bilateral  $AB$  subunion agreement at time 0, in order to entrench their position within a beneficial *interim status quo* and, thus, achieve greater bargaining leverage in subsequent policy negotiations at time  $t=1$ . In that case, the excluded party  $C$  will recognize that

<sup>21</sup> The synergistic benefits of integration are often hard to observe and quantify; hence, it would be plausible to assume that governments are willing to invest in technocratic expertise (expert commissions, technical reports) only with regard to international policy arrangements that affect them directly.

<sup>22</sup> An equivalent interpretation of this type of private information is that the representative of any country  $k$  does not know with certainty the level of his/her country’s *marginal contribution*  $y_{ABC}-y_{ij}$ ,  $k \neq i, j$  to the *grand* coalition. To make a clarifying analogy, think of the level of a country’s *marginal contribution* to the *grand* coalition as a poker hand; then this formulation of the game is tantamount to a game of poker where the players hold their cards against their forehead so that everyone else can see them but themselves.

only high  $y_{AB}$  types would find it in their interest to incur the cost of *strategic delay*, in order to induce a better  $x_{ABC}$  proposal at time  $t=1$ . The more extreme (relative to  $C$ ) of the ‘core’ countries, i.e., the one that has least to gain, will hold the enlargement process hostage, in order to achieve the best possible deal in the formation of the *grand coalition ABC*.

I now proceed to formally demonstrate the workings of this signaling mechanism of *strategic delay*. I first state the equilibrium in Proposition 2 and then go on to characterize it. Formal derivations can be found in the [Appendix](#).

**Proposition 2** *In the two-period bargaining game with asymmetric information there exists a perfect Bayesian Nash equilibrium, whereby the representative of either of the extreme countries  $i=A$  or  $C$  will propose  $x_0^{i*} = m_i$  at time  $t=0$  if and only if  $y_{iB} \in (\tilde{y}_{iB}, y_{ABC}]$  for some  $\tilde{y}_{AB} \in [y_{ABC} - (m_B - m_A), 2y_{ABC} - (m_C - m_A)]$  or  $\tilde{y}_{BC} \in [y_{ABC} - (m_C - m_B), 2y_{ABC} - (m_C - m_A)]$ , in which case coalition  $AB$  (or  $BC$  respectively) will form right away and may later expand to the grand coalition  $ABC$  at time  $t=1$  with strictly positive probability. Otherwise, for  $y_{iB} \in [0, \tilde{y}_{iB}]$ ,  $i=A$  or  $C$  will propose  $x_0^{i*} = \hat{x}^j(\delta)$  (see Fig. 2) such that  $j \neq i$ ,  $j=A, C$  is just indifferent between accepting and rejecting at time  $t=0$ , in which case the grand coalition  $ABC$  will form immediately. If median country  $B^{B*}$  gets to propose first, then all its types will pool on an equilibrium proposal  $x_0 = m_B$ , which will lead to the immediate formation of the grand coalition.*

In other words, gradualism in union formation will arise with positive probability for a non-degenerate support of high bilateral union types  $y_{iB} > \tilde{y}_{iB}$ . Otherwise, the extreme country ( $A$  or  $C$ ) finds it too costly to delay the *immediate* formation of the weakly efficient *grand coalition*. The following analysis derives equilibrium strategies starting from the second period for all subgames starting with  $A$  as the first-period proposer, while the [Appendix](#) completes the characterization of the equilibrium and contains the exact derivation of optimal compromising first-period proposals  $x_0^{i*} = \hat{x}^j(\delta)$  and equilibrium threshold types  $y_{iB} \in [0, \tilde{y}_{iB}]$ , where  $i = A$  or  $C$  and  $j \neq i$ .

Solving this sequential bargaining game backwards, let us first examine the proposal subgames in the *second* period, starting with an existing *partial* integration coalition structure  $AB(m_A)|C$ .<sup>23</sup> Depending on its type, once recognized, the representative of country  $A$  will either move to propose his/her ideal point for high enough values of  $y_{AB}$ , in order to preserve the existing *status quo* coalition structure  $AB(m_A)|C$ , or will otherwise propose  $m_C - y_{ABC}$ , which makes  $C$ ’s participation constraint just binding, thus leading to the formation of  $ABC(m_C - y_{ABC})$  as the final outcome. Formally,  $A$ ’s optimal proposal strategy at  $t=1$  in this subgame is the following:

$$x_1^{A*} = \begin{cases} m_A, & y_{AB} \in (2y_{ABC} - (m_C - m_A), y_{ABC}] \\ m_C - y_{ABC}, & y_{AB} \in [0, 2y_{ABC} - (m_C - m_A)] \end{cases}$$

<sup>23</sup> The analysis of the ‘mirror image’ *partial* integration subgame with existing coalition structure  $A|BC(m_C)$  is very similar and is, hence, omitted.

In equilibrium,  $B$  will accept any  $x_1^A \in [m_A - (y_{ABC} - y_{AB}), 2m_B - m_A + (y_{ABC} - y_{AB})]$  if and only if  $C$  accepts too; otherwise,  $B$  will only accept an amended *status quo* bilateral coalition  $AB$  such that  $x_1^A = x_{AB} \in [m_A, 2m_B - m_A]$  conditional on  $C$ 's rejection. Finally, it is a weakly dominant strategy for  $C$  to accept any  $x_1^A \in [m_C - y_{ABC}, m_C + y_{ABC}]$  regardless of  $B$ 's response, i.e., regardless of whether the *grand* coalition actually materializes or not.<sup>24</sup>

Country  $B$  in the *second* period will choose to preserve the *status quo* coalition structure so long as there is no *grand* coalition  $A$  will agree to; otherwise,  $B$  will seek to bring the *grand* coalition  $ABC$  together at the most favorable terms possible subject to the approval of the other two negotiating parties. Hence,  $B$ 's optimal second-period proposal strategy is as follows (the last column indicates the ensuing coalition structure):

$$x_1^{B*} = \begin{cases} m_A, & y_{AB} \in (2y_{ABC} - (m_C - m_A), y_{ABC}] & AB(m_A)|C \\ m_A + y_{ABC} - y_{AB}, & y_{AB} \in (y_{ABC} - (m_B - m_A), 2y_{ABC} - (m_C - m_A)] & ABC(m_A + y_{ABC} - y_{AB}) \\ m_B, & y_{AB} \in [0, y_{ABC} - (m_B - m_A)] & ABC(m_B) \end{cases}$$

Finally, once recognized in the *second* period, the representative of country  $C$ , which was excluded from the initial union at time  $t=0$ , will seek to maximize its expected returns from participating in the *grand* coalition  $ABC$  based on its beliefs about the existing union's orthogonal benefits  $y_{AB}$ . According to the proposed semi-separating equilibrium,  $C$ 's updated Bayesian beliefs following  $x_1^A = m_A$  will be such that  $y_{AB} \sim U(\tilde{y}_{AB}, y_{ABC})$ ; therefore, its optimal proposal strategy would be to maximize its expected payoff from participating in the *grand* coalition  $ABC$  subject to the approval of country  $A$ , i.e.,

$$\begin{aligned} x_1^{C*} &= \arg \max_{x_1^C \in [m_C - y_{ABC}, m_C]} \left\{ \text{Prob}(\alpha_1^{A*}(x_1^C) = 1 | y_{AB} \sim U(\tilde{y}_{AB}, y_{ABC})) \times (x_1^C - m_C + y_{ABC}) \right\} = \\ &= \arg \max_{x_1^C \in [m_C - y_{ABC}, m_C]} \left\{ \text{Prob}(y_{AB} \leq m_A - x_1^C + y_{ABC} | y_{AB} \sim U(\tilde{y}_{AB}, y_{ABC})) \times (x_1^C - m_C + y_{ABC}) \right\} = \\ &= \frac{m_A + m_C - \tilde{y}_{AB}}{2} \end{aligned} \tag{1}$$

Of course, there is a positive support of types  $y_{AB} \in \left( y_{ABC} - \frac{(m_C - m_A) - \tilde{y}_{AB}}{2}, y_{ABC} \right]$  that will reject  $C$ 's proposal at time  $t=1$ , thus giving  $C$  its autarchic equilibrium utility of 0 and generating the *ex post* inefficient outcome of *partial* coalition formation.

Now let the coalition structure at time  $t=1$  be  $A|B|C$ , namely the fully autarchic *status quo*. On the assumption that the responders will not coordinate on a Pareto

<sup>24</sup> Note that a condition for the existence of the proposed equilibrium in pure strategies is that any country will accept a policy proposal if indifferent between accepting and rejecting. Moreover, in equilibrium, any country will always opt for the larger coalition if indifferent between coalitions of different size.

suboptimal Nash equilibrium of  $(no, no)$ , the optimal proposal strategies in this second-period subgame are as follows:

$$\begin{aligned}
 x_1^{A*} &= \begin{cases} m_A, & y_{AB} \in (2y_{ABC} - (m_C - m_A), y_{ABC}] & AB(m_A)|C \\ m_C - y_{ABC}, & y_{AB} \in [0, 2y_{ABC} - (m_C - m_A)] & ABC(m_C - y_{ABC}) \end{cases}, \\
 x_1^{B*} &= m_B, \forall y_{AB}, y_{BC} & ABC(m_B), \text{ and} \\
 x_1^{C*} &= \begin{cases} m_C, & y_{BC} \in (2y_{ABC} - (m_C - m_A), y_{ABC}] & A|BC(m_C) \\ m_A + y_{ABC}, & y_{BC} \in [0, 2y_{ABC} - (m_C - m_A)] & ABC(m_A + y_{ABC}) \end{cases}.
 \end{aligned}$$

Finally, once the *grand* coalition  $ABC$  has formed in the first period, its common policy may only be amended at time  $t=1$  if and only if  $x_{ABC} \notin [m_A, m_C]$ , i.e., if and only if its existing common policy lies outside of the unanimity *core*.

