International Unions and Integration

Oğuzhan Çelebi*
Stanford University

Elias Papaioannou†
London Business School

July 22, 2024

Abstract

We consider a model of international unions in which countries have heterogeneous preferences for integration, and their integration decisions are strategic complements. We study various integration protocols that differ in flexibility to shed light on the formation, expansion, and cohesion of the European Union (EU). Unlike previous models with strategic substitutes, our results align with the EU's history, where enlargement and flexible integration went hand in hand with deepening integration, often spearheaded by the "core" countries. Extending the framework to study unions' integration with non-members (candidates, exiting countries, and others) reveals the necessity of restrictions to non-member integration to foster cooperation and make the union robust to changing preferences of its members. We conclude with an exploration of the trade-offs of two-tier unions, an increasingly topical issue. Our results demonstrate the important role complementarities play in expanding membership and deepening integration in international unions.

*Department of Economics, Stanford University, 579 Jane Stanford Way, Stanford, CA 94305. Email: ocelebi@stanford.edu

†London Business School, Economics Department, Regent's Park. London NW1 4SA. United Kingdom. Email: eliasp@london.edu.

We are grateful to Daron Acemoglu, Roberto Corrao, Glenn Ellison, Joel Flynn, Bård Harstad, Michal Kobielarz, David Myatt, Giacomo Ponzetto, Enrico Spolaore, Birger Wernerfelt, Michael Whinston, and seminar participants in LBS, MIT Theory Lunch, the Bank of Greece, and MIT Organizational Economics Lunch for helpful comments. First Draft: June 2021.

1. Introduction

International organizations play an instrumental role in the world economy and geopolitics, shaping national and international policies; their domains range from trade and banking to agriculture, industrial policy, product standardization, and climate. While international organizations' objectives differ, a common goal is to foster members' cooperation and integration. The European Union (EU) is arguably the most prominent and complex international organization, as its domains have expanded significantly after its establishment. Starting in 1951 as the European Coal and Steel Community among six countries, the EU has morphed into a complex international institution of 27 countries with a *single market for goods, services, capital and labor*, unified trade policy with non-members, a common agriculture policy, legislative and regulatory harmonization in financial services, product standardization, among others.[1] Defying skeptics who prophesied its demise, European integration progressed, covering a wider policy domain, with stronger ties. Deepening integration took place, even accelerated, as the EU welcomed new members: the United Kingdom (UK), Ireland, and Denmark in 1973, Greece, Portugal, and Spain in the 1980s, Austria, Finland and Sweden in the mid-1990s, and Eastern Europe in the 2000s. Although BREXIT illustrated that the Union may not be fully robust, the pandemic and the war in Ukraine have led many policymakers and the public to argue for deepening integration alongside an expansion in new domains, like health and security.[2] The EU's incremental approach towards integration and gradual membership expansion serves as a model for regional integration worldwide.[3]

The experience of the EU raises many first-order questions, especially topical nowadays, where geopolitics is at the core of the world economy. How did integration evolve in an increasingly broader set of areas alongside enlargement from a small core to the periphery? Why does the EU allow for enhanced cooperation with candidates and other non-members, and what are the main trade-offs of these relationships? Why does the BREXIT deal with the UK entail a lower integration than many anticipated after the referendum? What are the trade-offs of multiple tiers with heterogeneous integration that many propose in Europe?

Here, we develop a game-theoretic model to shed light on these issues. We study how

[1] See Spolaore (2013), Gilbert (2020), Sapir (2011), and Eichengreen (2006) for extensive reviews on the history, functions, and transformation of the EU.

[2] In an influential talk at the European Parliament on May 2023, Mario Draghi, former Italian prime minister and President of the European Central Bank pleaded for *"pragmatic federalism"* and *"acceleration of the integration process"* in defense and welfare state policies, alongside enlargement in the Western Balkans. Emmanuel Macron's 2022 Presidential election campaign and his recent public interventions focus on deeper EU ties on defense, calling Europe to achieve *"strategic autonomy"* from the US.

[3] For example, in 2021, the African Union established the African Continental Free Trade Area (AfCFTA), which aims to be the world's largest free trade area, connecting about 1.3*bn* people across 54 African countries.

international unions facilitate integration, both between members and non-members. Our model nests earlier approaches (*e.g.*, Alesina, Angeloni, and Etro (2005)), which study "public good" games where countries' actions are strategic substitutes. However, we focus on integration with strategic complementarities, as this setting is more tightly aligned with the EU, which fosters cooperation and legislative/regulatory harmonization in trade, product markets, and financial intermediation rather than coordinating public goods investments. Unlike earlier works, our framework can jointly explain the essential transformations of the EU. We show how enlargement to the periphery and implementation of flexible (non-mandatory) integration protocols such as the Eurozone (monetary union) or the Schengen Treaty (on migration) co-exist alongside deeper integration, often initiated by "core" countries. Leveraging the model's flexibility, we also touch upon some chief issues that earlier theoretical research either did not explore or study in isolation. Extending the model to allow for the Union's integration with non-members, we gain insights into the relationship between the EU and candidates like Turkey and Albania, exiting countries like the UK, and outsiders with preferential integration like Norway. In addition, we explore the formation of two-tier unions, another topical and controversial issue.

A Model with Strategic Complementarities. We develop a framework of international integration in which countries with heterogeneous preferences decide their *integration levels*, either forming a union with a common policy (*rigid union*), a union with a minimum threshold where further integration is possible (*flexible union*) or integrating independently without explicit enforcement (*non-union integration*). Unlike the earlier literature, countries' actions representing integration are strategic complements rather than substitutes. Thus, as the union's size (number of countries) and depth (the extent of integration) increase, members have an incentive to integrate more to reap scale and market size effects. This setting, albeit simplified and abstract, provides a more realistic description of the EU's core emphasis on fostering a single market for goods, services, capital, and labor, standardization of regulations and safety protocols, financial sector legislative harmonization, and legal convergence. Besides, the EU budget is small compared to the combined budget of members (around 2% of EU public spending), and classic public goods, such as education, health, policing, and defense are provided mainly at the national level.

Deepening Integration and Enlargement to Periphery. We employ our theoretical framework to shed light on the EU's two main changes (in Section 3): the introduction of flexible non-mandatory integration protocols (*e.g.*, Eurozone, and Schengen) and a constant enlargement to the periphery. Both changes have been accompanied by deeper integration in an increasingly larger set of domains. Besides, such policies have often been spearheaded by the "core" countries such as Germany, France, and the Netherlands.[4] We commence by characterizing the equilibrium integration under different integration protocols that differ in flexibility (Proposition 1). We then describe how earlier theoretical explorations where countries' actions are strategic substitutes yield a reduction of integration after enlargement to the periphery and/or adoption of flexible integration protocols, which contradict the experience of the EU. Our simple departure from a public investment setting to strategic complementarities, more relevant for capital, labor, product, and service market integration, is sufficient to explain EU's evolution, where enlargement, deeper integration, and flexibility have moved in tandem. With strategic complementarities, greater integration from some countries and enlargement to the periphery incentivize all members to increase their integration, which increases the depth of the union (Propositions 2 and 3). As such, when complementarities are strong, "core" countries, with strong preferences for integration, may support the union's expansion to the periphery and more flexible integration policies, which also raise integration in equilibrium. Therefore, our framework is aligned with some salient features of European integration, deepening integration alongside flexible integration protocols, and expansion to the periphery with multi-country enlargement rounds.

Non-member Integration, Exit, and Enlargement. In Section 4, we introduce integration with non-members, a critical issue that earlier works have abstracted from. We use the extended framework to explore the tradeoffs of EU's agreements with candidates (e.g., Kosovo, Serbia, Albania) and others (e.g., Switzerland, Norway, UK). *Restrictions* on the integration of non-members and exiting countries are chief determinants of the union's size and scope and are necessary for the union to enhance cooperation (Proposition 6) and be robust to preference shocks (Proposition 8). Our theoretical results therefore explain why the EU has not allowed the UK to "cherry-pick" the post-exit integration with the union, as some proponents of BREXIT argued before the referendum and during the heated UK-EU negotiations. Moreover, countries that prefer tighter exit restrictions are those with high integration preferences (Proposition 9), as they are to lose the most. This result is in line with the tough stance of core EU countries, like France, the Netherlands, and Germany, in the post-BREXIT negotiations with the UK.[5] We then study enlargement alongside non-member

[4]In Appendix B, we argue why in our model, the earlier members of the union like Germany, France and Netherlands correspond to high types (with a higher preference for integration), while more recent members, current candidates as well as the United Kingdom who have left the union are low type countries (with a lower preference for integration). Moreover, we describe how the two main transformations, enlargement to periphery and the introduction of flexible integration, were accompanied with deeper integration across the union.

[5]For example, in an interview with the *Financial Times* (7 May 2024), Friedrich Merz, the leader of the German center-right (CDU) party, acknowledged Germany's tough stance during the negotiations between the UK and the EU before and after the referendum.

integration and identify a new channel through which outside integration restrictions affect the union composition. When considering candidates, higher-type countries compare the benefits of integrating with more countries to the possibility of a lower union policy induced by the lower median after enlargement. If non-member integration is not restrictive, these countries have a stronger incentive to reject candidates, nudging them to integrate as non-members, as they can still reap the benefits of integration without giving candidates voting power. Thus, integration must be restricted to satisfy incentive constraints not only of the candidates but also of the initial members (Proposition 7).

Two-Tier Unions. We conclude by studying tiered unions, where a subset of members integrate further, a controversial issue in Europe nowadays. We differentiate between endogenous tiers, where some countries adopt similar/identical policies (laws/regulations) to integrate further without explicit enforcement, and *tier agreements*, where countries make binding and enforceable commitments for deeper integration. Both of these methods are used currently by the EU, under Open Method of Cooperation (OMC) and Enhanced Cooperation Agreements (ECA).[6] We show that which method is effective depends on countries' preferences and externalities on the policy domains, which explains why the EU has institutionalized two separate governance mechanisms in its Treaty. When benefits of integration require reciprocal adoption of the same policies, countries with similar preferences over integration form endogenous tiers without a formal agreement; besides, structured tier agreements cannot improve welfare (Proposition 10). In contrast, if there are considerable externalities, endogenous tiers cannot exist, and binding agreements between countries with similar preferences are welfare enhancing (Proposition 11).

Related Literature. Our paper relates to works examining international unions' formation, functions, and enlargement, focusing mainly on the EU. Alesina, Angeloni, and Etro (2005) study the formation and enlargement of international unions, building on the optimal country size literature pioneered by Alesina and Spolaore (1997) and Bolton and Roland (1997) and theories of federalism (Oates (1972)).[7] As Alesina, Angeloni, and Etro (2005), we consider equally sized countries that solely differ in their preferences for integration. However, we modify, extend, and reformulate their framework in multiple directions. First, while our model nests theirs, we focus on a setting where countries' actions are strategic complements rather than investments in a public good, as many EU policies entail standardization

[6]Examples of OMC include the European Employment Strategy and the European Higher Education Area (Bologna Process). Examples of ECA are the agreements on the Eurozone and the Unitary Patent Agreement.

[7]See Bolton, Roland, and Spolaore (1996) and Oates (1999) for thorough literature reviews on the size distribution of nations and fiscal federalism, respectively. Alesina and Spolaore (2005) offers a book-length treatment on the theory of the size of nations.

and harmonization in product and capital markets and are subject to market size effects. Besides, the EU budget is small, the role of the EU in the provision of standard public goods provision is negligible (e.g., Alesina, Tabellini, and Trebbi (2017)), and the initial discussions of including defense in the EU were abandoned. Second, we allow non-members to integrate with the union to explore the tradeoffs of EU's trade and investment treaties with candidates (Albania, Turkey, Serbia, Montenegro, and North Macedonia), currently integrating with the EU in various domains, such as trade, legal system (Copenhagen Criteria) and product standardization. Allowing for non-member integration and enlargement provides insights into the EU's arrangements with exiting countries. Third, we study multi-tier unions, formalizing the ongoing debate of allowing some countries for deeper integration in some domains.

Our paper also relates to works that explore certain aspects of international unions. Harstad (2006) shows that allowing for "inner clubs" enhances integration if heterogeneity is considerable and externalities small. Berglof, Burkart, Friebel, and Paltseva (2008) show that in a framework where cooperation requires unanimity, the threat of an "inner club" by higher-type ("core") members spurs contributions from lower-type members (periphery), strengthening the union's cohesion. Our different modeling approach allows differentiating between tiers formed endogenously and through agreements. Kobielarz (2024) extends the Alesina, Angeloni, and Etro (2005)'s model to study exit, allowing for transfers and exit costs. Fiscal transfers can prevent inefficient exits, and post-exit arrangements where the former member partially contributes to the public good in return for limited spillovers improve efficiency. Our results on exit are complementary, showing that similar mechanisms are present in a setting with different assumptions (strategic complements vs. substitutes) and interpretations (integration vs. investment in a public good) and without transfers (as in reality). Our result on the necessity of restriction on non-member integration and characterization of efficient restrictions echo similar mechanisms in Bolton and Roland (1996, 1997), who build a model with two countries deciding whether to separate.[8]

Our model builds on the extensive literature on the theory of clubs, initiated by Buchanan (1965). Roberts (1999), the closest to ours, develops a dynamic voting setting with an endogenous electorate under strategic complementarities.[9] We take a less abstract viewpoint,

[8]Other related papers that, however, focus on trade and redistribution rather than integration include Crémer and Palfrey (1996), which considers a model where districts of heterogeneous populations decide between centralization and decentralization; Casella (2001) explores coalition formation in a spatial club model; Alesina, Spolaore, and Wacziarg (2000) add border costs to Alesina and Spolaore (1997) to show that globalization yields a higher number of countries; Casella (2005) studies a two-region (core and periphery) model with heterogeneous countries deciding on redistribution; Gancia, Ponzetto, and Ventura (2020) study the effects on trade, income distribution, and welfare of economic unions differing in size and scope. Abramson and Shayo (2022) consider a "core-periphery" model in which countries with endogenously heterogeneous identities and preferences decide over policies.

[9]Our focus on international unions, alongside our core assumption of strategic complementarities, can be

as our objective is to characterize various types of international integration and understand the trade-offs of their formation, expansion, and even shrinkage, allowing, realistically, for partial outside integration. Finally, since, as we rely on the results of the characterization of extremal equilibria, our paper connects with the literature on supermodular games (Topkis, 1979; Vives, 1990; Milgrom and Roberts, 1990).

Outline. Section 2 sets up the theoretical framework. Section 3 first characterizes the equilibria under rigid union, flexible union, and integration without a union. It then compares the equilibrium union formation and enlargement policies under the different methods, exploring which policies are preferred by what country types. Section 4 extends the model allowing for non-member integration. In this richer setup, we study different extensive form games to gain insight into various important issues regarding international unions, including enlargement with non-member integration, exit, and tiered unions. In Section 5, we summarize and discuss avenues for future research.

2. Model

2.1. Preferences

$U = \{1, 2, ..., |U|\}$ denotes the finite set of union members. Each country has a type $\gamma_i \in \mathbb{R}_+$, measuring the strength of its preference for integration, where $\gamma_1 < \gamma_2 < ... < \gamma_{|U|}$ and $\gamma = \{\gamma_i\}_{i \in U}$.[10] $t_i \in \mathbb{R}_+$ denotes the action of country i; $t \in \mathbb{R}_+^{|U|}$ denotes the action profile. Country i's payoff depends on its action, the actions of other countries, and their type; the payoff is $u_i(t, \gamma_i) \equiv u(t_i, t_{-i}, \gamma_i)$, where $u : \mathbb{R}_+^{|U|+1} \to \mathbb{R}$.[11] We start with the main assumptions on the payoffs.

Assumption 1. *The utility function, u, satisfies the following conditions:*

(i) u is increasing in γ_i and satisfies increasing differences in γ_i and t_i and γ_i and t_{-i}.

(ii) $u(t_i, t_j, t_{-ij}, \gamma_i)$ is increasing in t_j and strictly increasing in t_j whenever $t_j < t_i$.

The first part ensures that high-type countries (preferring stronger integration) benefit more from other countries' higher actions. The second part ensures that integration is beneficial: countries prefer other countries to choose higher actions. Actions can be thought

viewed as a bridge between Alesina, Angeloni, and Etro (2005) and Roberts (1999).

[10] We do not take a stance of the origins of integration preferences. For example, they could reflect elites' desire to solidify post-World War II peace by bringing countries' policy-making closer, as in the early decades of the European integration project. They could also capture citizens' ideologies, values, and beliefs.