Reasoning backwards to the bargaining game in the *first* period, there are only two uniform non-empty supports of types  $y_{AB}$  that country  $A$  may credibly signal to country  $C$  in a semi-separating equilibrium:<sup>25</sup> low types  $y_{AB} \in [0, \tilde{y}_{AB}]$ , who will want to participate in the weakly efficient union  $ABC$  as soon as possible, and high types  $y_{AB} \in (\tilde{y}_{AB}, y_{ABC}]$ , whose transitory reservation utility in a *gradualist* equilibrium is high enough to justify *strategic delay* with the aim of eliciting a more favorable policy proposal from  $C$  at  $t=1$ . Hence, there are only two equilibrium common policy proposals  $x_0^{A*} \in X$  at  $t=0$ : high types in favor of gradualism will seek to maximize their transitory reservation utility stemming from a subcoalition  $AB$  by proposing their ideal point  $m_A$ , which country  $B$  will unconditionally accept regardless of  $C$ 's response and  $C$  will obviously reject in favor of setting its own autarchic policy. Low compromising types, on the other hand, will immediately propose a common policy  $x_0^{A*} = \hat{x}^C(\delta) \in [m_C - y_{ABC}, m_B]$ ,<sup>26</sup> such that  $C$  would be just willing to join the *grand* coalition  $ABC$  right away, instead of waiting for a potentially more favorable deal at a later time, i.e.,

$$\hat{x}^C(\delta) = \inf \{x_0^A \in [m_C - y_{ABC}, m_B] : V_{t=0}^C(\alpha_0^C(x_0^A) = 1 | \alpha_0^B(x_0^A) = 1) = V_{t=0}^C(\alpha_0^C(x_0^A) = 0 | \alpha_0^B(x_0^A) = 1)\},$$

where  $V$ 's denote first-period continuation values. In order to derive the compromising first-period proposal  $\hat{x}^C(\delta)$ , we need the following lemma:

**Lemma 1** *There exists for country C a threshold point on the real line  $\tilde{x}_0^C(\delta) \in [m_C - y_{ABC}, m_B]$ , such that for  $x_0^A < \tilde{x}_0^C(\delta)$  and  $\pi_0 = \{AB(x_0^A)|C\}$  a positive support of low types  $y_{AB} \leq \tilde{y}_{AB}$  will reject C's optimal second-period proposal  $x_1^{C*}$  given its equilibrium beliefs  $\sigma^{C*}(y_{AB}|x_0^A)$ , while for  $x_0^A \geq \tilde{x}_0^C(\delta)$  all low types  $y_{AB} < \tilde{y}_{AB}$  will accept  $x_1^{C*}$ . A similar threshold point  $\tilde{x}_0^A(\delta)$  exists for country A.*

<sup>25</sup> The number of distinct policy proposals that may be elicited in this *perfect Bayesian equilibrium* is limited by the number of players and potential subcoalitions.

<sup>26</sup> It is clear from the above analysis that  $\hat{x}^C(\delta) \geq m_C - y_{ABC}$ , otherwise for  $m_A \leq \hat{x}^C(\delta) < m_C - y_{ABC}$   $C$  would never want to participate in a *grand* union  $ABC$  whose ideological make-up is so far skewed to the left that it yields less than its reservation utility of 0 throughout both periods (since it would not be able to amend it at time  $t=1$ ).



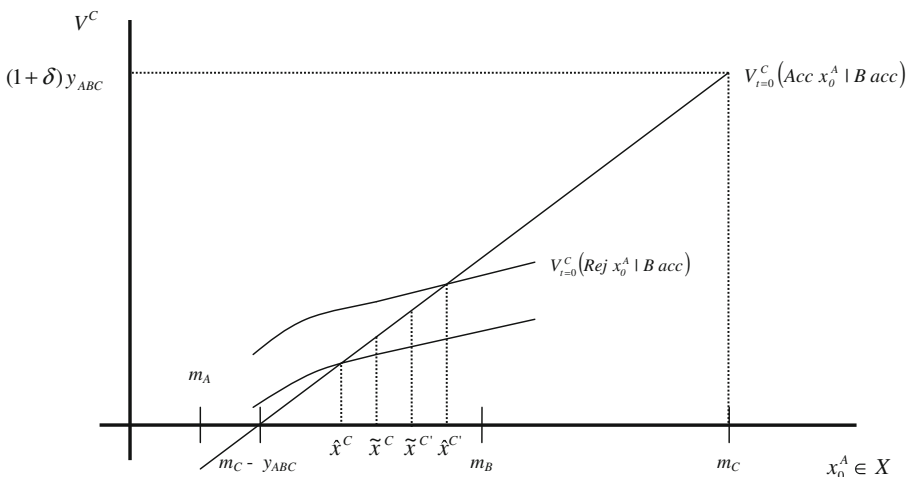
*Proof* See Appendix.

Figure 2 provides a graphical demonstration of how the location of  $\hat{x}^C(\delta)$  with respect to  $\tilde{x}^C(\delta)$  depends on the value of the discount factor  $\delta \in [0, 1]$ . It depicts  $C$ 's first-period *acceptance* and *rejection* continuation values  $V_{t=0}^C$  as a function of  $x_0^A$ .  $A$ 's optimal compromising first-period proposal lies at the intersection of these two lines. The value of the discount factor  $\delta$  will affect the location of the continuation value curves and the threshold point; it will, therefore, determine whether this point of intersection occurs above or below  $\tilde{x}^C(\delta)$ . If it occurs at a point  $\hat{x}^C(\delta) < \tilde{x}^C(\delta)$ , then it lies on the quadratic part of  $C$ 's *rejection* continuation value. Otherwise, if  $\hat{x}^C(\delta') \geq \tilde{x}^C(\delta')$ , then the point of intersection lies on the linear part of  $C$ 's *rejection* continuation value. Note that the continuation value curves for  $\delta' \neq \delta$  have been omitted from the graph for clarity of exposition.

In order to complete the characterization of the semi-separating equilibrium starting with the representative of country  $A$  as the first-period proposer, I also need to specify the out-of-equilibrium beliefs for  $C$  in the following manner:

$$\sigma^{C*}(y_{AB}|x_0^A) \rightarrow \begin{cases} y_{AB} \sim U[0, \tilde{y}_{AB}], & \text{for } x_0^A \in [m_C - y_{ABC}, \hat{x}^C] \cup (\hat{x}^C, +\infty) \\ y_{AB} \sim U(\tilde{y}_{AB}, y_{ABC}) & \text{for } x_0^A \in (-\infty, m_A) \cup (m_A, m_C - y_{ABC}) \end{cases}$$

where  $\sigma^{C*}(y_{AB}|x_0^A)$  denotes  $C$ 's updated beliefs about  $A$  and  $B$ 's reservation utility. In light of  $A$ 's perfect Bayesian equilibrium proposal strategy,  $C$  will reason that only non-compromising, high types would ever make an initial policy proposal outside of  $C$ 's second-period *grand* coalition acceptance interval, i.e., strictly less than  $m_C -$



**Fig. 2** Graphical determination of  $\hat{x}(\delta)$

$y_{ABC}$ , since  $C$  would never accept such a proposal regardless of  $B$ 's response.<sup>27</sup> On the other hand, only compromising, low types seeking the *immediate* formation of the *grand* coalition would ever make an offer within that interval. Given that the above line of reasoning would be common knowledge,  $A$ 's prescribed strategy would indeed be optimal in equilibrium.

As it turns out, in light of the assumptions on the efficiency of the *grand* coalition and the ideological configuration of the three players, the *gradualist* equilibrium of *strategic delay* may only come about once either of the more extreme countries gets to propose first. In effect, agenda-setting power is not imposed *ex ante* as an exogenous prerogative but is rather rationalized *de facto* through the *gradualist* equilibrium of the model. In other words, the proposal order determines the order of entry into the union, which is crucial for the determination of its shape and ideological character over time. It would be relatively straightforward to show that were the representative of country  $B$  to be recognized as a proposer in the first period, all of its types would necessarily pool on the same policy proposal by dint of its median ideological position, thus attracting both of its potential partners to the *immediate* formation of the *grand* union  $ABC$ . As the model has been set up,  $B$  would have no incentive to instigate an inefficient *gradual* coalition-formation process.

### 3.4 Welfare Analysis

Having explained how gradualism may arise within the context of a perfect Bayesian equilibrium, it is now incumbent to examine the efficiency properties of this equilibrium taking *without loss of generality* the representative of country  $A$  as the first-period proposer. From an *interim* efficiency perspective, the question arises whether there is an *immediate grand* coalition formation path  $\langle ABC(x_{ABC}), ABC(x_{ABC}) \rangle$  that makes everyone better off in comparison with the *gradual* equilibrium path for high  $y_{AB}$  types, i.e.,  $\langle AB(m_A)|C, ABC(x_1^{j*}) \rangle$ .

**Proposition 3** *For a non-degenerate support of high  $y_{AB}$  types  $(\tilde{y}_{AB}, \tilde{y}_{AB} + \varepsilon]$ ,  $\varepsilon > 0$ , there exists an immediate grand coalition formation policy  $x_{ABC}^* \in [m_C - y_{ABC}, m_B]$  such that  $\langle ABC(x_{ABC}^*), ABC(x_{ABC}^*) \rangle \succeq_i \langle AB(m_A)|C, ABC(x_1^{j*}) \rangle$  for all  $i, j = A, B, C$  and  $\langle ABC(x_{ABC}^*), ABC(x_{ABC}^*) \rangle \succ_i \langle AB(m_A)|C, ABC(x_1^{j*}) \rangle$  for at least one  $i$ , where  $\succeq$  and  $\succ$  denote the weak and strict preference relations respectively and  $x_1^{j*}$  denotes the optimal second-period proposal for any representative  $j$  to get recognized.*

<sup>27</sup> Note that, despite the fact that country  $B$  shares the same private information as  $A$ , I have omitted its response to  $A$ 's first-period proposal from the signaling mechanism, taking its acceptance for granted. In light of its moderate ideological position on the real line vis-à-vis the other two actors, the representative of country  $B$  would always accept a compromising *grand* union proposal  $x_0^{A*} = \hat{x}(\delta)$ . However, it would only accept the *gradualist* equilibrium proposal  $x_0^{A*} = m_A$  for  $m_C$  high enough and/or  $\delta$  low enough. In either case, its pooling acceptance strategy does not provide an informative signal to  $C$ . Otherwise,  $A$  would have to moderate its initial *gradualist* proposal to the extent that  $B$  is just indifferent between  $\langle AB(x_0^A)|C, ABC(x_{ABC}) \rangle$  and  $\langle A|B|C, ABC(x_{ABC}) \rangle$ .

*Proof* See [Appendix](#).

*Ex post* inefficiency, on the other hand, ensues whenever  $C$ 's optimal second-period proposal  $x_1^{C*}$  gets rejected by  $A$  at  $t=1$  (or conversely when  $x_1^{A*}$  gets rejected by  $C$ ), namely whenever the *grand* union  $ABC$  fails to materialize and the coalition formation process gets stalled in the *partial* coalition stage. In light of  $C$ 's non-degenerate beliefs about its partners' reservation utility, there is always a positive support of high types that will reject its second-period proposal and, hence, the partial union  $AB$  will fail to expand despite the existence of mutually beneficial enlargement policy deals. Note that, in the absence of *asymmetric information*, the unique equilibrium solution to the above two-period spatial bargaining game would prescribe *immediate* (*partial* or *comprehensive*) coalition formation and as such would necessarily be efficient. In other words, the presence of *asymmetric information* constitutes the real source of inefficiency by giving rise to the possibility of *gradual* and/or *partial* coalition formation.

#### 4 Case Study: UK Accession to the European Community<sup>28</sup>

The early evolution of the membership and institutional set-up of the European Economic Community provides an interesting illustration of the above mechanism of *strategic delay* in union formation.<sup>29</sup> Brought to life by the Treaty of Rome (signed and ratified by the founding members, namely France, Germany, Italy, The Netherlands, Belgium, and Luxembourg in 1957), the European Economic Community was a belated offspring of the European Coal and Steel Community (the ECSC was created by the Schuman Plan and the Treaty of Paris in 1950). The initial exclusion of the UK from the EEC led to the *de facto* division of Western Europe and gave rise to the core grouping of 'the Six' (also known as 'Little Europe'). On account of its heavy reliance upon its transatlantic and Commonwealth trade links, Britain did not show any interest in participating in such a novel economic and political integration project, especially if it had to cede much of its highly valued sovereignty over sensitive domestic policy areas. In the eyes of the British political establishment, the incipient EEC was a fragile and highly uncertain political experiment that could prove utterly unsuccessful and lead to the imminent collapse of the established supranational institutions.

All the more reason for British abstention from the Treaty of Rome was that the agreement was too politically ambitious and supranationalist in nature on account of the strong agenda-setting influence of the French on the institutional and political character of this new structure (which would suggest France as a first-period proposer in our model). By dint of its population and economic size, colonial power status, and its alignment with the victors in WWII,<sup>30</sup> France possessed the agenda-

<sup>28</sup> Following the Merger Treaty, which was signed on April 8, 1965 and came into force on July 1, 1967, the European Coal and Steel Community (ECSC), the European Atomic Energy Community (Euratom), and the European Economic Community (EEC) were consolidated into a single institutional structure and became collectively known as the European Communities (EC). The terms EC and EEC are, thus, used chronologically.

<sup>29</sup> For a concise literature review of approaches to the study of European integration see Hix (1994, 1998).

<sup>30</sup> By contrast, the devastating effects of the Second World War on Germany's economic infrastructure, international credibility, and morale dictated a more passive approach to restoring its international status through the medium of international political and economic cooperation.

setting initiative during the formative years of Western European reconciliation and integration.<sup>31</sup> Imbued by a high degree of supranationalism and bureaucratic *dirigisme*—a direct heir of the Monnet blueprint on the ECSC High Authority—, the EEC was both in character and design much more than an economic free trade association. As a concession primarily to the French, the Treaty of Rome, whose main agenda comprised the completion of the common market and subsequently the integration of agricultural markets, also allowed for the possibility of international cooperation in both the political and social spheres. As a result, the British perceived their national interests and preferences to lie elsewhere; in pursuit of an export outlet for their industrial production, they opted instead for the looser economic grouping of the European Free Trade Association (signed into existence by the Stockholm Convention in 1959),<sup>32</sup> whose economically diverse and geographically scattered membership also included Denmark, Norway, Sweden, Switzerland, Austria, and Portugal. The EEC, however, remained according to the Treaty of Rome open to all Western European countries, subject to certain political and economic eligibility requirements (later formalized as the Copenhagen criteria in 1993) and as long as they fully accepted the *acquis communautaire*.<sup>33</sup>

The first few years after the inception of the EEC, also known as the ‘honeymoon years’ (see Ludlow 1997, pp. 22–26), were marked by great success and dynamism in the creation of the common market and the dismantlement of intra-EEC barriers to trade, which in turn sent a clear signal of high ‘mutual gains’ among the founding members and strong interest in the continuation and expansion of cooperation. That signal combined with a gradual shift in its trade patterns towards the major economies of the EEC (and away from its Commonwealth partners) led to a radical change in British policy towards Europe under the Conservative Macmillan government. As a result, the UK (together with Denmark, Ireland, and Norway) applied for EEC membership in 1961.

The extensive negotiations that followed between the EEC-6 and the new candidates focused on (1) the harmonization of the latter’s domestic legislation with the extensive body of EEC legislation (collectively known as the *acquis communautaire*) through derogations and transitional periods, (2) their political weight in the supranational institutions, and (3) their financial contributions to the common budget. While the Dutch were the staunchest supporters of UK accession (given their strong political and economic links with the British), four of the other EEC members were also in favor each for their own reasons (see Ludlow 2006, Chapter 6). Only the French appeared lukewarm towards the prospect of EEC enlargement (especially in the case of Britain), albeit not opposed to it in principle. They therefore presented the British with the biggest obstacles towards membership.

<sup>31</sup> For an account of France’s leading role in the negotiations of the Treaty of Paris (1951), the Messina Conference (1955), and the Treaty of Rome, see Moravcsik (1998).

<sup>32</sup> See Gstöhl (2002) for an analysis of EC-EFTA relations.

<sup>33</sup> It should be noted that the Treaty of Rome did not specify any particular bargaining protocol for future accession negotiations other than that any enlargement decisions had to be agreed upon by unanimity (Art. 237).

Lest the negotiations be brought to a successful conclusion, General de Gaulle pronounced an effective veto upon British accession in a famous January 1963 press conference,<sup>34</sup> in an attempt to preempt a potential *yes* or *no* decision with respect to a fully negotiated but undesirable from the French point of view accession deal.

De Gaulle's unilateral and sudden decision to end accession negotiations by pronouncing 'Britain not ready for Europe' came to the *chagrin* of the 'Five' other EEC members and initially caused some disenchantment and malaise in the everyday workings of the Community, effectively slowing down the pace of integration. Even the Germans, who under Adenauer placed high political stakes in a French-German *rapprochement*<sup>35</sup> but were also eager to welcome a major trading partner and militarily powerful country like Britain into the Community, were negatively surprised by de Gaulle's actions; yet, apart from some ireful statements by politicians like Schroeder and Erhard, they chose not to confront France on the issue lest they jeopardize the heretofore achieved gains from cooperation and destabilize the internal institutional bargains of the EEC. Despite strong reactions by politicians of the other member-states, not one country chose to unilaterally withdraw from the Community in light of the high economic and political stakes at hand. A similar episode took place in 1967 during the second British application for EC accession under Wilson's Labor government, only this time the official excuse for de Gaulle's veto was the monetary instability of the sterling in light of its devaluation in 1967 (see Ludlow 2006, Chapter 5). However, in the aftermath of the second French veto, peace, unity, and integration momentum were unlikely to return to the Community until the enlargement controversy (also known as *la question anglaise*) had been addressed in a manner satisfactory to both applicants and member states.

Enlargement negotiations were finally reopened in the Hague Summit of 1969, concluded in 1971, and the UK, Denmark, and Ireland officially became the first EC expansion members in 1973 (Norway's accession was rejected by popular referendum). The French under their newly elected Gaullist president Georges Pompidou—but with the express approval of his predecessor and ideological kinsman de Gaulle—appeared much more accommodating during the enlargement process, having simultaneously achieved some much desired progress on the completion of the Community's initial agenda (*achèvement*) and the expansion of cooperation into new policy areas (*approfondissement*). The British, on the other hand, got a much worse deal than they would have in 1963, as it became apparent that the onus of adaptation to the *acquis* lay with the states wishing to join the EC, in light of the complexity of existing internal policy bargains. In the context of my

<sup>34</sup> Here is a translated excerpt from his press conference as quoted in Ludlow 1997, p. 207: "England is, indeed, insular and maritime, linked by her trade, her markets and her food supplies to diverse and often far-flung countries. She works primarily in industry and commerce, and hardly at all in agriculture. She has, in all her patterns of work, habits and traditions [...] which are highly distinctive and original. In short, the nature, the structure, the economic situation that characterize England, differ profoundly from the Continent. How then could England, as she lives, as she produces, as she trades, be incorporated into the Common Market as it was conceived and as it works?"

<sup>35</sup> As evidenced by the signing of the bilateral Elysée Treaty shortly after de Gaulle's press conference in January 1963, which remains until today a strong symbol of French-German *rapprochement* in the 1960s.

model, this bargaining outcome translates into a policy farther away from the expansion country's ideal point.<sup>36</sup>

The French vetoes of the British bids for EC membership in the 1960s<sup>37</sup> provide an interesting illustration of the above model, where France would be country *A*, Germany would be country *B*—making those two countries ‘natural partners’—, and the UK would be the ‘peripheral’ country *C*. The location of these three countries on the ideological line is justified by (1) the close economics ties between France and Germany, (2) the *détente* in their relations following the devastating effects of WWII, and (3) Britain's strong trade and political links with the USA and its Commonwealth partners. Restricting attention to the French–Anglo–German triptych of actors for the purposes of the case study does not detract from its accuracy and rigor considering the preponderant role of these three countries in the bargaining dynamics of the first enlargement. The UK was avowedly the most important and ‘controversial’ among the initial candidate-members (UK, Denmark, Ireland, and Norway),<sup>38</sup> while France was both on account of size and revealed preferences the unanimity *pivot* among existing EEC member states. Finally, Germany has to be singled out from the remaining EEC-5 as the ‘moderate’ country in our model by dint of its notable economic interests and its special political relationship with France (*rapprochement*). Firmly grounded on a rational choice perspective, one may interpret the above historical episode through the theoretical prism of *strategic delay* in the formation of the EC-9.