[11] The first argument, t_i, denotes the county's integration level, while t_{-i} is the vector of other countries' integration levels. We also use $u(t_i, t_j, t_{-ij}, \gamma_i)$ to specify the effect t_j has on the utility of country i. As u is common across all countries, we assume $u(t_i, t_{-i}, \gamma_i) = u(t_i, \hat{t}_{-i}, \gamma_i)$ if $\hat{t}_{-i}$ is a permutation of t_{-i}.

of as investment/spending decisions on public goods, as in earlier research (e.g., Alesina, Angeloni, and Etro (2005), Kobielarz (2024)), or perhaps more realistically as efforts towards common harmonization policies in goods and services fostering trade, movement of labor across borders, product market standardization, homogenizing laws and regulations on banking and capital markets (Alesina, Angeloni, and Schkuknecht (2005)).

If one views actions as investments in a common public good, it is reasonable for payoffs to satisfy the strategic substitutes. This implies that u satisfies decreasing differences in t_i and t_{-i}. Indeed, the payoff function in Alesina, Angeloni, and Etro (2005) satisfies both Assumption 1 and the strategic substitutes condition.[12] The next assumption emphasizes the critical difference between our thinking of integration and earlier works.

Assumption 2. *u satisfies increasing differences in t_i and t_{-i}.*

Our setting's generality allows for various interpretations of the action t_i. If we view integration as a bilateral process where both countries must take certain measures to benefit, we can interpret t_i as the *intended* integration of country i. Higher t_i implies that the country is willing to integrate more with other countries, but this only happens if others reciprocate. Our model nests the case where the *effective* integration between two countries, i and j, is given by the minimum of intended integration, $\min\{t_i, t_j\}$; the payoffs in this case become $u(t_i, t_{-i}, \gamma_i) = \tilde{u}(t_i, \min\{t_i, t_{-i}\}, \gamma_i)$, where $\min\{t_i, t_{-i}\}$ is the vector obtained by replacing t_j with $\min\{t_i, t_j\}$ for all j. In Section 4, we explicitly define and study these preferences.

Alternatively, action t_i may represent a country's unilateral steps to increase integration with another country, for example, policies and legislation facilitating the easiness of starting a new business, licensing, or the standardization of consumer safety provisions, which may benefit both countries but do not require reciprocation. Another example of action t_i would be investments in a public good that satisfies strategic complementarities, such as transportation investments. In such cases, an increase in t_j may increase $u(t_i, t_j, t_{-ij}, \gamma_i)$ even if $t_i < t_j$. In both cases, t_i does not necessarily determine how much country i integrates with another country j, as integration is jointly determined as a function of all countries' choices.

Assumption 3. *The utility function, u, satisfies the following conditions:*

(i) For all γ_i and t_{-i}, $u(0, t_{-i}, \gamma_i) = 0$.

(ii) For each γ_i, there exists a $\hat{t} > 0$ such that u is strictly decreasing in t_i for all $t_i > \hat{t}$.

(iii) u is strictly concave in t_i.

The first part of this technical assumption normalizes the payoff of no integration to zero. The second part ensures that the benefits of integration cannot be infinite, as the costs will

[12] Payoffs in Alesina, Angeloni, and Etro (2005) are given by $u(t_i, t_{-i}, \gamma_i) = \gamma_i H(t_i + \beta \sum_{j \neq i} t_j) - t_i$ for some positive β and strictly concave H, where t_i represents the investment of country i to the public good.

eventually dominate, guaranteeing the existence of an equilibrium.[13] The third part ensures that each country has a unique best response given the integration of other countries.

2.2. Integration Protocols

We define and analyze three main methods/protocols of integration: *(i) rigid union*; *(ii) flexible union*; and *(iii) non-union integration*.

Rigid Union. In a *rigid union*, all countries integrate at the common integration level, and this level is determined by majority voting.[14] Formally, the *rigid union equilibrium policy* is the integration level $r^* \in \mathbb{R}_+$ that is preferred to any alternative r' by a majority of countries; in other words, r^* satisfies the *Condorcet Criterion*.[15]

Flexible Union. There is a minimum integration threshold in a *flexible union*. Still, countries can take steps for deeper integration, for example, in monetary policy and banking. The *flexible union equilibrium policy* is a lower bound b, determined by majority voting. Given b, countries voluntarily choose any action in $[b, \infty)$. A vector of integration levels $T(b)$ is a flexible union equilibrium under threshold b if $T_i(b)$ is the optimal integration given $T_{-i}(b)$:

$$T_i(b) \in \arg\max_{t_i \geq b} u_i(t_i, T_{-i}(b), \gamma_i) \quad \forall i. \tag{1}$$

While there might be multiple equilibria under b, Proposition 1 shows that there is a highest equilibrium, which is Pareto dominant. Assuming the highest (lowest) equilibrium will be played, countries vote over b to maximize $u_i(T_i(b), T_{-i}(b), \gamma_i)$, and an integration level b^* is the flexible union equilibrium policy if it satisfies the *Condorcet Criterion*, where countries evaluate their payoffs in the highest (lowest) possible equilibrium under b. The lower bound represents core union issues. In the EU, for example, a "single market" with free movement of goods, services, capital, and labor. The flexible union allows further integration, like joining a currency area or harmonizing patent protection.

Non-Union Integration. Countries can also integrate without explicit bargaining, negotiation, or centralized enforcement by independently and simultaneously choosing their integration levels. *Non-union integration* is the simplest and most flexible form of international cooperation and is equivalent to a flexible union with a bound exogeneously set

[13] Alternatively, we can bound the action space, $t_i \in [0, 1]$, interpreting the maximum, $t_i = 1$, as becoming a single country.

[14] When $|U|$ is even, it is possible that two integration levels get the same votes. In these cases, we break the tie in favor of the lower policy.

[15] If a country is indifferent between two policies, we assume that the country breaks ties in favor of the lower policy.

at $b^* = 0$. An action profile t^* is a *non-union integration equilibrium* if for all $i \in U$, t_i is the optimal integration given t_{-i}, *i.e.*, $t_i^* \in \arg\max_{t_i} u_i(t_i, t_{-i}^*, \gamma_i)$. In what follows, we mainly concentrate on rigid and flexible unions and use non-union integration outcomes as a benchmark to evaluate the effect of a union.

3. Equilibrium Characterization and Comparison

3.1. Equilibrium

We first characterize the equilibria of the three protocols. For the rest of the paper, m denotes the country with the median type (if $|U|$ is even, $m = |U|/2$).

Proposition 1. *The following statements characterize the equilibria in each protocol:*

1. *In a rigid union, the most preferred policy of the median country is the Condorcet winner:* $r^* = \arg\max_r u(r, \ldots, r, \gamma_m)$
2. *In a flexible union, given a bound b, there are highest and lowest equilibria,* $\overline{T}(b, \gamma)$ *and* $\underline{T}(b, \gamma)$. $\overline{T}(b, \gamma)$ *is the Pareto dominant equilibrium. Under the highest (or lowest) equilibrium selection, the preferred policy of the median is the Condorcet winner:* $b^* = \arg\max_b u(\overline{T}(b, \gamma), \gamma_m)$. *Median country and all lower type countries choose* b^* *as their integration level at* $\overline{T}(b, \gamma)$.
3. *In non-union integration, there are highest and lowest equilibria,* $\bar{t}(\gamma)$ *and* $\underline{t}(\gamma)$, *that are increasing in countries' type* γ_i *(for all i).* $\bar{t}(\gamma)$ *is the Pareto dominant equilibrium.*

In the rigid union, as all countries integrate at the same rate, countries' preferences over the union policy satisfy single crossing: countries with higher preferences always prefer higher union policy. The characterization of the Condorcet winner follows from the median voter theorem with single-crossing preferences (Gans and Smart, 1996). In non-union integration and flexible union (given $b \geq 0$), the characterization of the extremal equilibria is a consequence of increasing differences and follows from results in supermodular games (*e.g.*, Topkis (1979)). We then show that the equilibrium integration choices in a flexible union are increasing in b, and show that countries' preferences satisfy single crossing in the union policy, completing the characterization. For the rest of the paper, we concentrate on the highest equilibrium for flexible unions.

3.2. Comparing Integration Protocols

We start with an example that illustrates how countries' decisions and the equilibrium policies under different integration protocols depend on the countries' preferences, in particular, the strength of complementarity.

Example 1. *There are three union members, $U = \{1, 2, 3\}$. The utility function is*

$$u_i(t, \gamma_i) = \gamma_i \sum_{j \neq i} t_i t_j - t_i^3. \tag{2}$$

The benefit of integration for country i of integrating with j is given by the product of the integration levels and the type, $\gamma_i t_i t_j$; the cost equals t_i^3. Here, γ_i denotes the strength of complementarity and the preference for integration, as $\frac{\partial u_i}{\partial t_i \partial t_j} = \gamma_i$. We fix the types of countries 1 and 2 as $\gamma_1 = 0.9$, $\gamma_2 = 1$; let $\gamma_3 > 1$ vary.

Plugging in $\gamma_m = \gamma_2 = 1$, Proposition 1 implies that the rigid union policy is determined as follows

$$\arg\max_r r(r + r) - r^3, \tag{3}$$

which is maximized at $r = 4/3$, the rigid union equilibrium policy. In the flexible union case, by Proposition 1, countries 1 and 2 will choose policy b as their integration level. Therefore, the payoff of country 3 is

$$u(t_3, b, b, \gamma_3) = \gamma_3 2bt_3 - t_3^3, \tag{4}$$

maximized at $t_3(b, \gamma_3) \equiv (\frac{2}{3}\gamma_3 b)^{1/2}$, assuming $t_3(b) \geq b$, which we will verify later. Unlike the rigid union case, the union policy affects the integration of country 3 only indirectly, and its effect on $t_3(b, \gamma_3)$ gets stronger as complementarity (parameterized by γ_3) is higher. Plugging in $\gamma_m = \gamma_2 = 1$, the flexible union policy is determined as follows:

$$\arg\max_b b\,(b + t_3(b)) - b^3. \tag{5}$$

Comparing Equations 3 and 5, the effect of flexibility on the union policy depends on the comparison of the direct effect in the rigid union and the indirect effect through $t_3(b)$ in the flexible union. Figure 1 plots the equilibrium policies in flexible and rigid unions and the integration of country 3 as a function of the complementarity, γ_3.

When γ_3 is close to 2, the complementarity of integration is relatively weak and the indirect effect on of raising b is lower than the direct effect in rigid union. Therefore, $b < r$. As γ_3 increases, country 3 responds more strongly to an increase in b, which in turn increases the incentive of the median country to choose a higher policy. Therefore, if complementarity for country 3 is strong enough, the indirect effect becomes stronger and integration under flexible union is higher than integration under rigid union.

We now explore which countries prefer flexible to rigid unions. First, all countries prefer a flexible union with policy to a rigid union with policy with the same policy, as in flexible union, they can choose their integration level at the rigid union, and all other countries must choose a (weakly) higher integration. Moreover, as b^* is the most preferred policy of

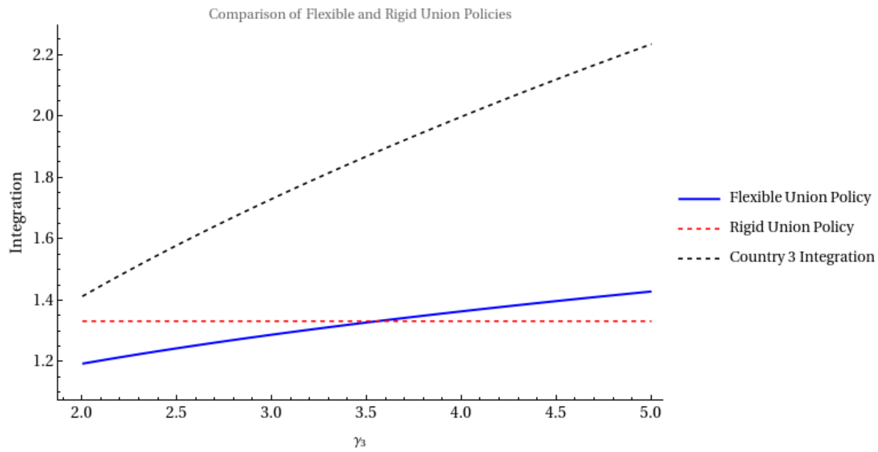

Fig. 1. **Flexible and Rigid Union Equilibrium Policies.** The red line denotes the equilibrium policy under rigid union, which does not depend on γ_3. The Blue and Black curves denote the flexible union policy and the integration of country 3 as a function of γ_3.

the median country, the median always prefers flexible to rigid union. When $b^* \geq r^*$, the median, and therefore any country with a higher type, benefits from the higher integration of the flexible union. Conversely, when $b^* < r^*$, the median and countries with lower types prefer flexible union, as they can integrate at b^* instead of higher r^*.

Proposition 2. *A majority of countries prefer flexible unions to rigid unions.*

- *If $b^* \geq r^*$, then (at least) the median country and all countries with higher types prefer flexible union.*
- *If $b^* < r^*$, then (at least) the median country and all countries with lower types prefer flexible union.*

Example 1 and Proposition 2 demonstrate that under strategic complementarities, flexible protocols (i) can increase the union-wide integration and (ii) be preferred by the higher-type countries, explaining an essential regularity in the adoption of flexible integration policies such as Schengen and Eurozone. It is instructive to compare the theoretical results when actions are complements with the model predictions when actions are strategic substitutes, as this has been the basic assumption of most of the literature.

Strategic Substitutes. First, with strategic substitutes, when higher-type countries choose higher actions, this causes the median country to choose a *lower* integration level, as effectively, the median and lower-type countries "free-ride" on higher-type members' investments. Moreover, a higher integration policy leads to *lower* actions from the higher-type countries, further lowering the incentives of the median to choose a higher policy. The negative indirect effect cannot dominate the direct effect of raising the rigid union policy. Both channels reduce the flexible union equilibrium policy compared to the rigid union, $b^* < r^*$; flexibility (always) causes low-type countries to free-ride on the contributions of higher-type countries. Hence, low-type countries prefer flexible unions when countries' actions are investments in a common public good, while higher-type countries prefer rigid unions.[16]

Strategic Complements and The Experience of the EU. Conversely, when countries' actions are complements, as in our setting, a majority always prefers a flexible to a rigid union. But, it is not *ex-ante* clear which countries select flexible arrangements. If the policy is lower under flexible unions, countries with lower types prefer it. This result follows from the fact that low-type countries prefer a flexible union with a low bound, as they pay a cost from integrating to a higher-than-desired level. In contrast, when complementarities are strong and integration is higher under a flexible union, high-type countries prefer it, as they can (and will) integrate further while maintaining the benefits of the low-type countries' integration. Not only does Proposition 2 reverse the surprising conclusion reached in Alesina, Angeloni, and Etro (2005), but is more in line with the dynamics of the EU: "core" countries (like Germany and France) have promoted flexible union, allowing Eastern and Southern European countries to join, while at the same time integrating themselves more (via monetary unification).[17] Public goods models with strategic substitutes cannot easily explain the co-evolution of deeper European integration and enlargement to periphery.

3.3. Enlargement to Periphery

We now endogenize the union U to study enlargement, a major feature of the EU and other Unions. N denotes the set of countries, $I \subset N$ initial members, and $C = N - I$ candidates. We analyze the following extensive-form game:

1. Candidate countries decide whether to apply for membership.
2. Each union member decides whether to admit or reject each candidate.
3. Equilibrium union U is the initial union plus the countries admitted unanimously.[18]

[16] Alesina, Angeloni, and Etro (2005) provide a formalization of this mechanism in their Proposition 2, writing: *"For instance, a surprising result emerges if, in a rigid union with a uniform provision of public goods, countries are allowed individually to add extra expenditure. One may think that countries with strong preferences on public spending would support such a reform: in reality, these are the only countries that may oppose the reform and prefer the rigid union. The reason is that this reform would reduce the uniform provision chosen in political equilibrium at the union level so as to rely on the extra-provision of individual countries."*

[17] In Appendix B, we extend our discussion of how the model maps to the EU and describe the main transformations EU has experienced.

[18] We focus on enlargement with unanimity since this is the paradigm the EU has followed since its inception. This does not play an important role in our analysis as our results are focused on characterizing which countries would support the enlargement and the effect of enlargement on integration.

4. Union members decide the equilibrium policy r^* of with majority voting. Members integrate at r^*, non-members do not integrate with the union.[19]

To consider enlargement, motivated by the history of the EU, we assume that the initial members have higher integration preferences than candidates. This assumption appears reasonable, as the process of European integration started with six "core" countries in a limited domain, and, over time, the union admitted new members while integration deepened considerably. Moreover, in Proposition 14 in Appendix D.2, we show that during the union formation phase, all countries with a higher type than the median country are union members, which offers another justification for this assumption. We keep this assumption for our results on enlargement (Propositions 3, 4, 5, and 7).

Assumption 4. *Initial union members are more numerous and have higher types than candidates. That is,* $|C| < |I|$ *and if* $i \in I$ *and* $j \in C$*, then* $\gamma_i > \gamma_j$*.*

Before moving to the results, we define a property that ensures the preferences satisfy a mild complementarity requirement. Let u_{t_i} denote the partial derivative of u with respect to t_i, if it exists. We say that integration is *minimally complementary* if $u_{t_i}(t_i, t_j, t_{-ij}, \gamma_i) > u_{t_i}(t_i, t'_j, t_{-ij}, \gamma_i)$ when $t_j < t_i \leq t'_j$. In other words, integration is more beneficial for country i when another country increases its integration from below i's to above i's integration level.

Proposition 3. *Suppose that* u *is continuously differentiable and integration is minimally complementary.*

1. *There exists a* $\hat{\gamma} < \gamma_m$ *such that integration increases after the admission of* C *to the union if and only if the median type in* $I \cup C$ *is higher than* $\hat{\gamma}$*.*
2. *Initial members with above-median types prefer enlargement whenever integration increases after enlargement.*

Enlargement to the periphery affects equilibrium integration through two main channels: an increase in the number of countries integrating and a fall in the type of the median country due to the admission of lower-type candidates. Proposition 3 establishes that the two consequences have countervailing effects under strategic complementarities; the first increases the equilibrium integration, while the second decreases it. Hence, integration increases after enlargement towards the periphery if the decrease in the median type is small enough.

[19] Formally, $t_i = r^*$ if $i \in U$ and $t_i = 0$ otherwise. We do not allow for non-member integration and focus on rigid unions to compare our results with Alesina, Angeloni, and Etro (2005). Although in their setting, non-members can invest in producing a *private* good, members do not benefit from the investments made by non-members. Likewise, non-members do not benefit from the union's public good. We extend Proposition 3 by allowing integration of non-members and flexible unions in Appendix D.6 (see Proposition 16).

It is instructive to compare this result with corresponding results under strategic substitutes (*e.g.*, Alesina, Angeloni, and Etro (2005)). With substitutes, a larger union causes all countries (including the median) to prefer lower actions, amplifying the effect of the lower type median following enlargement. Therefore, enlargement to the periphery is *always* accompanied by lower integration across the union, supported by the lower-type countries.

Proposition 4. *Suppose that* u *satisfies Assumptions 1, 3 and decreasing differences in* t_i *and* t_{-i}*. Then the rigid union equilibrium policy decreases after enlargement to periphpery.*

Our analysis shows that, the type of the post-enlargement median is not the main determinant of a successful enlargement under strategic complementarities. There are two main roadblocks to enlargement, depending on the post-enlargement union policy. First, high-type members may block entry as they do not want a reduced integration. This is the (direct) mechanism often discussed in media and policy circles when the EU expanded in the South in the 1980s and Eastern Europe in the 2000s. Besides, many make this point when discussing future EU enlargement in the Western Balkans (Albania, Serbia, Kosovo, North Macedonia), Turkey, and more recently, Ukraine and Moldova.[20] Second, low-type members may block entry as they do not want increased post-enlargement integration. This novel mechanism is not much discussed, although likely present, as many non-core EU countries with arguably low preferences for deeper integration, like Bulgaria and Hungary, are expressing concerns for its future expansion in the Balkans and the East. Besides, periphery countries in the South were not keen proponents of the EU's enlargement in Eastern Europe in the 2000s.[21]

As more countries integrate after enlargement, it has an additional, and unambiguously positive effect on payoffs. Whether a country supports enlargement or not is then determined by its potential loss from the changed union policy, and its benefit from integrating with a larger union.

Proposition 5. *Suppose that* u *is continuous. If the union policy after enlargement is close enough to the union policy in the initial union, then all countries are in favor of enlargement.*

Taken together, these findings indicate that complementarities could offset the impact of a decrease in the median type after expansion to the periphery, potentially easing the enlargement process. This is also consistent with various EU enlargement rounds when the union admitted many countries, as conditional on the post-enlargement median, admitting more

[20]For example, former French President Francois Mitterand was initially skeptical of the EU's enlargement in the South, expressing openly his fear that Greece, Portugal, and Spain's accession would prevent deeper integration. Likewise, former EU Commission President Romano Prodi was concerned that accepting more "peripheral" countries would leave the EU a "simple" enhanced free trade area.

[21]We provide explicit examples of both mechanisms in Examples 5 and 6 in Appendix D.3.

countries may balance the tradeoffs across countries with different preferences for integration.[22] This notion gains further credence considering the minimal monetary contributions from peripheral members during the EU's enlargement, suggesting that core members with higher types might be disinclined to support expansion. However, the presence of complementarities fosters integration, elucidating why core members were enthusiastic advocates for enlargement.

Taking Stock. We believe our results (in particular Propositions 2 and 3) align with the EU's history, where enlargement to periphery occured alongside deeper integration across more domains. Quite often, "core" countries with high preferences for European integration were the big promoters of the Union's enlargement and flexible policies. Therefore, our model focusing on integration with strategic complementarities provides a better description of the nature of the EU and makes more realistic predictions about the evolution of integration and enlargement. In the next section, we switch gears and study certain issues that have been absent in the previous literature, such as non-member integration and exit.

4. Moving Forward: Integration with Non-Members, Enlargement, Exit, and Two-Tier Unions

4.1. Preferences

For the analysis of non-member integration, entry and exit, and two-tier unions, we make the following assumption on payoffs.

Assumption 5. *$u(t_i, t_{-i}, \gamma_i) = \sum_{j\neq i} \tilde{u}(f(t_i, t_j), \gamma_i) - c(t_i, \gamma_i)$ where $f : \mathbb{R}^2_+ \to \mathbb{R}_+$ is increasing in both arguments, $\tilde{u}$ is concave, twice continuously differentiable, satisfies increasing differences and is strictly increasing in both arguments; c is strictly convex, twice continuously differentiable, satisfies strictly decreasing differences, is strictly increasing in t_i and decreasing in γ_i.*[23]

Under Assumption 5, the integration payoffs are additively separable across countries. The payoff of country i from integrating with country j depends on the integration levels of the two countries through $f(t_i, t_j)$, which denotes the effective integration between i and j. The total payoff of i is the sum of its integration with all countries. The additive separability of payoffs makes the analysis tractable without much sacrificing generalizability. Moreover, it allows us to make more precise assumptions about the nature of integration through f.

[22] See Appendix D.3 for an extended discussion of this effect.

[23] These assumptions on $\tilde{u}$ and c ensure that the utility function u is compatible with Assumptions 1, 2 and 3.

We focus on bilateral integration for our analysis of non-member integration and exit.

Assumption 6. $f(t_i, t_j) = \min\{t_i, t_j\}$.

Under Assumption 6, the benefit of country's i integration with j depends on $\min\{t_i, t_j\}$, countries' common integration (aligned with our interpretation of t_i as the intended integration): both countries must implement various integration measures for each of them to reap the benefits. We relax this assumption when we study multiple-tier unions in Section 4.5 and show how it affects the structure of tiers in unions.

4.2. Integration with Non-Members

Motivation. Assuming that non-members do not integrate at all with the union was a useful starting point, as it features in much of earlier research (e.g., Alesina, Angeloni, and Etro (2005); Roberts (1999)). However, international unions allow for integration with outsiders, and, in some domains, such policy integration is considerable. Article 238 of the Treaty of Rome specifies: *'The Community may conclude with a third country, a union of states, or an international organization agreements creating an association embodying reciprocal rights and obligations, joint actions and special procedures."* The EU has treaties with non-members covering multiple areas. For example, Turkey and the EU have a *Customs Union*, an enhanced trade agreement allowing imports-exports to flow across the border freely. Norway has access to the EU's single market, with some exceptions on agriculture, fishing, and food. Candidates like Serbia, Albania, Montenegro, the Republic of North Macedonia, and Turkey integrate with the EU on 35 "accession chapters", covering company law, public procurement, energy, taxation, financial services, and consumer and health protection, among others. In each area of the *acquis*, candidates are *"required to adapt their administrative and institutional infrastructures and to bring their national legislation into line with EU legislation in these areas."*

Setting. Outside integration allows countries with preferences far from the union's policies to integrate as non-members. Opting out of the union but without being totally excluded from integration may be optimal for some countries, balancing the costs and benefits under (relatively) low preferences for integration. Similarly, it the union members can benefit from integration with non-members who are not willing to integrate as much as members. Reasonable examples include, we believe, Switzerland and Norway, which have been integrating with the EU in various domains without joining, as they prefer keeping national policy-making in some areas. Besides, Western Balkan countries, like Serbia and Albania, with arguably low preferences for integration stemming from sizable economic differences and skepticism towards the EU, may want to integrate as non-members to grasp some benefits. To analyze how the union's policies toward non-members affect its stability, the integration of members, and the union's integration with non-members, we study the following union formation game's Subgame Perfect Equilibrium (SPE).

1. Countries (in N) decide between joining the union or integrating as non-members. Countries who join form the union U.
2. Members decide the equilibrium union policy b with majority voting.
3. Countries choose their actions; where $t_i \in [b, \infty)$ for members and $t_i \in [0, \overline{o}]$ for non-members.

$\overline{o}$ is the non-member integration bound and denotes the maximum integration level non-member countries can choose. We now study its effect on an equilibrium union's size and policies.

Analysis. We start by analyzing the relationship between non-members and the union. A union is *ineffective* if the integration and payoffs of all countries under non-union integration are weakly higher than their integration levels and payoffs in the union. An ineffective union does not increase integration, its primary function. Moreover, whenever a union is ineffective, all members are indifferent between joining or integrating as non-members. The following proposition shows that imposing restrictions on the degree of integration among non-members is necessary for an effective union.

Proposition 6. *Let U be an equilibrium union with union policy b^*. If $\overline{o} \geq b^*$, then U is ineffective.*

To prove Proposition 6, we first note that when non-member integration is unrestricted, countries with integration preferences below the median are not incentivized to join. By joining, their integration jumps to the union policy. Hence, these countries prefer to integrate as non-members, allowing them to choose their most preferred integration, which is below the union policy. As a result, no union can include a country whose type is below the median. The only possible equilibria are two-country unions, where the lower type member chooses the union policy, which equals its non-union integration level. However, such unions are ineffective, as they do not increase integration compared to non-union integration. Proposition 6 demonstrates a chief result: policies towards non-members are essential determinants of the union's size and scope.[24] However, when the non-member integration level is restricted

[24] Our result on the necessity of non-member integration restriction echoes the analysis of Bolton and Roland (1997) on barriers of trade: *"An unpleasant implication of our analysis is that barriers to trade and factor movements between the European Union and neighboring non-Union states play a role in cementing*

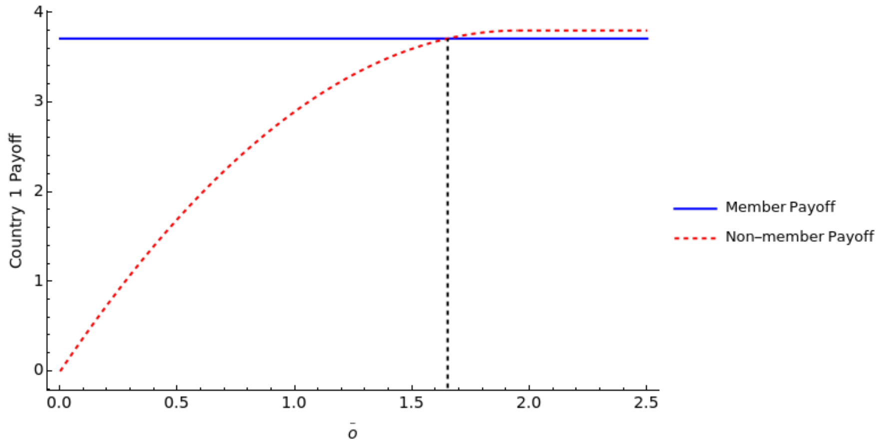

Fig. 2. **Country 1 Payoff.** The red line denotes the payoff of country 1 obtains when integtaring as a non-member. The blue line denotes the payoff of country 1 when it becomes a member. Black dashed line indicates the non-member integration level that equalizes these payoffs.

(as in the EU), the lower-type countries are incentivized to join the union, integrating more. The following example illustrates the mechanisms.

Example 2. *There are four countries,* $N = \{1, 2, 3, 4\}$ *with the following types:* $\gamma_4 = 4$, $\gamma_3 = 3.5$, $\gamma_2 = 1.5$ *and* $\gamma_1 = 1.3$. *For simplicity, we will consider the rigid union case. The utility function is:*

$$u_i(t_i, t_{-i}, \gamma_i) = \gamma_i \sum_{j \neq i} \min\{t_i, t_j\} - t_i^2 \tag{6}$$

We will compute when country 1 *will join the union. If country* 1 *integrates as a non-member,* $U = \{2, 3, 4\}$ *the median is country* 3. *As in equilibrium country* 1 *will integrate less than the union, country* 3 *will solve the following problem to determine the union policy:* $\max_r 2\gamma_3 r - r^2$, *which is maximized at* $\hat{r}_3 = \gamma_3$. *If country* 1 *joins, then country* 2 *becomes the median, and will solve* $\max_r 3\gamma_2 r - r^2$, *which is maximized at* $\hat{r}_2 = 3/2\gamma_2$.

Thus, as a member, country 1*'s payoff is* $3\gamma_1 \hat{r}_2 - (\hat{r}_2)^2$, *while as a non-member, country* 1 *solves the following problem*

$$\max_{t_1 \leq \overline{o}} 3t_1\gamma_4 - (t_1)^2 \tag{7}$$

and chooses $t_1^* = \min\{3/2\gamma_4, \overline{o}\}$. *The member and non-member payoffs of country* 1 *as a function of non-member integration restrictions is given in Figure 2. The two curves cross at* $\overline{o} = 1.65$, *which is the integration level that makes country* 1 *indifferent between integrating as a non-member and joining the union. Therefore, non-member integration must be restricted enough* ($\overline{o} < 1.65$), *to make* $U = \{1, 2, 3, 4\}$ *an equilibrium.*

the Union. In the absence of such barriers, a country would be less willing to join the Union if it can obtain most of the economic benefits of the Union by staying out and not paying the political costs in terms of loss of sovereignty."

As Example 2 illustrates, a restrictive policy for non-members may result in a larger union with less integration. In contrast, a less restrictive policy allows for more integration with non-members but results in a smaller union. In Proposition 15 of Appendix D.5, we characterize the efficient non-member restrictions and show that it is without loss of optimality to restrict attention to non-member integration bounds that make the lowest type member country indifferent between becoming a member and integrating as a non-member.

4.3. Enlargement with Non-member Integration

We now extend our analysis to consider *jointly* enlargement and non-member integration. As non-members do not have voting rights, integrating with them keeps integration high without moving the decisive median. Therefore, when enlargement to the periphery is accompanied by lower integration, non-member integration is desirable from the members' viewpoint, especially those favoring high integration. This may explain why the EU's "core" countries have historically favored expanding the number and depth of deals with non-members. For example, Germany has been historically a proponent of deepening ties with Turkey.

Motivation. Studying enlargement and non-union integration at the same time allows for a more in-depth analysis of both past and present mechanisms and considerations. First, all current EU members who joined in the various enlargement rounds had considerable integration agreements with the EU (or the European Economic Community, its predecessor) prior to becoming members. For example, Greece signed a customs union agreement in 1961, 20 years before joining, while Portugal had a free trade agreement since 1972, 14 years before entering. Likewise, Eastern European countries started signing integration agreements with the EU more than a decade before joining in 2004. Second, all candidate countries today have integration agreements with the EU and can maintain them without joining.

Setting. To study enlargement together with non-member integration, we analyze the extensive form game we introduced in Section 3.3, but we modify how countries choose their integration levels as follows. If $i \in U$, then i can choose an integration level $t_i \in [b^*, \infty]$, while if $i \notin U$, then i can choose an integration level $t_i \in [0, \overline{o}]$.