Viewing the process of European integration from 1957 (inception of the EEC by the Treaty of Rome) to 1973 (first expansion to EC-9) as a well-defined sequential game played by unitary forward-looking actors with stable preferences allows us to extrapolate the predictions of our theoretical model of gradualism to the unraveling of historical events. Evidence of domestic cross-partisan consensus among ruling parties in France, Germany, and Britain allows us to treat them as non-myopic strategic actors. The 1960s ushered in a period of trade expansion, high economic growth, and stability in the industrial West. The absence of any notable exogenous shocks to the global political and economic order (even though the collapse of the Bretton Woods system and the first oil crisis were just around the corner) affords us the analytical margin to zero in on endogenous strategic considerations as driving forces of the integration process during that time, rather than exogenously expanding ‘mutual gains’. In any event, the residual uncertainty with respect to the global

<sup>36</sup> Even though the *acquis* needs to be fully accepted and implemented by all new Community members, the single policy bargaining dimension in the model may be construed as the ‘institutional terms of accession’ to a union or, in other words, the degree of flexibility in the expansion country's adjustment to an existing body of legislation (e.g. through derogation clauses and transitional periods). This is why accession to an already formed and institutionalized union is not treated as a dichotomous decision to fully accept or reject the *acquis* as it is. Another way to rationalize the unidimensionality of the enlargement bargaining process is that it captures the degree of influence upon the character and institutional design of the enlarged union afforded to new members.

<sup>37</sup> See N. Piers Ludlow (1997) for a detailed analysis of the relationship between the UK and the EEC in the 1960s.

<sup>38</sup> After all, the issue of membership expansion was commonly referred to as the ‘English Question’ effectively suggesting that the first wave of enlargement could only take place with the participation of Britain.

economic system was commonly shared by all actors involved. On the political side, the Cold War balance of power seemed unwavered despite heightened moments of turmoil (Cuban missile crisis). It is within the above historical context that one may tease out the following hypotheses with respect to the interrelated course of events that led from the inception of the EEC to its gradual expansion in the early 1970s.

De Gaulle's successive vetoes of British membership were rational but costly actions aiming at extracting the highest possible bargaining surplus from future potential enlargement negotiations at the cost of delayed expansion. His main objective was to signal the political viability of the existing regime (parameter  $y_{AB}$  in the model) to the British. In undermining British accession negotiations with the Community, de Gaulle also wanted to ensure that (1) the customs union and its common external tariff had been completed subject to the timetable set out by the Treaty of Rome, (2) the Common Agricultural Policy and its financial regulation framework became *fait accompli*, and (3) the institutional make-up of the union strongly reflected French interests (captured by ideal point  $m_A$ ), before he would agree to enter into membership negotiations with Britain (see Moravcsik 1998, Chapter 3). That being said, De Gaulle never rejected British membership in principle as evidenced by his statements. Hence, the period between the Treaty of Rome (1957) and the Hague Summit (1969)—in effect overlapping with de Gaulle's second tenure in office (1958–1969)—corresponds to the first period in our model, where extreme country  $A$ 's (i.e., France's) proposed scheme of integration  $x_0^{A*} = m_A$  is only accepted by moderate country  $B$  (i.e., Germany along with the rest of the 'Five' of EEC-6) and, thus, becomes an entrenched and consolidated *interim status quo*.

France was apprehensive that premature British accession would derail the ongoing common market integration process, lead to the renegotiation of the CAP (through the formation of a strong British–German pro-industrial axis), and on the whole alter the character of the EEC integration project towards a looser Atlanticist free trade area.<sup>39</sup> This scenario corresponds to an 'accommodating' first-period proposal  $x_0^A = \hat{x}^C(\delta)$  (see Proposition 2) by proposing country  $A$ . On a further note, even though the possibility of a veto does arise along the equilibrium coalition-formation path (within a non-empty support of types  $y_{AB} \in \left( y_{ABC} - \frac{(m_C - m_A) - y_{AB}}{2}, y_{ABC} \right]$ ), the two-period time structure of the model may not capture the strategic nuances of repeated vetoes. To that end, one needs a multi-period model of successive proposals to generate a more complex signaling equilibrium with vetoes and Bayesian updating of beliefs. That being said, the logic of *strategic delay* that comes out of our model does provide the gist of the game-theoretic intuition in interpreting this prolonged historical episode.

Germany, i.e., moderate country  $B$ , and the rest of the EEC-6 members did not want to risk existing gains (namely variable  $y_{AB}$  in the model) by threatening the integrity of the Community. Caught within the 'strategic altercation' between the two extreme countries (namely France and the UK), it played a more passive, strategically

<sup>39</sup> The following quote (Ludlow 2006, p. 138) by General de Gaulle is quite indicative on this point: "Either it will have to be recognized that their [the British] entry into the Common Market, with all the exceptions that would inevitably accompany it, with all the quantitative and qualitative changes that it would entail, and with the participation of multiple other states that would certainly be its corollary, would amount to the establishment of an entirely new entity, all but erasing that which has been built. And where, then, would this lead us other perhaps than the creation of a type of European free trade area, which would in turn lead to an Atlantic zone that would deprive our continent of any real personality."

irrelevant role in this signaling game, since the fact that all country  $B$  types have an incentive to pool their strategies makes for uninformative signals. However, the longer British accession was being delayed, the more impatient they all became, given that they had more to gain on account of their stronger politico-economic links and shared economic outlook with the British. While the reaction of the ‘Five’ to the *first* French veto was only harsh in rhetoric but not in actions aiming to undermine the integration process, it became evident that in the aftermath of the *second* French veto the pace of integration would be substantially decelerated almost coming to a standstill, until the enlargement controversy had been successfully addressed.

The British, on the other hand, underestimated the political and economic stakes of existing members in the preservation of the European Economic Community (i.e., beliefs  $\sigma^C(y_{AB}|x_0^A)$  over partial union value  $y_{AB}$ ) as well as their willingness to adhere to the general principles of the *acquis* (captured by discount factor  $\delta$ ). They could not accurately anticipate De Gaulle’s commitment to a more competitive, open, and free-market oriented economy aimed at locking in the Germans in a beneficial subunion agreement. This is not a story of irrational misperception but rather of rational decision-making under *incomplete information*. Despite the common residual uncertainty with respect to the global economic environment, the founding members of the EEC (and more importantly France and Germany) had a more accurate assessment of the synergistic benefits of cooperation amongst themselves, given their incentive to commission expert reports on the costs and benefits of regional integration and to share expertise on the issue within intergovernmental fora (such as the Messina Conference in 1955 in which the British did not take part). Only they could correctly anticipate the synergistic benefits of market integration and trade liberalization together with the resulting ‘peace dividends’. As the model would predict, only if the French knew that these were high enough, could they afford to exclude the British from the incipient cooperation agreement.

When the British eventually came to realize how much they had to gain or by corollary how little the French would benefit from enlargement (i.e., country  $C$ ’s marginal contribution to the *grand* union  $y_{ABC}-y_{AB}$ ), they softened their bargaining stance and agreed to much more onerous accession terms (second-period accession policy outcome  $x_1^{i*} = x_{ABC}$ ). All in all, the British achieved a less favorable negotiation outcome in 1971, whose terms were avowedly worse for them than the 1963 foiled agreement, simply because they could not afford to remain excluded from a burgeoning European Community and to have their interests unaccounted for in the design of its policies and institutions.

This can avowedly not be a story of partisan re-orientation of the country, since (1) British accession negotiations were concluded under Conservative Prime Minister Edward Heath, who was also the Lord Privy Seal or main negotiator during the 1961–1963 period, and (2) a general cross-partisan consensus came about in Britain in the mid 1960s with regard to the desirability of entry into the EC. To summarize, a way to explain this dynamic through a rational-choice framework is by arguing that the French reluctance to expand the union in the 1960s was simply an attempt to signal its strong contentment with the EEC-6 *status quo*. Gradualism in this vein may be construed as a strategic ploy in an uncertain environment of overlapping interdependencies with the aim of tilting future expansion negotiations in the founding members’ favor. The clash between France and Britain in this instance became even more pronounced because of



their structurally distinct commercial interests and their diametrically opposed preferences over the ideological orientation of the European integration project.

Although the perception of French and British national interests by de Gaulle and Macmillan respectively may also be viewed through the prism of the wider geopolitical environment of that period (as many historians are apt to do),<sup>40</sup> strategic political economy factors (coupled with domestic political considerations) do certainly come into play in examining the dynamics of enlargement. Even though geopolitical incentives loomed in the background in the form of ‘security externalities’ (Gowa 1994), the primary strategic considerations driving the integration and enlargement process were essentially economic in nature, especially given that attempts towards greater political integration and security cooperation had not yet come to fruition following the failure of the Fouchet Plan in the early 1960s.

Even if the French perception of the UK as the ‘Trojan Horse’ of the Americans may go some way towards explaining the latter’s absence from the Treaty of Rome negotiations and de Gaulle’s subsequent vetoes, it may not, however, account for the subsequent successful resumption and completion of the enlargement negotiations. Why would the Gaullists alter their perception of Britain’s geopolitical role in the absence of any significant developments in the global balance of power? De Gaulle’s European policy was not dictated by illusions of *grandeur* or crude geopolitical considerations (Moravcsik 2000a, b); his perception of a ‘European Europe’ was above all of an economic nature and his primary concern was the promotion of shared commercial interests. It is within the above historical context that the above political economy story of gradualism becomes germane.