Analysis. Proposition 15 shows the importance of incentivizing low-type (potential) members. When we consider enlargement with non-member integration, an additional force limits enlargement. As discussed in Section 3.3, initial members may be resistant to admitting new members. When non-member integration is possible, a more permissive non-member integration restriction increases the payoffs of initial members from integrating with the candidates without admitting them to the union. For example, consider a high bound for outside integration. Suppose that post-enlargement, the union policy will decrease significantly. In

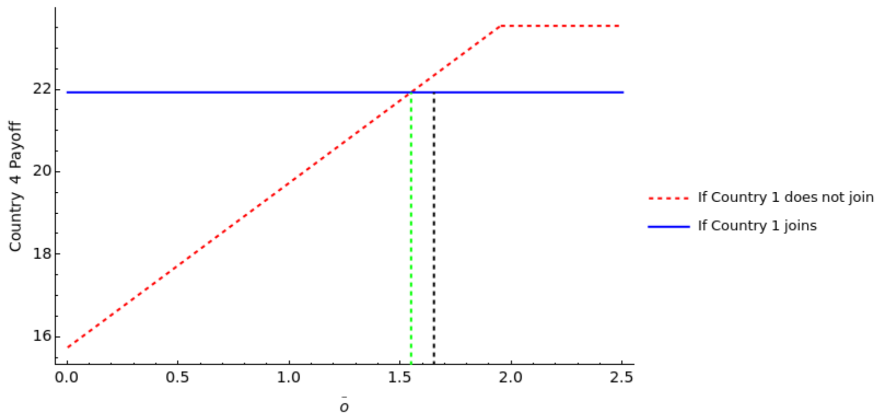

Fig. 3. **Country 4 Payoff.** The red line denotes the payoff of country 4 when country 1 integrates as a non-member under $\bar{o}$. The blue line denotes the payoff of country 4 when country 1 becomes a member. Green dashed line denotes the non-member integration restriction that equalizes these payoffs. Black dashed line indicates the necessary non-member restriction to make sure country 1 prefers to join the union.

that case, high-type members will reject the candidates, keeping union policy high (not giving non-members voting) while, at the same time, benefiting from integrating with them as non-members.[25] We now reconsider Example 2 in this setting.

Example 2 (continued). *In the setting of Example 2, let $I = \{2, 3, 4\}$ and $C = \{1\}$. When country 1 joins, country 4 integrates with all three countries at $\hat{r}_2 = 3/2\gamma_2 = 2.25$ and obtains a payoff of $3\gamma_4\hat{r}_2 - (\hat{r}_2)^2$. When country 1 integrates as a non-member, country 4 integrates with other two members at $\hat{r}_3 = \gamma_3 = 3.5$ and with country 1 at its equilibrium non-member integration level t_1^*, obtaining a payoff of $2\gamma_4\hat{r}_3 - (\hat{r}_3)^2 + t_1^*\gamma_4$. Figure 3 plots the payoff of country 4 as a function of $\bar{o}$.*

We have already showed that $\bar{o}$ must be less than 1.65 for country 1 to join the union. When country 1 is not a member, union policy is higher due to higher median. If non-member integration is not restricted, country 4 prefers to keep higher policy, while integrating with country 1 as a non-member. Therefore, for country 4 to admit country 1, non-member integration must be further restricted, $\bar{o} < 1.55$ (denoted by green line).

The following Proposition shows that for enlargement, the non-member integration bound must be lower than two bounds. The first bound is determined by the preferences of the candidates, $\bar{o}_C$, and is similar to the union formation case (denoted by the black dashed line in Figure 3). It assures that non-member integration is restricted to incentivize candidates to join. The second bound is determined by the preferences of the initial members, $\bar{o}_I$ (denoted by the green dashed line in Figure 3). It ensures that non-member integration is restricted so

[25] Example 8 in Appendix D.8 demonstrates how non-member integration restrictions may be necessary for high-type initial members to accept candidates.

that the initial members have an incentive to admit the candidates rather than integrating with them as non-members.

Proposition 7. *Consider an initial union I and possible enlargement $U = I \cup C$ with equilibrium policy b^*.*

1. *If U is not an equilibrium union under enlargement without non-member integration ($\bar{o} = 0$), then it is not an equilibrium under any $\bar{o} > 0$.*
2. *If U is an equilibrium union under enlargement without non-member integration, then there exists two cut-offs $\bar{o}_I$ and $\bar{o}_C$ such that enlargement to U is an equilibrium if and only if $\bar{o} \leq \min\{\bar{o}_I, \bar{o}_C\}$. Moreover, $\min\{\bar{o}_I, \bar{o}_C\} < b^*$, that is, non-member integration must be restricted for enlargement.*

4.4. Exit

Motivation. We now leverage the flexibility of our framework with non-member integration to study exit, an increasingly topical issue after BREXIT, the adversarial relationship of the EU with Hungary, and the rise of parties openly advocating for leaving the EU. To study exit, we allow for a preference shock, as in many countries, governments and people's views towards the EU and nationalism have changed over time.[26]

Setting. There is a union, $U = \{1, \ldots, |U|\}$, with $|U| > 2$. To ensure that all members prefer membership to leaving under their initial types, we assume that U is an equilibrium union with policy b in the union formation game studied in Section 4.2. A preference shock that reduces the type of a (random) country emerges with probability ρ; for example, an exogenous decline of citizens' preferences towards the EU, like the crisis in the European periphery or the election of a new administration skeptical of union integration, as in the UK. After the preference shock, the country can decide to exit the union and integrate up to $\bar{e}$, the upper bound of integration for countries leaving. If the country leaves, remaining members receive a cost $\kappa \geq 0$. This cost represents the time and effort needed to negotiate the exit deal, legislative amendments, and dealing with workers and companies from the leaving country. Formally, we analyze the SPE of the following game:

1. With probability ρ, a random country gets a shock, reducing its type to $\gamma_l \leq \gamma_1$.[27]

[26]See Guriev and Papaioannou (2022) for an overview of works on populism and euro-skepticism, and Guiso, Herrera, and Morelli (2016) and Alesina, Tabellini, and Trebbi (2017) on the role of cultural differences on EU's integration.

[27]We assume that the post preference shock type is homogeneous across countries for expositional ease. Our results go through as long as the post-preference shock type is (weakly) increasing in the initial country type.

2. The country facing the preference shock decides whether to exit the union.
   (a) If the country stays, it can choose any integration level above b.
   (b) If the country exits, it can choose any integration level below $\overline{e}$. The exit costs $-\kappa$ to all remaining members (where $\kappa \geq 0$).

A union U is *robust* under $\overline{e}$ if the preference shock does not lead the country to leave.

Analysis. In any SPE, the country with the preference shock stays in the union if and only if the payoff under staying is higher than the payoff under leaving.[28] The following proposition summarizes the core results linking the preference shock's size, the union's stability, and the associated restrictions on union's integration with non-members.

Proposition 8. *The following statements hold:*

1. *There is a $\tilde{\gamma}$ such that if $\gamma_l < \tilde{\gamma}$, then U is not robust under any $\overline{e}$.*
2. *If $\gamma_l \geq \tilde{\gamma}$, there exists $e(\gamma_l)$ where U is robust under $\overline{e}$ if and only if $\overline{e} \leq e(\gamma_l)$, i.e., the exiting country is restricted to have an integration level below $e(\gamma_l)$.*
3. *The exit restriction that makes the union robust, $e(\gamma_l)$, is increasing in γ_l.*

Proposition 8 shows that the union is robust when the upper bound on the integration of countries exiting the union is under $e(\gamma_l)$. This result appears intuitive since the former member that gets the negative preference shock becomes similar to a potential candidate: the exiting (candidate) country compares the payoff under exit (non-member) integration levels to staying (becoming) a member. Therefore, harsher restrictions contribute to the robustness (size) of the union. If the "deal" the union offers to the exiting (candidate) country is significantly worse than that of members, say restricting access to the single market and requiring licenses to conduct cross-border banking, the union becomes more cohesive (part (2)). The fierce negotiations and polemic between the UK and the EU following the BREXIT referendum are in line with these model corollaries. As many commentators and politicians argued at the time, it was unwise for the UK to expect a post-BREXIT deal with significantly *better* terms than the ones other non-EU-members have, as this would unravel the EU. Likewise, the UK's effort to "cherry-pick" parts of integration it favored during the negotiations failed. Ultimately, the BREXIT deal was far from a "soft" one where the UK would have kept most of the integration benefits.

As the magnitude of the shock increases (a lower γ_l), more restrictive non-member integration policies are necessary to keep the union intact (part (3)). If the preference shock is more likely to be stronger in the future, for example, due to the increasing euro-skepticism

[28] We focus on the equilibria where the country facing the shock stays in the union if it is indifferent between two options.

driven by immigration/refugee/economic stress, the post-exit deal with the Union will be much harsher to prevent exit. We coin the case where $\overline{e} \leq e(\gamma_l)$ as *exit restriction*, since in this scenario, the policy is strong enough to deter exit, resulting in a robust union.

Next, we consider countries' preferences towards exit, analyzing when the exit restrictions are Pareto optimal and/or implemented in equilibrium with majority voting. A country *prefers exit restriction* if its expected utility under a robust union with such restriction ($\overline{e} = e(\gamma_l)$) is greater than without it. A robust union with exit restriction results in a larger union, which is harmful after a preference shock as restrictions prevent exiting and keeping a desired level of integration but are beneficial otherwise. As the benefit of additional integration is higher for high-type countries, the following emerges:

Proposition 9. *If country i prefers exit restriction and $\gamma_j > \gamma_i$, then country j also prefers exit restriction.*

The following corollary characterizes when exit restriction is Pareto-improving and hence adopted under majority voting.

Corollary 1. *There is a cut-off country, $k(\gamma_l, \kappa)$, such that all higher-type countries prefer restrictions on exit. Exit restriction is Pareto improving if the country with the lowest type prefers it. Exit restriction is adopted in majority voting if the median member prefers it.*

Intuitively, if the remaining countries face higher costs after a former member exits, the benefits of a robust union are higher, while the cost of the restricted integration does not change. Therefore, with higher exit costs, countries prefer tighter exit policies. Conversely, if the preference shock is more significant, the benefit of a robust union for the remaining countries stays the same. At the same time, the cost of exit restriction is higher for the shock-hit country. The following corollary formalizes these points.

Corollary 2. *Fewer countries prefer exit restriction if the preference shock is greater or the exit cost is smaller.*

4.5. Tiered Unions

Motivation. We now analyze two-tier unions with heterogeneous integration across members, a controversial issue. The debate on the pros and cons of a multi-speed EU goes back to the enlargement in the South and the East and the decision to launch the euro, which has effectively created a two-tier Union. The discussion on multi-tier integration has intensified with the economic crisis in the European periphery and the contentious relationship with Hungary. Given rising geopolitical tensions have led politicians and administrators to argue for a two-tier (European) union. "Core" countries (with stronger preferences for integration) should be "allowed" to proceed further, deepening economic policy ties, for example, with stronger coordination of fiscal policy and debt mutualization.

In practice, the EU has already in place two governance mechanisms for further integration. The first is the *Enhanced Cooperation Agreement* (ECA) with binding controls, where a minimum of nine EU member states are allowed to establish advanced cooperation in an area within EU structures. The second is the *Open Method of Coordination* (OMC), formally initiated by the Lisbon European Council in 2000. OMC aims to achieve convergence towards EU goals in policy areas that fall under the partial or full competence of Member States by spreading best practices rather than by new EU legislation.[29] OMC examples are the European Employment Strategy, which coordinates EU members' reforms in labour markets and social policies (Tholoniat, 2010), and the European Higher Education Area (Bologna process), established to ensure comparability in the standards of higher-education qualifications (Veiga and Amaral, 2006).

Setting. Motivated by the EU's two distinct methods for deeper cooperation, we analyze when *tiers* (set of countries integrating at the same level) can emerge endogenously without binding agreements and when binding agreements can improve cooperation and welfare. We start by defining the two kinds of *tiers* unions can exhibit. A flexible union equilibrium admits *endogeneous tiers* if there is a set of countries J and an integration level $\tilde{t} > b^*$ such that $t_j = \tilde{t}$ for all $j \in J$. Endogenous tiers correspond to OMCs as they are self-enforcing: choosing $\tilde{t}$ is optimal for all countries in J even without enforcement through an agreement. Alternatively, a subset $\tau \subset U$ countries may sign an enforceable *tier agreement* (τ, b_τ) that guarantees that they integrate at some $b_\tau > b^*$, which correspond to enhanced cooperation agreements in the EU context.

Analysis. A tier agreement *increases welfare* if all tier members obtain weakly higher utility (with at least one strict) compared to the equilibrium without it. The following proposition demonstrates two results. First, when reaping the benefits of integration requires harmonizing policies, regulations, and standards (i.e., the adoption of the same measures, $f(t_i, t_j) = \min\{t_i, t_j\}$), tiers emerge endogenously for countries with similar preferences without the need for enforcement. Second, tier agreements cannot increase the welfare of its participants as a whole.

[29]Dehousse (2003) differentiates OMC from the earlier enhanced cooperation agreements. *"The absence of formal constraints: as the guidelines are formally devoid of any binding character, the peer assessment process is aimed at fostering learning processes. One counts on the "emulation between the Member States" to ensure the new strategy's success, rather than on classical community control mechanisms".*

Proposition 10. *Suppose that* $f(t_i, t_j) = \min\{t_i, t_j\}$*. If* i *integrates more than the union policy* ($\overline{T}_i(b^*) > b^*$)*, then there is an* ϵ *such that all* j *with* $t_j \in (\gamma_i - \epsilon, \gamma_i + \epsilon)$ *belong to the same endogeneous tier. Moreover, there does not exist a tier agreement that increases welfare.*

For the intuition behind the first part, consider two countries with consecutive types, $\gamma_i < \gamma_j$. When integration requires the adoption of same measures, the lower type country i already obtains their first best payoff given the integration levels of higher type countries. Therefore, country i will integrate at the same level regardless of higher type country's type or integration level and will not reciprocate further integration by country j. As a result, integrating more than country i is less beneficial for j, and as countries become more similar, the difference in types cannot compensate for the loss of integrating with one fewer country, forming the endogenous tier.

To grasp the intuition for the second part, note that to increase welfare, a tier agreement should increase integration of some countries compared to the initial equilibrium, which necessarily includes the lowest type member of the tier. However, as the lowest type member of any tier already obtains all the benefits it could get at their equilibrium integration level without the tier agreement, such tier agreement makes that country worse off. Thus, a tier agreement either does not increase the integration at all or makes (at least) the lowest type country worse off.[30]

We now contrast this result with the case where countries' efforts towards higher integration benefit other members regardless of their relative integration levels.

Proposition 11. *Suppose that* $f(t_i, t_j)$ *is strictly increasing in both arguments and continuously differentiable. Then, there are no endogenous tiers in equilibrium. Moreover, a tier agreement increases welfare whenever two countries have close enough types.*

The assumption that the externalities of higher integration are positive regardless of the relative integration levels yields two crucial results. First, higher integration will always be more beneficial for high-type countries; hence no endogenous tiers emerge. Second, in equilibrium, countries do not fully internalize their positive externalities on others. Countries of similar types can benefit from the enforcement power of tier agreements, as such agreements ensure that these countries increase their integration (considering the externalities they impose on each other) and improve the welfare of all members in the tier.

[30] This result does not necessarily rule out tiered structures. If there are barriers to integration such that higher-type members can prevent lower-type members from integrating into certain domains, then agreements can be used to increase the integration of lower-type countries. This would be the case if countries are allowed to integrate either at the union policy b^* or the tiered integration level b_τ, but are not free to choose any integration level in-between (b^*, b_τ). The optimal design of such tiers is an interesting question with arguably important implications, which we leave for future research.

Propositions 10 and 11 yield insights on EU's two major methods for deepening integration. The existence and effectiveness of tiers depend on the nature of the payoff function and the strategic complementarity. When the benefits of integration require reciprocal adoption of the same policies, a method like OMC, where the standards are set without enforcement can be enough to reap the benefits of integration. Conversely, ECA may be more beneficial in domains with considerable externalities due to their enforcement power.

5. Conclusion

We develop a model of international integration where countries with heterogeneous preferences decide to integrate, either joining a rigid or flexible international union or integrating without the commitment of the union. Inspired by the EU's focus on fostering a single market for goods, services, capital and labor, legislative and regulatory harmonization policies in capital markets, the ongoing banking union, and product market standardization, we model countries' actions as strategic complements, rather than contributions in a public good. In the first part of our analysis, we study the trade-offs of each integration method and examine union formation and enlargement. Unlike previous models focusing on public goods games, our strategic complementarities framework explains the evolution of the EU, where deeper integration in an increasing set of domains took place alongside enlargement to the periphery and adoption of flexible integration protocols, often spearheaded by the "core" countries. In the second part of the analysis, we extend our model to incorporate the union's integration with non-members to consider the pros and cons of a union's tendency to have special arrangements with candidates, other non-members, and exiting members. We demonstrate that placing restrictions on the (post-exit) integration level with outside (leaving) countries, even severe, is necessary to maintain the robustness of the union. Lastly, we explore endogenous and institutionalized multi-tier unions, shedding light on the ongoing debate on "two-speed" Europe and different methods EU uses for further integration such as OMC and ECA.

Our framework offers a baseline for thinking about the trade-offs of global integration with international unions, core institutions of the world economy. Our setting can be extended in further directions to study some first-order issues. First, integration can be modeled in a two-dimensional space, where countries have different preferences over different domains. One domain may be subject to strategic complementarities, reflecting trade and financial integration, while the second domain may represent investments in a public good, like defense. As many senior EU politicians are pushing towards joint European investments in security, studying a union with some domains under strategic substitutes and some under complementarities can shed light on one of the most profound policies of the following decades. Second, extending the model to include countries with different sizes may yield additional insights on enlargement and relations with non-members. Third, adding dynamics and endogenizing types can shed light on union stability. Fourth, allowing for union investments in citizens' and countries' preferences for integration may be useful in studying how unions may promote their own constituency and opponents. Fifth, one could consider alternative voting rules for enlargement or determination of integration protocols, such as a qualified majority, shedding light on a heated debate in EU policymaking. Sixth, allowing for within-country heterogeneity on preferences for integration, perhaps stemming from integration-promoting winners and losers, may bring insights into euro-skepticism. Finally, given rising geopolitical tensions, one could consider a multi-union world equilibrium framework where countries can decide to either join one union or integrate partially with multiple unions.

References

Abramson, B., and M. Shayo (2022): "International Integration and Social Identity," *The Journal of International Economics*, 137, 1035–77.

Alesina, A., I. Angeloni, and F. Etro (2005): "International Unions," *The American Economic Review*, 95(3), 602–615.

Alesina, A., I. Angeloni, and L. Schkuknecht (2005): "What Does the European Union Do?," *Public Choice*, 123(3/4), 275–319.

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

——— (2005): *The Size of Nations*. Mit Press.

Alesina, A., E. Spolaore, and R. Wacziarg (2000): "Economic Integration and Political Disintegration," *The American Economic Review*, 90(5), 1276–1296.

Alesina, A., G. Tabellini, and F. Trebbi (2017): "Is Europe an Optimal Political Area?," *Brookings Papers on Economic Activity*, pp. 169–213.

Baldwin, R. E., and P. Krugman (2004): "Agglomeration, integration and tax harmonisation," *European economic review*, 48(1), 1–23.

Baldwin, R. E., and P. Martin (2004): *Agglomeration and regional growth*vol. 4, pp. 2671–2711. Elsevier.

Bayoumi, T., and B. Eichengreen (1992): "Shocking aspects of European monetary unification," .

——— (1997): "Ever closer to heaven? An optimum-currency-area index for European countries," *European economic review*, 41(3-5), 761–770.

Berglof, E., M. Burkart, G. Friebel, and E. Paltseva (2008): "Widening and Deepening: Reforming the European Union," *The American Economic Review*, 98(2), 133–37.

Bolton, P., and G. Roland (1996): "Distributional Conflicts, Factor Mobility, and Political Integration," *The American Economic Review*, 86(2), 99–104.

——— (1997): "The Breakup of Nations: A Political Economy Analysis," *The Quarterly Journal of Economics*, 112(4), 1057–1090.

Bolton, P., G. Roland, and E. Spolaore (1996): "Economic Theories of the Break-Up and Integration of Nations," *European Economic Review*, 3(40), 697–705.

Buchanan, J. M. (1965): "An Economic Theory of Clubs," *Economica*, 32(125), 1–14.

Campos, N. F., and C. Macchiarelli (2021): "The dynamics of core and periphery in the European monetary union: A new approach," *Journal of International Money and Finance*, 112, 102325.

Casella, A. (2001): "The Role of Market Size in the Formation of Jurisdictions," *The Review of Economic Studies*, 68(1), 83–108.

——— (2005): "Redistribution Policy: A European Model," *The Journal of Public Economics*, 89(7), 1305–1331.

Celi, G., G. Andrea, G. Dario, S. Annamaria, et al. (2018): *Crisis in the European Monetary Union. A Core-Periphery Perspective*, vol. 45. Routledge.

Crémer, J., and T. R. Palfrey (1996): "In or Out?: Centralization by Majority Vote," *European Economic Review*, 40(1), 43–60.

Dehousse, R. (2003): "The Open Method of Coordination: A New Policy," *Les Cahiers européens de sciences Po*, 3.

Eichengreen, B. (2006): "The Parallel-Currency Approach to Asian Monetary Integration," *The American Economic Review*, 96(2), 432–436.

Fujita, M., P. R. Krugman, and A. Venables (2001): *The spatial economy: Cities, regions, and international trade*. MIT press.

Gancia, G., G. A. Ponzetto, and J. Ventura (2020): "A Theory of Economic Unions," *The Journal of Monetary Economics*, 109, 107–127.

Gans, J. S., and M. Smart (1996): "Majority Voting with Single-Crossing Preferences," *Journal of Public Economics*, 59(2), 219–237.

Gilbert, M. (2020): *European Integration: A Political History*. Rowman & Littlefield Publishers.

Guiso, L., H. Herrera, and M. Morelli (2016): "Cultural Differences and Institutional Integration," *The Journal of International Economics*, 99, S97–S113.

Guriev, S., and E. Papaioannou (2022): "The Political Economy of Populism," *The Journal of Economic Literature*, 60(3), 753–832.

Harstad, B. (2006): "Flexible Integration? Mandatory and Minimum Participation Rules," *The Scandinavian Journal of Economics*, 108(4), 683–702.

James, H. (2012): *Making the European Monetary Union*. Harvard University Press, Cambridge, MA, United States.

Kobielarz, M. L. (2024): "A Theory of International Unions with Exits," *The Journal of International Theory*, 215(1).

Krotz, U., and J. Schild (2013): *Shaping Europe: France, Germany, and Embedded Bilateralism from the Elysée Treaty to Twenty-First Century Politics*. Oxford University Press, Oxford, United Kingdom.

Krugman, P. (1991): "Increasing returns and economic geography," *Journal of political economy*, 99(3), 483–499.

Milgrom, P., and J. Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58(6), 1255–1277.

Oates, W. E. (1972): *Fiscal Federalism*. Edward Elgar Publishing.

Oates, W. E. (1999): "An essay on fiscal federalism," *The Journal of Economic Literature*, 37(3), 1120–1149.

Paterson, W. E. (2011): "The reluctant hegemon: Germany moves centre stage in the European Union," *J. Common Mkt. Stud.*, 49, 57.

Roberts, K. W. (1999): "Dynamic Voting in Clubs," *LSE STICERD Research Paper*.

Sapir, A. (2011): "European Integration at the Crossroads: A review Essay on the 50th Anniversary of Bela Balassa's Theory of Economic Integration," *The Journal of Economic Literature*, 49(4), 1200–1229.

Spolaore, E. (2013): "What is European Integration Really About? A Political Guide for Economists," *The Journal of Economic Perspectives*, 27(3), 125–44.

Tholoniat, L. (2010): "The Career of the Open Method of Coordination: Lessons from a 'Soft' EU Instrument," *West European Politics*, 33(1), 93–117.

Topkis, D. M. (1979): "Equilibrium Points in Nonzero-Sum n-person Submodular Games," *Siam Journal on Control and Optimization*, 17(6), 773–787.

——— (1998): *Supermodularity and Complementarity*. Princeton University Press.

Tsoukalis, L. (1997): *The New European Economy Revisited*. Oxford University Press, Oxford, United Kingdom.

Veiga, A., and A. Amaral (2006): "The Open Method of Coordination and the Implementation of the Bologna Process," *Tertiary Education & Management*, 12(4), 283–295.

Vives, X. (1990): "Nash Equilibrium with Strategic Complementarities," *The Journal of Mathematical Economics*, 19(3), 305–321.

Appendices

A. Proofs

A.1. Proof of Proposition 1

We first prove the first part. To simplify notation, let $t^{\tilde{r}}$ denote the integration level vector where all countries choose integration level $\tilde{r}$ while $(\hat{r}, t^{\tilde{r}}_{-i})$ denotes the vector where all countries other than i choose integration level $\tilde{r}$ and i chooses $\hat{r}$.

Lemma 1. *Let $r < r'$ be two integration levels. There is a cut-off country $n(r, r') \in U$ such that if $i < n(r, r')$ then i prefers r while if $i > n(r, r')$, then i prefers r'.*

Proof. Let $\gamma_i > \gamma_j$ and $r < r'$. We will show that if j prefers $t^{r'}$ to t^r, then so does i.[31] If

[31]Remember that if j is indifferent between r and r', j breaks the tie in favor of the lower policy, r.

j prefers $t^{r'}$ to t^r, we have $u(t^{r'}, \gamma_j) - u(t^r, \gamma_j) > 0$. Note that by increasing differences, we have

$$u(t^{r'}, \gamma_i) - u(r', t^r_{-i}, \gamma_i) \geq u(t^{r'}, \gamma_j) - u(r', t^r_{-j}, \gamma_j) \tag{8}$$

$$u(r', t^r_{-i}, \gamma_i) - u(t^r, \gamma_i) \geq u(r', t^r_{-j}, \gamma_j) - u(t^r, \gamma_j) \tag{9}$$

Summing these to equations, we obtain $u(t^{r'}, \gamma_i) - u(t^r, \gamma_i) \geq u(t^{r'}, \gamma_j) - u(t^r, \gamma_j) > 0$ and i prefers r' to r. Let $n(r, r')$ be the country with lowest type that prefers r' to r. Then all countries k with $\gamma_k > \gamma_{n(r,r')}$ also prefer r' to r. From the definition of $n(r, r')$, all countries with $\gamma_k < \gamma_{n(r,r')}$ prefer r to r', which proves the result. □

Let $m' = m + 1$, *i.e.,* the lowest-type country among countries that have higher types than the median. Let r^* denote the most preferred integration level of m and let r' denote the most preferred integration level of m'.

First, $\gamma_{m'} > \gamma_m$ and increasing differences imply that $r' \geq r^*$. Let $\tilde{r} \neq r^*$ be any integration level. Note that if $\tilde{r} < r^*$, then $n(\tilde{r}, r^*) \leq m$, thus more than half of the countries prefer r^* to $\tilde{r}$. If $\tilde{r} > r'$, then $n(r', \tilde{r}) \geq m'$, thus again more than half of the countries prefer r to $\tilde{r}$. If $\tilde{r} \in [r^*, r']$, then $n(r^*, \tilde{r}) \geq m'$. If $|U|$ is odd, then $m > |U| - m$ and more than half of the countries prefer r^* to $\tilde{r}$. If $|U|$ is even and $n(r^*, \tilde{r}) > m'$, then more than half of the countries prefer r^* to $\tilde{r}$. If $|U|$ is even and $n(r^*, \tilde{r}) = m'$, then both policies get same amount of votes. As we break the tie in favor of the lower policy, r^* is selected. Thus r^* is the Condorcet winner.

We now prove the second and third parts of the proposition. Fix $b \geq 0$. From part 2 of Assumption 3, for any country with type γ_i, there is a $t(\gamma_i)$ such that all integration levels above $t(\gamma_i)$ is strictly dominated for country i. Let $t_{max} = \max_i t(\gamma_i)$. Then we can restricting attention to $[b, t_{max}]^N$ as the strategy space eliminates only strictly dominated strategies and does not change the set of equilibria. Moreover, $[b, t_{max}]^N$ is compact, $u(t_i, t_{-i}, \gamma)$ is continuous in t_i and t_{-i} and u satisfies increasing differences in t_i and t_{-i}. Thus, this is a supermodular game, and first part of the result follows from Theorem 4.2.1 in Topkis (1998).

To prove the comparative statics, note that $u(t, \gamma)$ has increasing differences in γ_i, t_i and t_{-i}. Then second part of the proposition follows from Theorem 4.2.2 in Topkis (1998). Finally, we prove that $\overline{T}(b, \gamma)$ is the pareto dominant equilibrium. For simplicity, let $t = \overline{T}(b, \gamma)$ and $\hat{t}$ denote another equilibrium. For all i, we have $u(\hat{t}_i, \hat{t}_{-i}, \gamma_i) \leq u(\hat{t}_i, t_{-i}, \gamma_i) \leq u(t_i, t_{-i}, \gamma_i)$, where the first inequality follows from the fact that t is the highest equilibrium (i.e., $t_{-i} \geq \hat{t}_{-i}$) and second follows from the fact that t is a best response to t_{-i}.

This characterizes the of flexible union equilibria under a given b and proves Part 3 by setting $b = 0$. For notational simplicity, we suppress γ in $\overline{T}(b, \gamma)$. The following lemma shows that equilibrium integration levels are increasing in b.

Lemma 2. *$\overline{T}(b)$ is increasing in b.*

Proof. See Appendix C.1. □

The following lemma shows that that if the union policy is the most preferred policy of the median country, then the median and all lower type countries choose the policy as their integration level.

Lemma 3. *Suppose that $\gamma_i < \gamma_m$ and let b denote the most preferred policy of the median country. Then $\overline{T}_i(b) = \overline{T}_m(b) = b$.*

Proof. See Appendix C.2 □

Next, we prove following the lemma, which shows that the most preferred integration bound of the median country is the Condorcet winner in flexible union and proves the proposition:

Lemma 4. *The most preferred integration bound of the median country, b, is the Condorcet winner.*

Proof. Let $b' \neq b$. We will show that b wins against b' in majority voting. First, let $b' < b$ denote an alternative policy level that is lower than b. From Lemma 2, $\overline{T}(b) \geq \overline{T}(b')$. From the definition of b, median country votes for b and not for b'. Let j be another country such that $\gamma_j > \gamma_m$. There are two cases: either $\overline{T}_j(b') \geq b$ or $\overline{T}_j(b') < b$.

Case 1: If $\overline{T}_j(b') \geq b$, we will prove the result in two subcases, either $\overline{T}_j(b) \neq \overline{T}_j(b')$ or $\overline{T}_j(b) = \overline{T}_j(b')$.

Case 1.1: $\overline{T}_j(b) \neq \overline{T}_j(b')$. As $b > b'$, this implies $\overline{T}_j(b) > \overline{T}_j(b')$, we have

$$u(\overline{T}_j(b), \overline{T}_{-j}(b), \gamma_j) > u(\overline{T}_j(b'), \overline{T}_{-j}(b), \gamma_j) \geq u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_j) \tag{10}$$

where the first inequality follows from optimality of $\overline{T}_j(b)$ and strict concavity of u in its first argument, and second follows as $\overline{T}(b) \geq \overline{T}(b')$. This shows that j votes for b.

Case 1.2: $\overline{T}_j(b) = \overline{T}_j(b')$. First, observe that as the median country chooses b to b', we have $u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_m(b'), \overline{T}_{-m}(b'), \gamma_m) > 0$. In this case, we will consider two further subcases, $\overline{T}_m(b) > \overline{T}_m(b')$ or $\overline{T}_m(b) = \overline{T}_m(b')$. If $\overline{T}_m(b) > \overline{T}_m(b')$, then as $\overline{T}_j(b') = \overline{T}_j(b) \geq \overline{T}_m(b) > \overline{T}_m(b')$, the payoff of j strictly increases due to the increase of the integration level of m, which shows that j votes for b. If $\overline{T}_m(b) = \overline{T}_m(b')$, first note that, as b is the equilibrium integration policy, from Lemma 3, $b = \overline{T}_m(b) = \overline{T}_m(b')$. We have the following claim.

Claim 1. *There exists at least one member i such that $\overline{T}_i(b') < b$ and $\overline{T}_i(b') < \overline{T}_i(b)$*

Proof. First, observe that from $u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_m(b'), \overline{T}_{-m}(b'), \gamma_m) > 0$, as $\overline{T}_m(b) = \overline{T}_m(b')$, at least one other member must strictly increase its integration level, as otherwise this inequality cannot be strict, thus $\overline{T}(b) \geq \overline{T}(b')$ with $\overline{T}(b) \neq \overline{T}(b')$. For a contradiction suppose that all such members already had integration levels weakly higher than b. But then, $\overline{T}(b)$ would be an equilibrium under b', which contradicts that $\overline{T}(b')$ is the highest equilibrium under policy b. □

Given this claim, the payoff of j strictly increases under b compared to b' as $\overline{T}_j(b') = \overline{T}_j(b) > \overline{T}_i(b')$ for i given in the claim.

Case 2: If $\overline{T}_j(b') < b$, then we will consider two cases, first is $\overline{T}_j(b) > b$ and the second is $\overline{T}_j(b) = b$. In the first case, we have that

$$u(\overline{T}_j(b), \overline{T}_{-j}(b), \gamma_j) > u(\overline{T}_j(b'), \overline{T}_{-j}(b), \gamma_j) \geq u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_j) \tag{11}$$

where first inequality holds from optimality of $\overline{T}_j(b)$, strict concavity of u in its first argument and $\overline{T}_j(b') < b < \overline{T}_j(b)$ and second follows since $\overline{T}_{-j}(b) \geq \overline{T}_{-j}(b')$.