## 5 Discussion and Extensions

This paper represents a step towards formalizing and conceptualizing the ‘big-picture’ dynamics and strategic incentives inherent in the process of political integration and union formation. It provides an explanation for the paradox of union expansion to formerly eligible countries (UK, Scandinavian countries, Austria) or in other words piece-meal coalition formation in the absence of binding exogenous (e.g., geopolitical) constraints. I have managed to derive equilibrium delay in the formation of the *grand* union through the postulate of *asymmetric information*. In the proposed semi-separating equilibrium, certain types of the extreme first-period proposing country will seek to participate in a partial union with the moderate country, in order to signal its high *interim* reservation utility and to extract the highest possible surplus from the excluded expansion country in the formation of the *grand* union.

Tampering with the dimensionality and the information structure of the model suggests one of the ways to proceed in making it richer, more comprehensive, and more realistic, by obviating the need for complex exogenous assumptions. Given the broad, non-issue specific nature of political unions, it seems more than plausible to assume more than one policy dimensions in the negotiation process, thus giving rise to opportunities for issue-trading, issue linkages, as well as ‘enhanced cooperation’

---

<sup>40</sup> See, for example, Vaïsse 1998, de la Serre 1992.

in the form of policy-specific subunions. Adding a second policy dimension or public good opens up a range of possibilities with regard to the equilibrium relationship between union size and scope and the optimal rules of union formation. An efficiency comparison of the bargaining equilibria of various schemes of union formation, such as ‘federal package deals’, ‘enhanced cooperation’, and ‘open partnerships’<sup>41</sup> could be potentially quite enlightening about the future of the European project. Alesina et al. (2001, 2005) use a public goods approach to predict a bias towards excessive centralization and small union size. However, the historical record of European integration particularly in the 1990s has shown that union expansion may be concomitant with deepening under the appropriate rules. In this extension to the model with multiple policy jurisdictions or public goods, it would then seem appropriate to examine how the coalition formation protocol affects the equilibrium relationship between size and scope.

One related avenue for future research would be to explore a dynamic model of enlargement and union deepening, whereby piece-meal integration helps current and prospective members refine and signal their beliefs about the common uncertain benefits of integration through a number of random sample draws proportional to the degree of integration. Subunions (or ‘enhanced cooperation agreements’) could serve as policy laboratories experimenting on the actual effects of policy coordination in particular areas subject to highly variable exogenous shocks. Moreover, it would be of theoretical interest to model the supranational bargaining process as a continuous-time war-of-attrition game with uncertainty (see for example Admati and Perry 1987; Cramton 1992), in order to derive *strategic delay* through a perfect separation of types.

One may also wish to conjecture how the bargaining outcomes of the above model would be affected by successive enlargement negotiations in a multi-period dynamic framework. What happens when there is an exogenous increase in the pool of eligible expansion countries? Arguably, the more entrenched and institutionalized an existing union is, the easier it is to gauge the potential effect of a new member on its collective synergies and policy orientation. For example, the strategic calculus and cost-benefit analysis inherent in the Southern and Eastern expansions of the European Community was much more clear-cut compared to its first enlargement. This helps explain why the strict conditionality clauses imposed on the most recent accession countries appeared as unduly harsh and discriminatory, even though these countries were much better off joining the EU than staying out (Moravcsik and Vachudova 2003). One could thus provide a theoretical account of the gradual shift in the EU enlargement process from open-ended accession negotiations (as evidenced in the earlier waves of enlargement) to the more rigorous application of

---

<sup>41</sup> Models of ‘variable geometry’, such as the *concentric circles* approach proposed by Karl Lamers and the *eccentric circles* approach proposed by Edouard Balladur, essentially distinguish between a core and a periphery of countries integrating over overlapping and non-overlapping jurisdictions and subunions. In addition, *generalized subsidiarity* and *open partnerships* refer to a model of flexible integration put forward by Dewatripont et al. (1995) that advocates the need for commitment to a common base of integration, allowing at the same time for discretion on the part of member-states to experiment and engage in optional new forms of cooperation in other policy areas.

the *conditionality* principle (most notable in the recent expansion from 15 to 27 and increasingly applied in ongoing accession negotiations with candidate-members).<sup>42</sup>

Finally, in light of the economic and political interdependence of otherwise sovereign states, it would also be instructive for our purposes to allow for policy *externalities* or *spillovers* (positive or negative), whereby the welfare of the autarchic country is affected by the common policy adopted by the bilateral union. It should be expected that the presence of externalities would have a significant impact on equilibrium payoff allocations and coalition formation paths.<sup>43</sup>

**Acknowledgement** I would like to thank John Londregan, Andrew Moravcsik, and Adam Meirowitz for their invaluable guidance as well as Kristopher Ramsay, Marco Battaglini, Andrea Vindigni, Thomas Romer, Thomas Palfrey, and two anonymous referees for their helpful comments and suggestions. I have also greatly benefited from personal discussions with Paola Conconi, Abdul Noury, Micael Castanheira, André Sapir, Paul Magnette, and Jean Pisani-Ferry while at ECARES/ULB and helpful feedback during my presentations at APSA 2006, MPSA 2007, the CESifo Venice Summer Institute Workshop on 'Reinventing Europe' (2007), and the KOF-ETH Political Economy of International Organizations Conference (2008). Finally, I would like to thank the Alexander S. Onassis Public Benefit Foundation for their financial support during the time of writing.

## Appendix

**Proposition 1** For any  $\delta \in [0, 1]$  and  $y_c \geq 0$ ,  $c \in 2^N \setminus \emptyset$  subject to assumptions 1, 2, and 3, there can be no gradual coalition-formation path in any subgame-perfect Nash equilibrium of the baseline game with complete information.

*Proof* To characterize the subgame-perfect Nash equilibria of this bargaining game, we first need to define the players' *proposal* and *acceptance* strategies, such that they satisfy *sequential rationality* in every subgame. Following a common policy proposal  $x_t^i$  by proposer  $i \in \{A, B, C\}$  at time  $t=0, 1$ , then players  $j, k \neq i$  need to simultaneously decide whether to accept or reject. Depending on the proposal and the underlining parameters of the model, accepting may be a (weakly) dominant strategy or just a best response. In a subgame-perfect Nash equilibrium, the *acceptance* strategies are characterized by *conditional* and *unconditional acceptance sets* defined for  $i \neq j \neq k$  as follows:

Conditional acceptance sets :

$$CA_t^i(\pi_t, y_{ij}, y_{jk}) = \{x_t^i : V_t^j(\alpha_t^j(x_t^i) = 1) \geq V_t^j(\alpha_t^j(x_t^i) = 0) | \alpha_t^k(x_t^i) = 1\}$$

<sup>42</sup> Reluctance (or inability) on the part of candidate-members to pledge full adherence to the existing *acquis* points to the adoption of more flexible modes of enlargement in the future by manner of 'privileged partnership' agreements and a more inclusive *European Neighborhood Policy* ([http://ec.europa.eu/world/enp/index\\_en.htm](http://ec.europa.eu/world/enp/index_en.htm)). Quoting Enlargement Commissioner Olli Rehn in his recent speech at the Finnish Institute of International Affairs (October 27, 2006): "By keeping our word and sticking firmly to the accession perspective, we can create a virtuous circle of credible commitment, rigorous conditionality, and reinforced reforms."

<sup>43</sup> Etro (2002) has analyzed the model with three countries and *spillovers* in international policy coordination and has found that, if union policy is characterized by strategic complementarities, then the *grand* union is much more likely to form than in the case of strategic substitutabilities.

Unconditional acceptance sets :

$$UA_t^j(\pi_t, y_{ij}, y_{jk}) = \{x_t^i : V_t^j(\alpha_t^j(x_t^i) = 1) \geq V_t^j(\alpha_t^j(x_t^i) = 0), \forall \alpha_t^k(x_t^i)\}$$

In light of our assumption of *weak superadditivity*, it quite obvious that  $UA_t^j \subseteq CA_t^j, \forall t, \pi_t, j, k \neq i$ , since the *marginal contribution* of any additional member to the orthogonal *worth* of a coalition can only be positive. Moreover, given the unidimensionality of the model and the linearity of the utility functions, these *acceptance sets* are *convex* and *compact*.

The simultaneity of the acceptance game gives rise to a multiplicity of equilibria in subgames with common policy proposals  $x_t^i \in (CA_t^j \cap CA_t^k) \setminus (UA_t^j \cup UA_t^k), i \neq j, k$ , i.e., such that pure acceptance is only a best response but not a dominant strategy for either non-proposing player  $j$  or  $k$ . This implies that an action profile of *(no, no)* would also constitute a pure Nash equilibrium in such a proposal subgame, albeit Pareto (weakly) dominated by the *(yes, yes)* equilibrium profile. The indeterminacy of such coordination voting games generates a continuum of subgame-perfect Nash equilibria, whereby the proposer’s optimal proposal is conditioned by the responders’ mutually reinforced expectations of equilibrium play in any proposal subgame with two distinct pure strategy Nash equilibria. This gives rise to the possibility of credible rejection threats by the responders for such proposed common policies that make them both better off *if and only if* the *grand* coalition *ABC* forms.

Now turning to optimal proposal strategies, *subgame-perfection* and the underlying coalition *value* parameters of the game, i.e.,  $y_c \geq 0, \forall c \in 2^N \setminus \emptyset$ , would imply that in equilibrium no autarchic player would make a common policy proposal that would never be acceptable to either of his/her coalition partners. There always exists a potential non-singleton coalition which any country  $i$  prefers to the autarchic state. However, the content of that proposal can determine the size of the proposer’s optimally preferred coalition at time  $t$  and for given coalition structure  $\pi_t$ .

*Gradual* coalition-formation in this model occurs whenever a player joins an existing coalition in the second period of bargaining or any new coalition forms after a first-period bargaining impasse. To show that *gradualism* may not arise in any *subgame-perfect Nash equilibrium* of the complete-information workhorse model, I examine each possible coalition-formation-path and then use *proof by contradiction*:

1. Let  $\pi_{t=0} = \{A|B|C\}$ , i.e., no coalition has formed after the first round of bargaining for some  $\delta \in [0, 1]$  and some first-period proposer  $i_{t=0} \in \{A, B, C\}$ . Then in any subgame-perfect equilibrium, for given conjectures about play off the equilibrium path, proposer  $i_{t=1}$  will propose some common policy to players  $j, k \neq i$  such that

$$x_1^{i*} = \arg \max_{x_1^i} \left\{ y_{j\alpha_t^j(x_1^i)=1}^{j*} - |x_1^i - m_i| |x_1^i \in (CA_1^{j*} \cap CA_1^{k*}) \cup (UA_1^{j*} \cup UA_1^{k*}) \right\}, j \neq k \neq i.$$

Depending on the coalition values of the underlying cooperative structure, the proposer will choose the best possible two-country or three-country coalition in accordance with *sequential rationality* and *subgame perfection*. Note that the non-proposing players  $j$  and  $k$  may only threaten to coordinate on a suboptimal *(no,*

no) equilibrium for  $x_1^i \in (CA_1^{j*} \cap CA_1^{k*}) \setminus (UA_1^{j*} \cup UA_1^{k*})$ , since any policy proposal within a player's *unconditional acceptance set* will surely be accepted in equilibrium (dominant strategy). Using backwards induction, one may infer the players' optimal *acceptance sets* in the *first* period in anticipation of equilibrium play in the *second* period. If second-period *acceptance sets* are non-empty, then so are first-period sets. Hence, because of intertemporal discounting and random recognition, it would never be optimal for the first-period proposer to make a proposal unacceptable to either coalition partner, since that would imply foregoing the immediate benefits of coalition formation as well as one's proposal prerogative. First-period *acceptance sets* are always non-empty even for very low values of bilateral coalition formation ( $y_{AB}, y_{BC}, y_{AC}$ ), since we know by assumption that there always exists a non-degenerate set of policies such that the *grand* coalition  $ABC$  is always Pareto superior to the autarchic state; therefore, autarchy may not persist as a first-period bargaining outcome within any *subgame-perfect Nash equilibrium*.

2. Now let  $\pi_{t=0} = \{AB(x_{AB})|C\}$ , i.e., partial union  $AB$  has formed after round one at some common policy within the Pareto set  $[m_A, m_B]$  of coalition  $AB$ . That can only be the outcome of a first-period proposal by  $A$ , such that  $x_0^{A*} = \arg \max_{x_0^A} \{V_0^{A*}(x_0^A) | x_0^A \in UA_0^B \setminus CA_0^C\}$ , since  $B$ 's subgame-perfect first-period

proposal of  $m_B$  would have been accepted by both  $A$  and  $C$ . We need to show that  $ABC(x_{ABC})$  cannot be the outcome of the second round of bargaining in a subgame-perfect NE for any  $x_{AB}, x_{ABC}$ , and  $\delta$ . Assume *by contradiction* that  $\pi_{t=1} = \{ABC(x_{ABC})\}$ . *Subgame perfection* would imply that  $x_{ABC} \geq m_C - y_{ABC}$  ( $C$ 's participation constraint) and  $x_{ABC} \leq y_{ABC} - y_{AB} + x_{AB}$  ( $A$ 's participation constraint). For the *grand* coalition unanimity *acceptance set* to be non-empty, we need that  $x_{AB} \geq m_C + y_{AB} - 2y_{ABC}$  (\*). This effectively rules out any  $x_{AB} < m_C - 2y_{ABC}$  as possible equilibrium first-period proposals. We proceed to prove the contradiction by showing that profitable deviations exist for any other possible  $x_{AB}$ :

- a. Let  $x_{AB} \in [m_C - 2y_{ABC}, m_A]$ : this is a Pareto-dominated set of proposal for both  $A$  and  $B$ , since both would be unambiguously better off with a proposed policy of  $x_{AB} = m_A$ ; it is not only closer to their ideal positions, but it also enhances their bargaining leverage vis-à-vis  $C$  by shrinking the  $ABC$  Pareto set.
- b. Let  $x_{AB} = m_A$ : from (\*) this implies that  $y_{AB} \leq 2y_{ABC} - (m_C - m_A)$ . For low types  $y_{AB} \in [0, y_{ABC} - (m_B - m_A)]$ , the contradiction follows by showing that  $V_0^A(\langle ABC(x_{ABC}^*), ABC(x_{ABC}^*) \rangle) \geq EV_0^A(\langle AB(m_A)|C, ABC(x_1^{j*}) \rangle), \forall \delta, y_{ABC}$ , where the expectation is taken over the identity of the proposer  $j$  at  $t = 1$  (hence  $x_1^{j*}$  is *ex ante* unknown in equilibrium) and  $x_{ABC}^* = \arg \max_{x_0^A} \{V_0^A(x_0^A) | \alpha_0^{j*}(x_0^A) = 1, \forall j \neq A, x_0^A \in X\}$ . A simple algebraic calculation shows that this holds for any  $y_{AB}$  within the above interval. Hence,  $A$  would have an incentive to deviate to a better proposal given the *subgame-perfect acceptance strategies* of  $B$  and  $C$ . Similarly, for intermediate types  $y_{AB} \in [y_{ABC} - (m_B - m_A), 2y_{ABC} - (m_C - m_A)]$ , there always exists a globally acceptable, Pareto efficient *grand* coalition first-period proposal  $x_{ABC}^*$  that makes  $A$  weakly better off compared to the *gradual* coalition-formation subgame.

Going through all the possible subgames and subcases for  $y_{AB} \in [y_{ABC} - (m_B - m_A), 2y_{ABC} - (m_C - m_A)]$  and  $\delta \in [0, 1]$ , it turns out that for  $x_{ABC} = m_A + y_{ABC} - y_{AB} - \varepsilon$ ,  $\varepsilon \geq 0$ , there always exists an  $\varepsilon \geq 0$  such that the following conditions are satisfied:

- i.  $V_0^A(x_0^A = x_{ABC}^*) \geq EV_0^A(x_0^A = m_A)$  (*A*'s optimization problem)
  - ii.  $x_{ABC}^* \in UA_0^{B*} \cap CA_0^{C*}$  (*B* and *C*'s mutual *acceptance* forms a Nash equilibrium)
  - iii.  $x_{ABC}^* = m_A + y_{ABC} - y_{AB} - \varepsilon \geq m_C - y_{ABC}$  (*C*'s participation constraint)
- c. Finally, for any  $x_{AB} > m_A$  it would be enough to show that, given that only *B* will accept, i.e.,  $x_{AB} \in UA_0^{B*} \setminus CA_0^{C*}$ , *A* would profit from deviating to a proposal  $x_0^{A*} = \arg \max \{V_0^A(x_0^A) \mid x_0^A \in UA_0^{B*} \setminus CA_0^{C*}\}$ , i.e., the one closest possible to  $m_A$ , such that only country *B* would accept. Indeed, this is a profitable deviation, since it would bring both immediate policy gains and enhanced second-period bargaining leverage (because of a restricted *grand* coalition Pareto set). For some  $x_0^A \geq x_{ABC}^* \geq m_C - y_{ABC}$ , *subgame perfection* and *sequential rationality* also imply that *C* should also have accepted *A*'s first-period proposal.
3. Now let  $\pi_{t=0} = \{A|BC(x_{BC})\}$ , i.e., partial union *BC* has formed after round one (following a proposal by country *C*). Using a similar reasoning as above, the assumption that  $\pi_{t=1} = \{ABC(x_{ABC})\}$  leads to a contradiction, since that coalition-formation path cannot be part of a *subgame-perfect Nash equilibrium*.
4. Finally the case of  $\pi_{t=0} = \{AC(x_{AC})|B\}$  may never arise in equilibrium, since either *A*'s or *C*'s *unconditional acceptance* of each other's policy proposals implies that the median country *B* is always better off accepting too. QED

**Proposition 2** *In the two-period bargaining game with asymmetric information there exists a perfect Bayesian Nash equilibrium, whereby the representative of either of the extreme countries  $i=A$  or  $C$  will propose  $x_0^{i*} = m_i$  at time  $t=0$  if and only if  $y_{iB} \in (\tilde{y}_{iB}, y_{ABC}]$  for some  $\tilde{y}_{AB} \in [y_{ABC} - (m_B - m_A), 2y_{ABC} - (m_C - m_A)]$  or  $\tilde{y}_{BC} \in [y_{ABC} - (m_C - m_B), 2y_{ABC} - (m_C - m_A)]$ , in which case coalition *AB* (or *BC* respectively) will form right away and may later expand to the grand coalition *ABC* at time  $t=1$  with strictly positive probability. Otherwise, for  $y_{iB} \in [0, \tilde{y}_{iB}]$ ,  $i=A$  or  $C$  will propose  $x_0^{i*} = \hat{x}^j(\delta)$  such that  $j \neq i$ ,  $j=A, C$  is just indifferent between accepting and rejecting at time  $t=0$ , in which case the grand coalition *ABC* will form immediately. If median country *B* gets to propose first, then all its types will pool on an equilibrium proposal  $x_0^{B*} = m_B$ , which will lead to the immediate formation of the grand coalition.*

**Proof** Having characterized how the proposed equilibrium plays out in the subgames, where *A* is the first-period proposer, I will now proceed to derive the equilibrium threshold type  $\tilde{y}_{iB}$  for either extreme country  $i=A, C$  as the first-period proposer. Lemma 1 below completes the characterization by deriving the threshold point  $\tilde{x}_0^i(\delta)$ ,  $i=A, C$  and the corresponding compromising first-period proposal  $\hat{x}^j(\delta)$ ,  $j=A, C, j \neq i$ .