In the second case, note that $\overline{T}_j(b) = \overline{T}_m(b) = b$ and therefore $\overline{T}_{-j}(b) = \overline{T}_{-m}(b)$. As b is preferred to $b' < b$, we have $u(b, \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_m(b'), \overline{T}_{-m}(b'), \gamma_m) > 0$. From optimality of $\overline{T}_m(b')$, this implies $u(b, \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_j(b'), \overline{T}_{-m}(b'), \gamma_m) > 0$. As $\overline{T}_{-m}(b) \geq \overline{T}_{-j}(b')$, $u(b, \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_m) > 0$. As $b > \overline{T}_j(b')$ and $\overline{T}_{-m}(b) = \overline{T}_{-j}(b) \geq \overline{T}_{-j}(b')$, by increasing differences of u and that $\gamma_j > \gamma_m$, we have $u(b, \overline{T}_{-m}(b), \gamma_j) - u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_j) > 0$. As $\overline{T}_m(b) = \overline{T}_j(b)$, we have $\overline{T}_{-m}(b) = \overline{T}_{-j}(b)$, which implies

$$u(b, \overline{T}_{-j}(b), \gamma_j) - u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_j) > 0$$

proving that j also prefers b to b'.

Next, suppose that $b' > b$. Moreover, let $\hat{t}$ denote the smallest $t \geq b'$ such that $\overline{T}_m(\hat{t}) = \hat{t}$ (of course, it is possible that $\hat{t} = b'$). Such $\hat{t}$ exists as the series defined by $t_0 = b'$ and $t_k = \overline{T}_m(t_{k-1})$ is increasing and bounded. Moreover, as t_k is a best response to t_{k-1} and t_k is an increasing sequence, we have:

$$u(\overline{T}_m(t_k), \overline{T}_{-m}(t_k), \gamma_m) \geq u(\overline{T}_m(t_{k-1}), \overline{T}_{-m}(t_{k-1}), \gamma_m) \text{ for all } k \tag{12}$$

Thus $u(\overline{T}_m(\hat{t}), \overline{T}_{-m}(\hat{t}), \gamma_m) \geq (\overline{T}_m(b'), \overline{T}_{-m}(b'), \gamma_m)$. Let $\gamma_j < \gamma_m$. From optimality of b for the median country and $\hat{t} > b$, we know that $u(\overline{T}(b), \gamma_m) - u(\hat{t}, \overline{T}_{-m}(\hat{t}), \gamma_m) \geq 0$. Moreover, as $\overline{T}_m(\hat{t}) = \hat{t}$ and $\gamma_j < \gamma_m$, we have $\overline{T}_j(\hat{t}) = \hat{t}$. Thus $\overline{T}_{-m}(\hat{t}) = \overline{T}_{-j}(\hat{t})$ and we have $u(b, \overline{T}_{-j}(b), \gamma_m) - u(\hat{t}, \overline{T}_{-j}(\hat{t}), \gamma_m) \geq 0$. From optimality of $\hat{t}$ and that $\overline{T}_j(\hat{t}) = \overline{T}_m(\hat{t})$, we have, $u(b, \overline{T}_{-j}(b), \gamma_m) - u(b', \overline{T}_{-j}(\hat{t}), \gamma_m) \geq 0$. As $b' > b$ and $\overline{T}_{-j}(\hat{t}) \geq \overline{T}_{-j}(b)$, from $\gamma_j < \gamma_m$ and increasing differences we have $u(b, \overline{T}_{-j}(b), \gamma_j) - u(b', \overline{T}_{-j}(\hat{t}), \gamma_j) \geq 0$. As $\hat{t} \geq b'$, we have $\overline{T}_{-j}(\hat{t}) \geq \overline{T}_{-j}(b')$, thus, $u(b, \overline{T}_{-j}(b), \gamma_j) - u(b', \overline{T}_{-j}(b'), \gamma_j) \geq 0$. This inequality, together with $b' > b$ shows that j prefers b to b', which proves the result. □

A.2. Proof of Proposition 2

We first prove the following lemma. Let t^x denote the integration levels under a rigid union with policy x.

Lemma 5. *For any x, $u(\overline{T}(x), \gamma_i) \geq u(t^x, \gamma_i)$ for all γ_i, i.e. all countries prefer a flexible union with policy x to a rigid union with policy x.*

Proof. First, note that $\overline{T}_{-i}(x) \geq t^x_{-i}$. Then $u(\overline{T}(x), \gamma_i) \geq u(t^x_i, \overline{T}_{-i}(x), \gamma_i) \geq u(t^x, \gamma_i)$, where the first inequality follows from the optimality of $\overline{T}_i(x)$ and second from $\overline{T}_{-i}(x) \geq t^*_{-i}$. which proves the result. $\square$

Let r denote the rigid union equilibrium policy level and t^r denote the policy vector. From Lemma 5, $u(\overline{T}(r), \gamma_i) \geq u(t^r, \gamma_i)$ so median country prefers flexible union to rigid union.

Next, let b denote the equilibrium policy level under flexible union. There are three cases, $b = r$, $b > r$ and $b < r$. In the first case ($b = r$), all countries prefer flexible union to rigid by Lemma 5.

In the second case ($b > r$), we will show that countries that have higher type than the median prefers flexible union to rigid union. Let $\gamma_i > \gamma_m$. There are two sub-cases, either $\overline{T}_i(b) > b$ or $\overline{T}_i(b) = b$. In the first sub-case, note that $u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_i) > u(r, \overline{T}_{-i}(b), \gamma_i) \geq u(r, t^r_{-i}, \gamma_i)$ where first inequality follows from the optimality of $\overline{T}_i(b)$, $\overline{T}_i(b) \geq b > r$ and the strict concavity of u in its first argument. The second inequality follows from the fact that $\overline{T}_{-i}(b) \geq t^r_{-i}$ (which is implied by $b > r$). This proves the first sub-case. In the second sub-case, $u(b, \overline{T}_{-m}(b), \gamma_m) > u(r, \overline{T}_{-m}(b), \gamma_m) \geq u(r, t^r_{-m}, \gamma_m) \geq 0$ where first inequality follows from strict concavity of u and second from $\overline{T}_{-m}(b) \geq t^r_{-m}$. Note that $\overline{T}_i(b) = b$ implies that $\overline{T}_{-m}(b) = \overline{T}_{-i}(b)$. Therefore $u(b, \overline{T}_{-i}(b), \gamma_m) - u(r, t^r_{-i}, \gamma_m) > 0$. Since $b > r$ and $\overline{T}_{-i}(b) \geq t^r_{-i}$, by increasing differences we have $u(b, \overline{T}_{-i}(b), \gamma_i) - u(r, t^r_{-i}, \gamma_i) > 0$ which proves the result for the case $b > r$.

In the third case, ($r > b$), we will show that countries that have lower type than the median prefers flexible union to rigid union. Suppose that $\gamma_i < \gamma_m$. Since b is the equilibrium policy, we have

$$u(b, T_{-m}(b), \gamma_m) \geq u(r, T_{-m}(r), \gamma_m) \tag{13}$$

As $\gamma_i < \gamma_m$, from Lemma 3, $\overline{T}_i(b) = \overline{T}_m(b) = b$. Therefore, $\overline{T}_{-i}(b) = \overline{T}_{-m}(b)$. Then equation 13 and increasing differences imply $u(b, T_{-i}(b), \gamma_i) \geq u(r, T_{-i}(r), \gamma_i) \geq u(r, t^r, \gamma_i)$ which proves the result for the case $r > b$.

A.3. Proof of Proposition 3

Let r_S denote the integration vector where S countries integrate at r and other integrate at 0. We can express the rigid union payoff of a country with type γ in an S member union with union policy r as $u^R(r, \gamma, S) \equiv u(r_S, \gamma)$. As u is continuously differentiable, so does $u^R(r, \gamma, S)$. Strictly increasing differences then follows from minimal complementarity of integration, as whenever $S' = S \cup \hat{S}$ for some $\hat{S} \neq \emptyset$, $u_{t_i}(r_{S'}, \gamma) > u_{t_i}(r_S, \gamma)$ for all i.

Given I, γ_m and the integration level in the initial union r_I, we have that $\frac{\partial u^R(r,\gamma_m,|I|)}{\partial r}\Big|_{r=r_I} = 0$, as otherwise, the median country would either decrease or increase r_I. As u^R satisfies strictly increasing differences in integration level and union size, $\frac{\partial u^R(r,\gamma_m,|U|)}{\partial r}\Big|_{r=r_I} > 0$ Moreover, from the optimality of r_I, we have $u^R(r_I, \gamma_m, |I|) \geq u^R(r', \gamma_m, |I|), \quad \forall r' < r_I$. As u^R satisfies strictly increasing differences in integration level and union size and $|U| > |I|$, $u^R(r_I, \gamma_m, |U|) > u^R(r', \gamma_m, |U|), \quad \forall r' < r_I$. As u^R is continuous, $u^R(r_I, \gamma, |U|) > u^R(r', \gamma, |U|), \quad \forall r' < r_I$ whenever γ is close to γ_m. Thus there exists a non-empty set of types below γ_m that would choose r_I or a higher policy after enlargement. For each $\gamma < \gamma_m$, define

$$\underline{u}(\gamma) = \max_{r<r_I} u^R(r, \gamma, |U|) \quad \text{and} \quad \overline{u}(\gamma) = \max_{r\geq r_I} u^R(r, \gamma, |U|) \tag{14}$$

Claim 2. *Suppose that $\gamma < \gamma'$, $\overline{u}(\gamma) > \underline{u}(\gamma)$. Then $\overline{u}(\gamma') > \underline{u}(\gamma')$.*

Proof. Let $\underline{r}$ and $\overline{r}$ denote an element of maximizer set of equations 14 for γ. Let $\underline{r}'$ and $\overline{r}'$ denote an element of maximizer set of equations 14 for γ'. As $\overline{u}(\gamma) > \underline{u}(\gamma)$, we have $u^R(\overline{r}, \gamma, |U|) > u^R(\underline{r}, \gamma, |U|) \geq u^R(\underline{r}', \gamma, |U|)$. Then by increasing differences, $u^R(\overline{r}, \gamma', |U|) > u^R(\underline{r}', \gamma', |U|)$. From optimality of $\overline{r}'$, $u^R(\overline{r}', \gamma', |U|) \geq u^R(\overline{r}, \gamma', |U|) > u^R(\underline{r}', \gamma', |U|)$, which proves the result. □

This claim shows that there exists a cut-off type $\hat{\gamma} \geq 0$ such that for all $\gamma'_m < \hat{\gamma}$, the new union policy will be lower than r_I, while whenever $\gamma'_m > \hat{\gamma}$, the new union policy will be weakly higher than r_I.

Finally, whenever $r_U \geq r_I$, for all i such that $\gamma_i \geq \gamma_m$, we have the following

$$\begin{aligned} u^R(r_U, \gamma_i, |U|) - u^R(r_I, \gamma_i, |I|) &> u^R(r_U, \gamma_i, |U|) - u^R(r_I, \gamma_i, |U|) \\ &\geq u^R(r_U, \gamma'_m, |U|) - u^R(r_I, \gamma'_m, |U|) \geq 0 \end{aligned} \tag{15}$$

where first inequality holds $u^R(r_I, \gamma_i, |I|) < u^R(r_I, \gamma_i, |U|)$ to due higher integration of new members, second inequality holds by increasing differences in integration and type and third holds as r_U is chosen as the integration level under U. This shows that all above median initial union members are in favor of enlargement and finishes the proof.

A.4. Proof of Proposition 4

Let r_L and r_S denote the equilibrium integration levels in the larger union after enlargement and smaller union before enlargement respectively. Let m_L and m_S denote the median countries in these respective cases. We also introduce the following notation. For $k \in \{L, S\}$, the vector (r_k^m, r_l^I, r_m^C) denote the integration profile for countries, where superscripts indicate the median country (m), the initial members (I) and the candidates (C), and the subscripts denote the equilibrium policies in rigid union before and after enlargement. Suppose for a contradiction $r_L > r_S$. As r_S is the equilibrium integration level before enlargement, we have that $u_{m_S}(r_S^{m_s}, r_S^I, r_S^C, \gamma_{m_S}) > u_{m_S}(r_L^{m_s}, r_L^I, r_L^C, \gamma_{m_S})$. As utility is increasing in other countries integration, this implies that $u_{m_S}(r_S^{m_S}, r_S^I, r_L^C, \gamma_{m_S}) \geq u_{m_S}(r_L^{m_s}, r_L^I, r_L^C, \gamma_{m_S})$. By decreasing differences and $\gamma_{m_L} \leq \gamma_{m_S}$, we have $u_{m_L}(r_S^{m_L}, r_S^I, 0, \gamma_{m_S}) \geq u_{m_L}(r_L^{m_L}, r_L^I, 0, \gamma_{m_S})$. This shows that setting r_S increases the utility of the median country in the larger union, which is a contradiction.

A.5. Proof of Proposition 5

First, let r_I and r_U denote the equilibrium policies under initial and larger unions. Note that if $r_I = r_U$, then all countries strictly prefer the larger union as all non-members who become members strictly increase their integration levels. As u is continuous, there exists ϵ such that all countries strictly prefer whenever $|r_U - r_I| < \epsilon$.

A.6. Preliminary Results for Non-member Integration

In this section, we characterize the highest equilibrium under non-union integration and flexible union with non-member integration. We also prove some claims about the properties of equilibrium integration levels. The characterizations and claims are useful for proving the propositions.

We start by characterizing the non-union integration equilibrium. The following lemma shows that the integration level of a country is monotone in its type in any equilibrium.

Lemma 6. *Let t^* denote a non-union integration equilibrium and $\gamma_j > \gamma_k$. Then $t_j^* \geq t_k^*$.*

Proof. See Appendix C.3 □

Lemma 6 suggests an iterative procedure to characterize the equilibrium integration levels in the highest equilibrium. Define the following vector of integration levels iteratively:

$$t_i = \begin{cases} \tilde{t} & \text{where } (|U-1|)\tilde{u}'(\tilde{t},\gamma_1) = c'(\tilde{t},\gamma_1), \text{ for } i = 1 \\ \min\{\tilde{t}, t_{-i}\} & \text{where } (|U-i|)\tilde{u}'(\tilde{t},\gamma_i) = c'(\tilde{t},\gamma_i), \text{ for } i > 1 \end{cases} \tag{16}$$

This defines a unique vector as the equations $(|U-i|)\tilde{u}'(\tilde{t},\gamma_i) = c'(\tilde{t},\gamma_i)$ have a unique solution due to concavity of $\tilde{u}$ and strict concavity of c. The following Lemma shows that this construction yields the highest non-union integration equilibrium.

Lemma 7. *Vector t defined above is the highest non-union integration equilibrium.*

Proof. See Appendix C.4 □

We will now characterize the equilibrium policy and integration levels for a given union U with median country m and lowest type country j in the highest equilibrium. As in the non-union integration case, we will order countries in a particular way and compute a vector of integration levels such that each countries integration level is (i) the highest integration level they can have in any equilibrium and (ii) is a best-response when all countries higher in the order choose a weakly higher integration level. To this end, divide countries to five sets

- $\underline{N} = \{\underline{n}_1, \ldots, \underline{n}_{|\underline{N}|}\}$, non-members with types below j
- $\hat{N} = \{\hat{n}_1, \ldots, \hat{n}_{|\hat{N}|}\}$, non-members with types above j but below m
- $\overline{N} = \{\overline{n}_1, \ldots, \overline{n}_{|\overline{N}|}\}$, non-members with types above m
- $\underline{U} = \{\underline{u}_1, \ldots, \underline{u}_{|\underline{U}|}\}$, members with types below m
- $\overline{U} = \{\overline{u}_1, \ldots, \overline{u}_{|\overline{U}|}\}$, members with types above m (including m).

Within each set, countries are ordered according to their types, therefore $\overline{u}_1$ is the median country. We will denote the equilibrium integration levels by h and construct the highest equilibrium. We keep using t to denote the highest non-union integration equilibrium.

For countries in $\underline{N}$, set $h_{\underline{n}_i} = \min\{t_{\underline{n}_i}, \overline{o}\}$. From our construction of non-union integration equilibirum, for each country $\underline{n}_i$, $h_{\underline{n}_i}$ is (i) highest integration level they can choose in equilibrium and (ii) a best-response when all other countries choose weakly higher integration levels.

We will now characterize the equilibrium integration levels for $\hat{N}$, assuming all countries in $\overline{N}$, $\underline{U}$ and $\overline{U}$ choose weakly higher integration levels, which we will subsequently show is the case. For country $\hat{n}_i$,

$$h_{\hat{n}_i} = \begin{cases} \min\{\tilde{t}, t_{\underline{n}_{|\underline{N}|}}, \overline{o}\} & \text{where } (|U| + |\overline{N}| + |\hat{N}| - 1)\tilde{u}'(\tilde{t}, \gamma_{\hat{n}_1}) = c'(\tilde{t}, \gamma_{\hat{n}_1}), \text{ for } i = 1 \\ \min\{\tilde{t}, h_{\hat{n}_{i-1}}, \overline{o}\} & \text{where } (|U| + |\overline{N}| + |\hat{N}| - i)\tilde{u}'(\tilde{t}, \gamma_{\hat{n}_i}) = c'(\tilde{t}, \gamma_{\hat{n}_i}), \text{ for } i > 1 \end{cases} \tag{17}$$

By construction, for each $\hat{n}_i$, $h_{\hat{n}_i}$ is the highest integration level they can choose conditional on countries in $\underline{N}$ and lower type countries in $\hat{N}$ choosing weakly lower integration levels. Moreover, it is a best-response when all other countries choose weakly higher integration levels. To determine the equilibrium policy and the rest of the equilibrium integration levels, we will consider two cases.

Case 1: $(|U| - 1)\tilde{u}'(\overline{o}, \gamma_m) - c'(\overline{o}, \gamma_m) \geq 0$. In this case, even if all non-members' integration levels are restricted to be below $\overline{o}$, the median still sets and integration bound b^* above $\overline{o}$. Moreover, this b^* is given by the unique solution to the following equation:

$$b^* = t \text{ where } (|U| - 1)\tilde{u}'(t, \gamma_m) - c'(t, \gamma_m) = 0 \tag{18}$$

Equation 18 implies that, for all $t' < b^*$ and $i \in \overline{N}$,

$$(|U|)\tilde{u}'(t', \gamma_i) - c'(t', \gamma_i) > 0 \tag{19}$$

Equation 19 implies that for all $i \in \overline{N}$, integrating at $\overline{o}$ is a best-response to any integration vector where union members choose integration levels above $\overline{o}$, which is the case as $b^* \geq \overline{o}$. Moreover, as non-members, $i \in \overline{N}$ cannot integrate above $\overline{o}$. Thus, we set $h_{\overline{n}_i} = \overline{o}$ for countries in $\overline{N}$. Next, since all countries in $\underline{U}$ has lower types than m, for all $i \in \underline{U}$ and $t'' > b^*$, Equation 18 implies

$$(|U| - 1)\tilde{u}'(t'', \gamma_i) - c'(t'', \gamma_i) < 0 \tag{20}$$

As there are at most $(|U| - 1)$ other countries with integration levels above b^*, Equation 20 implies that any $i \in \underline{U}$ cannot integrate over b^* in any equilibrium. Thus we have $h_{\underline{u}} = b^*$ for countries in $\underline{U}$, which is their best response in $[b, \infty)$. Finally, the highest integration level any country $\overline{u}_i$ in $\overline{U}$ (where $1 < i < |\overline{U}|$) can have is given by the solution to the equation, which is defined inductively, starting with $\overline{u}_2$, as follows

$$h_{\overline{u}_i} = \min\{h_{\overline{u}_{i-1}}, \tilde{t}\} \text{ where } (|\overline{U}| - i)\tilde{u}'(t, \gamma_{u_i}) = c'(t, \gamma_m) \tag{21}$$

Moreover, $h_{\overline{u}_i}$ is best response whenever all higher type countries in $\overline{U}$ has weakly higher integration level. Therefore, we have iteratively constructed an increasing sequence of integration levels, where at each step, each country's action is the highest best-response given the integration levels already computed and conditional on remaining countries choosing higher integration levels. Thus, h is the highest equilibrium under Case 1.

Case 2: $(|U| - 1)\tilde{u}'(\overline{o}, \gamma_m) - c'(\overline{o}, \gamma_m) < 0$. In this case, the equilibrium policy is given by b^* that solves:

$$b^* = t \text{ where } (|U| + |\overline{N}| - 1)\tilde{u}'(t, \gamma_m) = c'(t, \gamma_m) \tag{22}$$

Given b^*, using the same arguments as Case 1, we can show that for all countries in $\underline{U}$, we have $h_{\underline{u}_i} = b^*$ as they have lower types than γ_m. Moreover, 22 implies that all countries in $\overline{N}$ and $\overline{U}$ choose weakly higher integration levels than b^*, since they have higher types than the median. We do not need to further characterize the integration levels of the countries in $U \cup \overline{N}$ to prove our results, but the existence of an equilibrium where all these countries choose integration levels higher than b^* follows from Topkis Theorem with action spaces $[b^*, \infty]$ for members and $[b^*, \overline{o}]$ for non-members, which finishes the (partial) characterization of equilibrium integration levels in Case 2.

We now show some useful claims that follow immediately from the construction of the highest equilibrium. Given a set of countries and non-member integration bound $\overline{o}$, we use t^N to denote the integration levels under non-union integration and t^U to denote the integration levels under the union U.

Claim 3. *Fix an equilibrium union U. If $i \in \underline{N}$, then $t_i^N \geq t_i^U$.*

Proof. Follows as we had $t_i^U = \min\{t_i^N, \overline{o}\}$. □

Claim 4. *Fix a union U where all non-members have lower types than the median. Let b' and $\hat{b}$ denote the equilibrium policies under non-member integration bounds o' and $\hat{o}$, where $o' < \hat{o}$. If $b' = \hat{b} > \hat{o}$ then (i) all non-member choose weakly higher integration levels and (ii) all members choose the same integration levels.*

Proof. As $b' = \hat{b} > \hat{o} > o'$ and all non-members have below-median types, their integration levels are determined the same way in both cases, with the only difference being each non-member are allowed to choose a higher integration level, which proves (i). As $b' = \hat{b}$, the integration levels of all members are also determined the same way, which proves (ii). □

Claim 5. *Consider a union U under two different non-member integration restrictions $\hat{o} > o'$. Let b' and $\hat{b}$ denote the equilibrium policies, and $\hat{t}$ and t' denote the integration levels. Then $\hat{b} \geq b'$ and $\hat{t}_i \geq t'_i$ for all i.*

Proof. First, note that if the equilibrium under $\hat{o}$ satisfies Case 1, then so does the equilibrium under o'. Then $b' = \hat{b}$. Let $BR_i^{o,b}(t_{-i})$ denote the best response of i to t_{-i} given the membership/non-membership of i and non-member integration restriction o and union policy b. As t' is an equilibrium, $BR_i^{o',b'}(t'_{-i}) = t'_i$. Then under $\hat{o}$, we have $BR_i^{\hat{o},\hat{b}}(t_{-i'}) \geq t'_i$ as all non-members can choose higher integration levels, which implies that $\hat{t}_i \geq t'_i$.

Second, if the equilibrium under $\hat{o}$ satisfies Case 2, then the equilibrium under o' can satisfy either Case 1 or Case 2. If it satisfies Case 1, then $b' < \hat{b}$. As t' is an equilibrium, $BR_i^{o',b'}(t'_{-i}) = t'_i$. Then under $\hat{o}$ and $\hat{b}$, we have $BR_i^{\hat{o},\hat{b}}(t_{-i'}) \geq t'_i$ as $\hat{o} > o'$ and $\hat{b} > b'$, which implies that $\hat{t}_i \geq t'_i$. If it satisfies Case 2, then $b' \leq \hat{b}$. As t' is an equilibrium, $BR_i^{o',b'}(t'_{-i}) = t'_i$. Then under $\hat{o}$ and $\hat{b}$, we have $BR_i^{\hat{o},\hat{b}}(t_{-i'}) \geq t'_i$ as $\hat{o} > o'$ and $\hat{b} \geq b'$, which implies that $\hat{t}_i \geq t'_i$. □

A.7. Proof of Proposition 6

Suppose that $b^* \leq \overline{o}$ in an equilibrium union U. Let t^U denote the integration level vector under U. We will show that either there exists a country that strictly increases its utility by leaving the union (thus, U is not an equilibrium) or the integration levels under U are weakly lower than non-union integration (thus, U is ineffective). Let j denote the lowest type member, and $\hat{U} = U \setminus \{j\}$. As $\gamma_j \leq \gamma_m$, $t_j^U = b^*$. Let H denote the set of members and non-members that has weakly higher integration level than b^* at t_i^U. Let $BR_i(t_{-i})$ denote the best-response of i given t_{-i}, which is unique by the strict concavity of u. We will consider two cases.

Case 1: $(|H|-1)\tilde{u}'(b^*,\gamma_j) - c'(b^*,\gamma_j) \geq 0$. Observe that this implies $BR_j(t^U_{-j}) = b^*$. As j is the lowest type member, for any $i \in U$, $BR_i(t_{-i}) \geq b^*$. In other words, the best-response of all members are weakly above b. Then U is still an equilibrium even if all members do not have to choose an integration level above b, but non-members are still restricted to choose integration levels below $\overline{o}$. Thus, for any country j, we have $BR_j(t^U_{-j}) \geq t^U_j$, which implies that $t^N_j \geq t^U_j$. Thus, the union is ineffective.

Case 2: $(|H|-1)\tilde{u}'(b^*,\gamma_j) - c'(b^*,\gamma_j) < 0$. In this case, integrating at $BR_j(t_{-j})$ gives j the highest payoff it can receive conditional on (i) all countries in $\underline{N}$ integrate as they do under t^U and (ii) all other countries choose weakly higher integration levels than $BR_j(t_{-j})$. Let u^* denote this payoff. From strict concavity of u, u^* is strictly higher than the payoff j receives as a union member. From Claim 3, all countries in $\underline{N}$ integrate as they do under t^U. Moreover, from construction of the equilibrium, j chooses $BR_j(t_{-j})$ and all higher type countries choose weakly higher integration levels than $BR_j(t_{-j})$. Thus, j is strictly better off by not joining U and U is not an equilibrium.

A.8. Proof of Proposition 7

First, take an arbitrary country $i \in C$. We first compare the utility of i as a member in U and as a non-member when the union is $\tilde{U} = U \setminus \{i\}$, under a given $\overline{o}$. Let b and $\tilde{b}$ denote the respective equilibrium policies under U and $\tilde{U}$, determined by the median countries m and $\tilde{m}$. From definitions of b and $\tilde{b}$ and the fact that initial members have higher types and are more numerous than the candidates, $\gamma_{\tilde{m}} \geq \gamma_m \geq \gamma_i$. If i becomes a member, then its utility is given by

$$u^i_{join} = (|U|-1)\tilde{u}(b,\gamma_i) - c(b,\gamma_i) \tag{23}$$

Note that this value is not affected by $\overline{o}$. If i decides not to join, then there are two changes. First, the union policy is $\tilde{b}$, and the integration of i is restricted at $\overline{o}$. We now compute the payoff of i as a non-member, as a function of $\overline{o}$. Let $\hat{t}_i$ denote the unique solution to $(|U|-1)\tilde{u}'(t_i,\gamma_i) - c'(t_i,\gamma_i) = 0$. Let $t^*_i = \min\{\hat{t}_i, \overline{o}\}$. As $\gamma_i < \gamma_{\tilde{m}}$, we have $(|U|-1)\tilde{u}'(\hat{t}_i,\gamma_{\tilde{m}}) - c'(\hat{t}_i,\gamma_{\tilde{m}}) > 0$. Thus, $\tilde{b} \geq \hat{t}_i \geq t^*_i$ and t^*_i is the integration level of i as a non-member under $\tilde{U}$. Therefore, the payoff of i as a non-member is given by $u^i_{non-member}(\overline{o}) = \max_{t_i \leq \overline{o}}(|U|-1)\tilde{u}(t_i,\gamma_i) - c(t_i,\gamma_i)$. As a payoff of 0 is attainable by choosing $t_i = 0$, if i does not join to $\tilde{U}$ when $\overline{o} = 0$, that is also the case for any alternative $\overline{o}$.

Now, suppose that $u^i_{join} \geq u^i_{non-member}(0)$. We will show that there exists $\overline{o}_i < b$ such that $u^i_{non-member}(o') > u^i_{join}$ whenever $o' > \overline{o}_i$. Observe that $(|U|-1)\tilde{u}'(b,\gamma_m) - c'(b,\gamma_m) = 0$. As $\gamma_i < \gamma_m$, $(|U|-1)\tilde{u}'(b,\gamma_i) - c'(b,\gamma_i) < 0$, which implies that $\hat{t}_i < b$. From strict concavity of u, the payoff of i is strictly higher if (1) i integrates at $\hat{t}_i$ and (2) all other countries choose weakly higher integration levels compared to what they do under $\tilde{U}$. As $\tilde{b} \geq \hat{t}_i$, (2) is satisfied. Whenever $o' \geq \hat{t}_i$, then (1) is also satisfied. Then there exists $o' < b$ such that $u^i_{non-member}(o') > u^i_{join}$. The existence of a cut-off $\overline{o}_i$ then follows from the fact that $u^i_{non-member}(o)$ is increasing in o and u^i_{join} does not depend on o. To find $\overline{o}_C$, take $\overline{o}_C = \min_{i \in C} \overline{o}_i$. As all $\overline{o}_i < b$, $\overline{o}_C < b$. This shows that enlargement happens only if $\overline{o} \leq \overline{o}_C < b$.

We now consider the incentives of initial members. Fix $i \in I$ and $\tilde{U}$ such that $I \subseteq \tilde{U} \subset U$. We will now show if i strictly prefers $\tilde{U}$ to U under some o', then i also strictly prefers $\tilde{U}$ to U under less restrictive non-member integration bounds.

Lemma 8. *If i strictly prefers $\tilde{U}$ to U under some $o' < \overline{o}_C < b$, then i strictly prefers $\tilde{U}$ to U under all $\hat{o} \in (o', \overline{o}_C]$.*

Proof. Let b' and $\hat{b}$ denote the equilibrium policies in union $\tilde{U}$ under o' and $\hat{o}$. Let $\hat{t}$ and t' denote the equilibrium integration levels under o' and $\hat{o}$.

From Claim 5, we have $\hat{b} \geq b'$ and $\hat{t} \geq t'$. We will prove the lemma in two cases.

Case 1: $\hat{b} = b'$. By Claim 4, i chooses the same integration level under both $\hat{o}$ and o', while all other countries choose a weakly higher integration levels under $\hat{o}$. Thus, i obtains a weakly higher utility under $\tilde{U}$ and $\hat{o}$ compared to $\tilde{U}$ and o'. As the utility from U does not depend on non-member integration, the result follows.

Case 2: $\hat{b} > b'$. We first prove the following claim.

Claim 6. *There exists $j \in C$ such that $j \notin \tilde{U}$ and j integrates at $\hat{b}$ under $\hat{o}$.*

Proof. First, note that as all non-members have lower types than the median, none would integrate over the union policy. Let $\tilde{m}$ denote the median under $\tilde{U}$. Suppose for a contradiction no non-member integrates at $\hat{b}$ under $\hat{o}$. Then it must be that $(|\tilde{U}|-1)\tilde{u}'(\hat{b}, \gamma_{\tilde{m}}) - c'(\hat{b}, \gamma_{\tilde{m}}) = 0$, which implies, as $b' < \hat{b}$, $(|\tilde{U}|-1)\tilde{u}'(b', \gamma_{\tilde{m}}) - c'(b', \gamma_{\tilde{m}}) < 0$, which contradicts that b' is the equilibrium policy under o' and $\tilde{U}$. Thus, there exists such a $j \in C$. □

Let j denote the non-member that chooses $\hat{b}$. From strict concavity of u, for all $t'' < \hat{b}$, we know that

$$u(\hat{b}, \hat{t}_{-j}, \gamma_j) > u(t'', \hat{t}_{-j}, \gamma_j) \tag{24}$$

Take any $i \in I$. We will consider two sub-cases for Case 2.

Case 2.1: $\hat{t}_i = \hat{b}$. In this case, as $\hat{t}_i = \hat{t}_j = \hat{b}$, we have $\hat{t}_{-i} = \hat{t}_{-j}$. From Equation 24, and $\gamma_i > \gamma_j$, for all $t'' < \hat{b}$ $u(\hat{b}, \hat{t}_{-i}, \gamma_i) > u(t'', \hat{t}_{-i}, \gamma_i)$. From optimality of $\hat{b}$ for i, for all $t''' > \hat{b}$, $u(\hat{b}, \hat{t}_{-i}, \gamma_i) \geq u(t''', \hat{t}_{-i}, \gamma_i)$. These two inequalities give the following $u(\hat{b}, \hat{t}_{-i}, \gamma_i) \geq$

$u(t'_i, \hat{t}_{-i}, \gamma_i)$. As $\hat{o} > o'$ and $\hat{b} > b'$, Claim 5 implies that $t' \leq \hat{t}$. Then we have $u(\hat{b}, \hat{t}_{-i}, \gamma_i) \geq u(t'_i, t'_{-i}, \gamma_i)$, which shows that i is weakly worse off under o' compared to $\hat{o}$. As the utility from U does not depend on non-member integration, if i strictly prefers $\tilde{U}$ to U under o', then that is also the case under $\hat{o}$.