The unique threshold type  $\tilde{y}_{AB}$  has to be such that the representative of country  $A$  would be *ex ante* indifferent between the *immediate*  $\langle ABC(\hat{x}^C), ABC(\hat{x}^C) \rangle$  and the *gradual*  $\langle AB(m_A)|C, ABC(x_{ABC}) \rangle$  equilibrium coalition formation paths. Hence, to ensure the incentive-compatible truthful revelation of types at  $t=0$  within the context of the above semi-separating equilibrium, the cutoff type  $\tilde{y}_{AB} \in [y_{ABC} - (m_B - m_A), 2y_{ABC} - (m_C - m_A)]$  for country  $A$  will have to satisfy the following incentive constraint:

$$\begin{aligned}
 V_0^A(x_0^A = m_A | y_{AB} = \tilde{y}_{AB}) &= V_0^A(x_0^A = \hat{x}^C | y_{AB} = \tilde{y}_{AB}) \Leftrightarrow \\
 \tilde{y}_{AB} + \frac{\delta}{3} \left[ 2y_{ABC} - (m_C - m_A) + \tilde{y}_{AB} + m_A - \frac{m_A + m_C - \tilde{y}_{AB}}{2} + y_{ABC} \right] &= (1 + \delta)[m_A - \hat{x}^C + y_{ABC}] \Leftrightarrow \\
 \tilde{y}_{AB} = m_A + \frac{\delta}{2 + \delta} m_C - \frac{2(1 + \delta)}{2 + \delta} \hat{x}^C + \frac{2}{2 + \delta} y_{ABC} & \tag{2}
 \end{aligned}$$

In order to derive the above expression for  $\tilde{y}_{AB}$ , I make use of the optimal *proposal* and *acceptance* strategies in the second-period subgames analyzed in the model and also of the fact that the *lowest* of high types  $\tilde{y}_{AB}$  will always accept  $C$ 's second-period proposal  $x_1^{C*} = \frac{m_A + m_C - \tilde{y}_{AB}}{2}$  with certainty. Also note that each second-period subgame equilibrium utility is discounted by a factor  $\delta/3$  reflecting temporal discounting and equal recognition probabilities. Given the spatial location of  $\hat{x}^C(\delta) \in [m_C - y_{ABC}, m_B]$  derived in Eq. 5 below, it is fairly straightforward to confirm that the cutoff type  $\tilde{y}_{AB}$  indeed lies within the interval  $[y_{ABC} - (m_B - m_A), 2y_{ABC} - (m_C - m_A)]$ . Using a similar approach for  $C$ , where proposal  $\hat{x}^A(\delta)$  is computed in Eq. 6 below, one can derive the following cutoff proposer type:

$$\tilde{y}_{BC} = \frac{2}{2 + \delta} y_{ABC} + \frac{2(1 + \delta)}{2 + \delta} \hat{x}^A - \frac{\delta}{2 + \delta} m_A - m_C. \tag{3}$$

Finally, for all subgames starting with country  $C$  as the first-period proposer, extreme country  $A$ 's equilibrium beliefs are as follows:

$$\sigma^{A*}(y_{BC} | x_0^C) \rightarrow \begin{cases} y_{BC} \sim U[0, \tilde{y}_{BC}], & \text{for } x_0^C \in (-\infty, \hat{x}^A) \cup (\hat{x}^A, m_A + y_{ABC}] \\ y_{BC} \sim U(\tilde{y}_{BC}, y_{ABC}), & \text{for } x_0^C \in (m_A + y_{ABC}, m_C) \cup (m_C, +\infty) \end{cases}$$

**Lemma 1** *There exists for country  $C$  a threshold point on the real line,  $\tilde{x}_0^C(\delta) \in [m_C - y_{ABC}, m_B]$ , such that for  $x_0^A < \tilde{x}_0^C(\delta)$  and  $\pi_0 = \{AB(x_0^A)|C\}$  a positive support of low types  $y_{AB} \leq \tilde{y}_{AB}$  will reject  $C$ 's optimal second-period proposal  $x_1^{C*}$  given its equilibrium beliefs, while for  $x_0^A \geq \tilde{x}_0^C(\delta)$  all low types  $y_{AB} < \tilde{y}_{AB}$  will accept  $x_1^{C*}$ . A similar threshold point  $\tilde{x}_0^A(\delta)$  exists for country  $A$ .*

**Proof**  $C$ 's optimal second-period proposal given its updated equilibrium beliefs following a first-period proposal  $x_0^A \in [m_C - y_{ABC}, m_B]$  that falls within the *grand* union Pareto set of common policies becomes

$$\begin{aligned}
 x_1^{C*} &= \arg \max_{x_1^C \in [m_C - y_{ABC}, m_C]} \left\{ \Pr\left(\alpha_1^{A*}(x_1^C) = 1 | y_{AB} \sim U[0, \tilde{y}_{AB}]\right) \times (x_1^C - m_C + y_{ABC}) \right\} \\
 &= \frac{m_C + x_0^A}{2}. \tag{4}
 \end{aligned}$$

Hence,  $\tilde{x}^C(\delta)$  has to be such that the threshold  $y_{AB}$  type is just indifferent between *accepting* and *rejecting*  $x_1^{C*}$ , i.e.,

$$\begin{aligned} u_{t=1}^A \left( \text{Acc } x_1^{C*} = 1/2(m_C + \tilde{x}^C) | y_{AB} = \tilde{y}_{AB} \right) \\ = u_{t=1}^A \left( \text{Rej } x_1^{C*} = 1/2(m_C + \tilde{x}^C) | y_{AB} = \tilde{y}_{AB} \right). \end{aligned}$$

So for  $x_0^A < \tilde{x}^C(\delta)$  and  $\pi_0 = \{AB(x_0^A) | C\}$ , there is a positive support of low country  $A$  types  $y_{AB} \leq \tilde{y}_{AB}$  that reject  $x_1^{C*} = \frac{m_C + x_0^A}{2}$ , which implies that  $C$ 's first-period *rejection* continuation value  $V_{t=0}^C(\text{Rej } x_0^A | B \text{ acc})$  becomes quadratic. For  $x_0^A \geq \tilde{x}_0^C(\delta)$  and  $\pi_0 = \{AB(x_0^A) | C\}$ , all low types  $y_{AB} \leq \tilde{y}_{AB}$  will accept  $x_1^{C*} = \frac{m_C + x_0^A}{2}$ , since  $C$ 's optimal second-period proposal would make the highest possible low type  $\tilde{y}_{AB}$  just indifferent. This implies that  $C$ 's *rejection* continuation value becomes linear, in which case it is quite straightforward to derive the value of  $\tilde{x}^C(\delta)$  in the following manner:

$$\tilde{x}^C(\delta) = \frac{6 + 2\delta}{6 + 5\delta}(m_C - y_{ABC}) + \frac{2\delta}{6 + 5\delta}m_B + \frac{\delta}{6 + 5\delta}m_C. \tag{5}$$

Using a similar approach to derive  $\hat{x}^A(\delta)$  for country  $C$  as the first-period proposer, it turns out that:

$$\hat{x}^A(\delta) = \frac{6 + 2\delta}{6 + 5\delta}(m_A + y_{ABC}) + \frac{2\delta}{6 + 5\delta}m_B + \frac{\delta}{6 + 5\delta}m_A \leq \tilde{x}^A(\delta). \tag{6}$$

**Proposition 3** *For a non-degenerate support of high  $y_{AB}$  types  $(\tilde{y}_{AB}, \tilde{y}_{AB} + \varepsilon]$ ,  $\varepsilon > 0$ , there exists an immediate grand coalition formation policy  $x_{ABC}^* \in [m_C - y_{ABC}, m_B]$  such that  $\langle ABC(x_{ABC}^*), ABC(x_{ABC}^*) \rangle \succeq_i \langle AB(m_A) | C, ABC(x_1^{j*}) \rangle$  for all  $i, j = A, B, C$  and  $\langle ABC(x_{ABC}^*), ABC(x_{ABC}^*) \rangle \succ_i \langle AB(m_A) | C, ABC(x_1^{j*}) \rangle$  for at least one  $i$ , where  $\succeq$  and  $\succ$  denote the weak and strict preference relations respectively and  $x_1^{j*}$  denotes the optimal second-period proposal for any representative  $j$  to get recognized.*

**Proof** Let  $y_{AB} \in (\tilde{y}_{AB}, \tilde{y}_{AB} + \varepsilon]$ ,  $\varepsilon > 0$ . We first need to find the set of *grand union* policy proposals that make  $A$  weakly better off in an *immediate* coalition formation path rather than a *gradual* one, i.e., find  $x_{ABC} \geq m_A$  such that

$$V_0^A(\langle ABC(x_{ABC}), ABC(x_{ABC}) \rangle) \geq EV_0^A(\langle AB(m_A) | C, ABC(x_1^{j*}) \rangle), j = A, B, C.$$

So we need the following inequality to hold:

$$\begin{aligned} (1 + \delta)(m_A - x_{ABC} + y_{ABC}) &\geq y_{AB} + \frac{\delta}{3} \left[ 2y_{ABC} - (m_C - m_A) + y_{AB} + m_A - \frac{m_A + m_C - \tilde{y}_{AB}}{2} + y_{ABC} \right] \\ \Leftrightarrow m_A \leq x_{ABC} &\leq \frac{1}{1 + \delta} \left( \frac{2 + \delta}{2} m_A + \frac{\delta}{2} m_C + y_{ABC} - \frac{3 + \delta}{3} y_{AB} - \frac{\delta}{6} \tilde{y}_{AB} \right) = \bar{x}_{ABC}. \end{aligned}$$

Given that  $\tilde{y}_{AB} \leq 2y_{ABC} - (m_C - m_A)$ , it turns out that  $\bar{x}_{ABC} \geq m_C - y_{ABC}$ , which implies that there may be such an immediate *grand* coalition proposal that could make  $C$  weakly better off. Since  $B$ , the moderate country, will trivially have a strict preference to participate in an immediate *grand* union with a common policy much



closer to its own ideal point, all we need to show is that  $C$  is weakly better off under such an *immediate* agreement compared to the *gradual* equilibrium, whereby  $C$  believes  $A$  and  $B$  to be of a high type, i.e.,

$$\begin{aligned} V_0^C((ABC(\bar{x}_{ABC}), ABC(\bar{x}_{ABC})) &\geq EV_0^C((AB(m_A)|C, ABC(x_1^*)) \\ \Leftrightarrow (1 + \delta)(\bar{x}_{ABC} - m_C + y_{ABC}) &\geq \frac{\delta}{3} \left[ (m_A + y_{ABC} - y_{AB}) - m_C + y_{ABC} + \frac{m_A + m_C - \tilde{y}_{AB}}{2} - m_C + y_{ABC} \right] \\ \Leftrightarrow y_{AB} &\leq 2y_{ABC} - (m_C - m_A). \end{aligned}$$

Since we know from before that  $\tilde{y}_{AB} \leq 2y_{ABC} - (m_C - m_A)$ , then the latter expression has to be true for some  $y_{AB} \in (\tilde{y}_{AB}, \tilde{y}_{AB} + \varepsilon]$ ,  $\varepsilon > 0$  and will hold as a strict inequality for any interior cutoff type  $\tilde{y}_{AB} < 2y_{ABC} - (m_C - m_A)$ . We have thus shown that an *immediate grand* coalition formation path under  $\bar{x}_{ABC}$  will be a Pareto superior solution, hence the *interim* inefficiency of the *gradualist* equilibrium. *QED*

## References

- Admati, A. R., & Perry, M. (1987). Strategic delay in bargaining. *The Review of Economic Studies*, 54(3), 345–364.
- Aghion, P., Antràs, P., & Helpman, E. (2007). Negotiating free trade. *Journal of International Economics*, 73(1), 1–30.
- Alesina, A., Angeloni, I., & Etro, F. (2001). *The political economy of international unions*. CEPR Discussion Papers no. 3117.
- Alesina, A., Angeloni, I., & Etro, F. (2005). International unions. *American Economic Review*, 95(3), 602–615.
- Alesina, A., & Grilli, V. (1993). *On the feasibility of a one or multi-speed European Monetary Union*. NBER Working Paper no. 4350.
- Alesina, A., & Perotti, R. (2004). *The European Union: A politically incorrect view* (Revised), mimeo. Harvard University, NBER and CEPR IGER—Universita Bocconi and CEPR.
- Alesina, A., & Spolaore, E. (2003). *The size of nations*. Cambridge, MA: MIT Press.
- Alesina, A., & Wacziarg, R. (1999). *Is Europe going too far?* *Carnegie-Rochester Conference Series on Public Policy*, 51.
- Baldwin, R. (1999). A domino theory of regionalism. In J. N. Bhagwati, P. Krishna, & A. Panagariya (Eds.), *Trading blocs: Alternative approaches to analyzing preferential trade agreements*. Cambridge, MA: MIT Press.
- Bednar, J. (2007). Valuing exit options. *Publius: The Journal of Federalism*, 37(2), 190–208.
- Berglöf, E., Eichengreen, B., Roland, G., Tabellini, G., & Wyplosz, C. (2003). *MEI 12 built to last: A political architecture for Europe*. London: Centre for Economic Policy Research.
- Besley, T., & Coate, S. (2003). Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of Public Economics*, 87(12), 2611–2637.
- Bolton, P., & Roland, G. (1997). The breakup of nations: a political economy analysis. *The Quarterly Journal of Economics*, 112(4), 1057–1090.
- Bordignon, M., & Brusco, S. (2001). Optimal secession rules. *European Economic Review*, 45(10), 1811–1834.
- Bordignon, M., & Brusco, S. (2006). On enhanced cooperation. *Journal of Public Economics*, 90(10–11), 2063–2090.
- Brou, D., & Ruta, M. (2006). Special interests and the gains from political integration. *Economics & Politics*, 18(2), 191–218.
- Casella, A. (1992a). On markets and clubs: economic and political integration of regions with unequal productivity. Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association. *The American Economic Review*, 82(2), 115–121.
- Casella, A. (1992b). Participation in a currency union. *The American Economic Review*, 82(4), 847–863.
- Casella, A., & Feinstein, J. S. (2002). Public goods in trade: on the formation of markets and jurisdictions. *International Economic Review*, 43(2), 437–462.
- Cramton, P. C. (1992). Strategic delay in bargaining with two-sided uncertainty. *The Review of Economic Studies*, 59(1), 205–225.

- Cremer, J., & Palfrey, T. R. (2000). Federal mandates by popular demand. *Journal of Political Economy*, 108(5), 905–927.
- de la Serre, F. (1992). De Gaulle et la Candidature Britannique aux Communautés Européennes. In Institut Charles de Gaulle (Ed.), *De Gaulle en Son Siècle*, 5.
- Dewatripont, M., Giavazzi, F., Harden, I., Persson, T., Roland, G., Sapir, A., et al. (1995). *Flexible integration: Towards a more effective and democratic Europe. Monitoring European Integration 6*. London: Centre for Economic Policy Research.
- Downs, G. W., & Rocke, D. M. (1995). *Optimal imperfection?: Domestic uncertainty and institutions in international relations*. Princeton, N.J.: Princeton University Press.
- Downs, G. W., Rocke, D. M., & Barsoom, P. N. (1998). Managing the evolution of multilateralism. *International Organization*, 52(2), 397–419.
- Ellingsen, T. (1998). Externalities vs internalities: a model of political integration. *Journal of Public Economics*, 68(2), 251–268.
- Etro, F. (2002). International policy coordination with economic unions. *Rivista Internazionale di Scienze Sociali*, 110(2), 187–211.
- European Neighborhood policy. Retrieved from [http://ec.europa.eu/world/enp/index\\_en.htm](http://ec.europa.eu/world/enp/index_en.htm).
- Fratianni, M. (1998). Variable integration for the European Union. *Tijdschrift voor Economie en Management*, XLIII(3), 315–336.
- Gilligan, M. J. (2004). Is there a broader-deeper trade-off in international multilateral agreements? *International Organization*, 58(3), 459–484.
- Gomes, A. (1999). A theory of negotiations and formation of coalitions. *CARESS Working Paper 99–12*.
- Gowa, J. S. (1994). *Allies, adversaries, and international trade*. Princeton, NJ: Princeton University Press.
- Grieco, J. M. (1990). *Cooperation among nations: Europe, America, and non-tariff barriers to trade*. Ithaca: Cornell University Press.
- Gruber, L. (2000). *Ruling the world: Power politics and the rise of supranational institutions*. Princeton, N.J.: Princeton University Press.
- Gstöhl, S. (2002). *Reluctant Europeans: Norway, Sweden, and Switzerland in the process of integration*. Boulder, Colo.; London: Lynne Rienner Publishers.
- Hafer, C., & Landa, D. (2007). Public goods in federal systems. *Quarterly Journal of Political Science*, 2(3), 253–275.
- Harstad, B. (2007). Harmonization and side payments in political cooperation. *The American Economic Review*, 97(3), 871–889.
- Hix, S. (1994). Approaches to the study of the EC: the challenge to comparative politics. *West European Politics*, 17(1), 1–30.
- Hix, S. (1998). Approaches to the study of the European Union: the ‘new governance’ agenda and its rival. *Journal of European Public Policy*, 5(1), 38–65.
- Jacquemin, A., & Sapir, A. (1995). Is a European hard core credible? A statistical analysis. *C.E.P.R. Discussion Papers No. 1242*.
- Keohane, R. O. (1984). *After hegemony: Cooperation and discord in the world political economy*. Princeton, NJ: Princeton University Press.
- Krasner, S. D. (1983). *International regimes*. Ithaca: Cornell University Press.
- Ludlow, N. P. (1997). *Dealing with Britain: The six and the first UK application to the EEC*. Cambridge: Cambridge University Press.
- Ludlow, N. P. (2006). *The European Community and the crises of the 1960s: Negotiating the gaullist challenge*. London; New York: Routledge.
- Maskin, E. (2004). Bargaining, coalitions, and externalities. *13th WZB conference on markets and political economy*. WZB & CEPR.
- Mearsheimer, J. J. (1994/1995). The false promise of international institutions. *International Security*, 19(3), 5–49.
- Milner, H. V. (1997). *Interests, institutions, and information: Domestic politics and international relations*. Princeton, NJ: Princeton University Press.
- Moravcsik, A. (1997). Taking preferences seriously: a liberal theory of international politics. *International Organization*, 51(4), 513–553.
- Moravcsik, A. (1998). *The choice for Europe: Social purpose and state power from Messina to Maastricht*. Ithaca, NY: Cornell University Press.
- Moravcsik, A. (2000a). De Gaulle between grain and grandeur: The political economy of French EC policy, 1958–1970 (Part 1). *Journal of Cold War Studies*, 2(2), 3–43.
- Moravcsik, A. (2000b). De Gaulle between grain and grandeur: The political economy of French EC policy, 1958–1970 (Part 2). *Journal of Cold War Studies*, 2(3), 4–68.

- Moravcsik, A., & Vachudova, M. (2003). National interest, state power, and EU enlargement. *East European Politics and Societies*, 17(1), 42–57.
- Oates, W. E. (1972). *Fiscal federalism*. New York: Harcourt Brace Jovanovich.
- Oates, W. E. (1999). An essay on fiscal federalism. *Journal of Economic Literature*, 37(3), 1120–1149.
- Padoan, P. C. (1997). Regional agreements as clubs: The European Case. In E. D. Mansfield, & H. V. Milner (Eds.), *The political economy of regionalism*. New York: Columbia University Press.
- Persson, T., & Tabellini, G. (1996). Federal fiscal constitutions: Risk sharing and moral hazard. *Econometrica*, 64(3), 623–46.
- Pisani-Ferry, J. (1995). L'Europe à Geometrie Variable, Une Analyse Economique. *CEPII Working Paper No. 4*.
- Rehn, O. (2006). Speech at the Finnish Institute of International Affairs. Retrieved 17 March 2008, from [http://www.upi-fiia.fi/eng/events/events\\_2006/europes\\_next\\_frontiers/](http://www.upi-fiia.fi/eng/events/events_2006/europes_next_frontiers/).
- Roberts, K. Dynamic voting in clubs. STICERD—Theoretical Economics Paper Series, January 1999.
- Ruta, M. (2005). Economic theories of political (dis)integration. *Journal of Economic Surveys*, 19(1), 1–21.
- Seidmann, D. J., & Winter, E. (1998). A theory of gradual coalition formation. *Review of Economic Studies*, 65(4), 793–815.
- Vaisse, M. (1998). *La Grandeur: Politique Étrangère du Général De Gaulle, 1958–1969*. Paris: Fayard.