Case 2.2: $\hat{t}_i > \hat{b}$. From concavity of u and strict concavity of c, $u(\hat{t}_i, \hat{t}_{-i}, \gamma_i) \geq u(t'_i, \hat{t}_{-i}, \gamma_i)$. As $t' \leq \hat{t}$, we have $u(\hat{t}_i, \hat{t}_{-i}, \gamma_i) \geq u(t'_i, t'_{-i}, \gamma_i)$, which shows that i is weakly worse off under o' compared to $\hat{o}$. As the utility from U does not depend on non-member integration, if i strictly prefers $\tilde{U}$ to U under o', then that is also the case under $\hat{o}$. □

Lemma 8 shows that if i prefers $\tilde{U}$ to U when $\overline{o} = 0$, then that would be the case for any higher non-member integration bound. This proves the first part of the proposition. Moreover, by Lemma 8, if i prefers U to $\tilde{U}$ when $\overline{o} = 0$, then either there exists $\overline{o}_i < \overline{o}_C$ such that the preference is reversed for all $o' > \overline{o}_i$ or i always prefers U to $\tilde{U}$. Letting $\overline{o}_I = \min_{i \in I} \overline{o}_i$, we find that enlargement happens only if $\overline{o} \leq \min\{\overline{o}_I, \overline{o}_C\} < b$. Moreover, we have also shown that whenever $\overline{o} < \min\{\overline{o}_I, \overline{o}_C\}$, then all countries weakly prefer U to $\tilde{U}$ for all $\tilde{U}$ with $\tilde{U} \subset U$ and $I \subseteq U$, thus enlargement is an equilibrium in that case.

A.9. Proof of Proposition 8

To prove part 1, observe that $\tilde{u}(t, 0) = 0$ and $\tilde{u}$ is continuous. Therefore, when γ_l is low enough, the country with the preference shock prefers no integration and exits under any $\overline{e}$. Moreover, as U is an equilibrium of the union formation game, if $\overline{e} = 0$, then all countries prefer to remain members. Then, if $\gamma_l = \gamma_1$, then even after a shock, the shocked country still prefers to be a member.Existence of a $\tilde{\gamma}$ follows from the fact that $\tilde{u}$ is continuous and increasing in γ. To prove part 2, we first show the following lemma.

Lemma 9. *If U is robust under $\overline{e}$, then U is robust under all $\overline{e}' < \overline{e}$.*

Proof. Assume i prefers to stay in the union after a shock under $\overline{e}$. Then the union is robust under $\overline{e}$. As $\overline{e}' < \overline{e}$, any payoff that an exiting country can attain under $\overline{e}'$ is attainable under $\overline{e}$ and staying in the union is preferred under $\overline{e}'$. □

For any $\gamma_l \geq \tilde{\gamma}$, take the maximum of all $\overline{e}$ such that the country that gets the preference shock stays in the union, $e(\gamma_l)$, which exists by Lemma 9 and continuity of $\tilde{u}$. From definition of $e(\gamma_l)$, if $\overline{e} > e(\gamma_l)$, then the country leaves the union after a shock. From Lemma 9, if $\overline{e} \leq e(\gamma_l)$, the the country stays as a member after getting a shock, proving the second part.

To prove part 3, take γ_l with $e(\gamma_l)$ and fix a $\gamma'_l > \gamma_l$. If $e(\gamma_l) = 0$, then we are done as $e(\gamma'_l) \geq 0$. If $e(\gamma_l) > 0$, then we have

$$((|U| - 1)\tilde{u}(b, \gamma_l) - c(b, \gamma_l)) - ((|U| - 1)\tilde{u}(e(\gamma_l), \gamma_l) - c(e(\gamma_l), \gamma_l)) = 0 \tag{25}$$

As $b > e(\gamma_l)$ and $\gamma_l' > \gamma_l$, from increasing differences,

$$((|U|-1)\tilde{u}(b,\gamma_l') - c(b,\gamma_l')) - ((|U|-1)\tilde{u}(e(\gamma_l),\gamma_l') - c(e(\gamma_l),\gamma_l')) > 0 \tag{26}$$

Thus, $e(\gamma_l')$ value that satisfies this equation as equality must be higher than $e(\gamma_l)$, proving the third part.

A.10. Proof of Proposition 9

Fix γ_i. Let t^e denote the most preferred integration level of a country after a preference shock, which is given by $(|U|-1)\tilde{u}'(t^e,\gamma_l) = c'(t^e,\gamma_l)$. Observe that as $\gamma_l < \gamma_m$, $t^e \leq b$. We can characterize the difference in payoff of country i with and without exit restriction as

$$E(\gamma_i,\gamma_l,\kappa) \equiv \frac{\rho}{|U|-1}\left(\tilde{u}(b,\gamma_i) - (\tilde{u}(t^e,\gamma_i) - \kappa)\right) + \rho\left((\tilde{u}(b,\gamma_l) - c(b,\gamma_l)) - (\tilde{u}(t^e,\gamma_l) - c(t^e,\gamma_l))\right)$$

where first term corresponds to the event that another country gets the preference shock and second term corresponds to the even that i gets the preference shock. As $b \geq t^e$, first term is increasing in γ_i, while second term does not depend on γ_i. Thus, if $E(\gamma_i,\gamma_l,\kappa) \geq 0$ and $\gamma_j > \gamma_i$, then $E(\gamma_j,\gamma_l,\kappa) > 0$, which proves the result.

A.11. Proof of Proposition 10

Fix γ and let t denote the equilibrium. As t is an equilibrium,

$$(n-i)\tilde{u}(t_i,\gamma_i) - c'(t_i,\gamma_i) \leq 0 \tag{27}$$

which implies that $(n-i-1)\tilde{u}(t_i,\gamma_i) - c'(t_i,\gamma_i) < 0$. As $\tilde{u}$ and c are twice continuously differentiable, there exists $\epsilon > 0$ such that if $\gamma_j \in (\gamma_i, \gamma_i+\epsilon)$, then $(n-i-1)\tilde{u}(t_i,\gamma_j) - c'(t_i,\gamma_j) < 0$, which implies that $t_i = t_j$ if $\gamma_j \in (\gamma_i,\gamma_i+\epsilon)$.

To show the case for $\gamma_j \in (\gamma_i-\epsilon,\gamma_i)$, we will consider two subcases. First, suppose that Equation 27 holds with equality. Then we have $(n-i+1)\tilde{u}(t_i,\gamma_i) - c'(t_i,\gamma_i) > 0$. Thus, there exists $\epsilon > 0$ such that if $\gamma_j \in (\gamma_i-\epsilon,\gamma_i)$, then $(n-i+1)\tilde{u}(t',\gamma_j) - c'(t',\gamma_j) > 0$ for all $t' < t_i$, which implies that $t_i = t_j$ if $\gamma_j \in (\gamma_i-\epsilon,\gamma_i)$. Second, suppose that Equation 27 holds with strict inequality. This is only possible if $t_{i-1} = t_i$, as otherwise decreasing t_i would be a profitable deviation for i. Setting $\epsilon = \frac{\gamma_i-\gamma_{i-1}}{2}$ yields the result.

To see why a tier agreement never increases welfare, fix a tier agreement τ, b_τ. Let i denote the lowest type country in τ. First, if $t_j = b_\tau$, then all countries integration levels are weakly lower under the tier agreement compared to the initial equilibrium, which implies that all are weakly worse off. Second, suppose that $t_j < b_\tau$. As i denote the lowest type country in τ, the number of countries that integrate above t_j is the same with and without the tier agreement. As any country with lower type than j have the same integration level with and without the tier agreement, t_j is the unique best response of j in both cases and j

obtains strictly lower utility under tier agreement.

A.12. Proof of Proposition 11

To prove the first part, suppose that for a contradiction $t_i = t_j$. Without loss of generality, let $\gamma_i > \gamma_j$. As t is an equilibrium, the following equations must hold, where u_1 denotes the partial derivative of u with respect to the first argument:

$$u_1(t_i, t_j, t_{-ij}, \gamma_i) = 0 \text{ and } u_1(t_j, t_i, t_{-ij}, \gamma_j) = 0 \tag{28}$$

which is a contradiction as $t_i = t_j$ and $\gamma_i > \gamma_j$ implies that $\tilde{u}_1(f(t_i, t_k), \gamma_i) \geq \tilde{u}_1(f(t_j, t_k), \gamma_j)$ for all k and $c_1(t_i, \gamma_i) < c_1(t_j, \gamma_j)$.

To prove the second part, fix i and $j = i + 1$ who are integrating above the union policy in equilibrium. Let t^{γ_j} denote the largest equilibrium integration levels when country j has type $\gamma_j \in [\gamma_i, \gamma_{i+2}]$. Define $h(x, t_{-ij}, \gamma_i) \equiv u_i(x, x, t_{-ij}, \gamma_i)$.

Suppose that $\gamma_j = \gamma_i = \hat{\gamma}$, which implies that $t_i^{\hat{\gamma}} = t_j^{\hat{\gamma}}$ and $t_{-i}^{\hat{\gamma}} = t_{-j}^{\hat{\gamma}}$. As f is continuously differentiable, so does u, and as $t^{\hat{\gamma}}$ is an equilibrium, we have that $\frac{\partial u(x, t_{-i}^{\hat{\gamma}})}{\partial x}|_{x=t_i^{\hat{\gamma}}} = 0$. As f is strictly increasing in all arguments, $\frac{\partial h(x, t_{-ij}^{\hat{\gamma}}, \gamma_i)}{\partial x}|_{x=t_i^{\hat{\gamma}}} > 0$. Thus, there exists $x^* > t_i^{\hat{\gamma}}$ and δ such that $u_i(x^*, x^*, t_{-ij}^{\hat{\gamma}}, \gamma_i) > u_i(t_i^{\hat{\gamma}}, t_j^{\hat{\gamma}}, t_{-ij}^{\hat{\gamma}}, \gamma_i) + \delta$.

We now show that this inequality still holds whenever γ_j is close enough to $\gamma_i = \hat{\gamma}$. Keeping the types of all countries but j the same, consider $\gamma_j = \gamma' \in (\gamma_i, \gamma_{i+2})$. As the utility functions are twice continuously differentiable, the best responses of all countries are continuous. Moreover, as $\gamma' > \hat{\gamma}$ implies that $t^{\gamma'} \geq t^{\hat{\gamma}}$, there exists $\epsilon > 0$ such that whenever $\gamma_j' - \gamma_i < \epsilon$, we have the following:

$$x^* > t_j^{\gamma'} \geq t_i^{\gamma'} \tag{29}$$

$$u_i(t^{\gamma'}, \gamma_i) \in (u_i(t^{\hat{\gamma}}, \gamma_i), u_i(t^{\hat{\gamma}}, \gamma_i) + \delta/4) \tag{30}$$

$$u_j(t^{\gamma'}, \gamma_j) \in (u_i(t^{\gamma'}, \gamma_i), u_i(t^{\gamma'}, \gamma_i) + \delta/4) \tag{31}$$

which implies that

$$u_i(x^*, x^*, t_{-ij}^{\gamma'}, \gamma_i) > u_i(t^{\gamma'}, \gamma_i) \tag{32}$$

$$u_j(x^*, x^*, t_{-ij}^{\gamma'}, \gamma_j) > u_j(t^{\gamma'}, \gamma_j) \tag{33}$$

showing that both i and j are better off after the tier agreement.

B. Discussion: Mapping the Model to the EU

In this section, we detail how one can view the EU through the lens of our model. We first argue why initial members of the EU correspond to high type countries in our setting. Next, we discuss how EU has admitted lower type countries in the periphery in each enlargement round and implemented flexible integration policies, while also increasing the integration in domains that all countries have participated simultaneously.

The higher-type countries in our framework correspond to the EU's founding members: Germany, France, Belgium, Luxembourg, the Netherlands, and Italy. Lower-type countries are the more recently admitted, mainly in the South and the East, such as Poland, Malta, Cyprus, Hungary, and Croatia. The distinction between the European "core" and "periphery" is prominent among policymakers, commentators, and researchers. Flagship EU policies, such as the structural and cohesion funds and the common agricultural policy, were put forward to attenuate core-periphery productivity gaps. The vast literature in political science and international relations on European integration takes more or less the core-periphery distinction as given (e.g., Tsoukalis (1997), James (2012)), while many studies analyze the hegemonic role of Germany and France, which often in collaboration with Belgium, Netherlands, and Luxembourg have taken the lead in designing and reforming EU policies and institutions, such as the euro, the banking union, and patenting policies (e.g., Krotz and Schild (2013), Paterson (2011), Celi, Andrea, Dario, Annamaria, et al. (2018)).

Likewise, theoretical works in economic geography and more applied research on European integration economics embrace the core-periphery distinction (e.g., Fujita, Krugman, and Venables (2001), Krugman (1991), Baldwin and Martin (2004), Baldwin and Krugman (2004)). Some empirical works use factor analysis techniques to identify core and periphery economies based on their similarities with the EU economy and the synchronization of their business cycle. In influential work, Bayoumi and Eichengreen (1992) and Bayoumi and Eichengreen (1997) first identified important dissimilarities between "core" (Germany, France, Belgium, Netherlands, and Denmark) and "periphery" (Greece, Ireland, Italy, Portugal, Spain, and the UK) countries at the onset of the European Monetary Unification project, studying the business cycle properties of EU12 countries. Campos and Macchiarelli (2021) employ cluster analysis to compile a continuum measure of how similarly EU member countries' economies respond in the short and medium run to aggregate supply shocks. Their estimates suggest a periphery (in decreasing order of similarity with the EU cycle) of Latvia, Ireland, Lithuania, Estonia, Luxembourg, Czech Republic, Greece, Portugal, Slovakia, Poland, Hungary, Finland, and Spain, with the core consisting of the UK, Sweden, Denmark, Germany, Austria, France, Netherlands, Slovenia, Belgium, and Italy.

Country	EU Agreements					
	Eurozone	Schengen	Divorce	Patent	Couples	EU Prosecutor
Italy	✓	✓	✓	✓	✓	✓
Germany	✓	✓	✓	✓	✓	✓
France	✓	✓	✓	✓	✓	✓
Austria	✓	✓	✓	✓	✓	✓
Belgium	✓	✓	✓	✓	✓	✓
Slovenia	✓	✓	✓	✓	✓	✓
Portugal	✓	✓	✓	✓	✓	✓
Malta	✓	✓	✓	✓	✓	✓
Luxembourg	✓	✓	✓	✓	✓	✓
Netherlands	✓	✓		✓	✓	✓
Spain	✓	✓	✓		✓	✓
Finland	✓	✓		✓	✓	✓
Greece	✓	✓	✓		✓	✓
Latvia	✓	✓	✓	✓		✓
Lithuania	✓	✓	✓	✓		✓
Estonia	✓	✓	✓	✓		✓
Croatia	✓	✓			✓	✓
Bulgaria			✓	✓	✓	✓
Slovakia	✓	✓				✓
Cyprus	✓				✓	✓
Sweden		✓		✓	✓	
Czechia		✓			✓	✓
Romania		✓	✓			✓
Denmark		✓		✓		
Hungary		✓	✓			
Ireland	✓					
Poland		✓				

Table 1: The participation of EU members to Eurozone, Schengen and Enhanced Cooperation Agreements on Applicable Divorce Law, Unitary Patent, Property Regimes of International Couples and European Public Prosecutor, as of January 2024.

Second, the EU has already taken steps corresponding to further integration in the shape of six Enhanced Cooperation Agreements, including the Schengen Area, the Eurozone, and the Unitary Patent Agreement. Table 1 documents countries' participation in each agreement. As opt-outs are possible, it is reasonable to conclude that participating countries have higher preferences for integration. Countries participating in all six agreements are Austria, Belgium, France, Germany, Italy, Luxemburg, Malta, Portugal, and Slovenia. Moreover, the countries lagging in integration and participating in the fewest agreements are those often viewed as peripheral members: Poland and Ireland participate only in one domain, while Hungary and Denmark participate in two. Although this is a rough measure, it aligns with our interpretation of the core and peripheral countries.

Since the initial enlargement round of 1973, when Denmark, the United Kingdom, and Ireland joined, the EU has admitted twenty-two new members (while losing the UK), primarily countries with lower preferences for integration (see discussion in Section B), simultaneously deepening and expanding integration in many areas, including the formation of a single market. This effect is most apparent in the enlargement rounds in 2004 and 2007, which included countries with considerable differences with the existing members, but accompanied with efforts toward greater integration such as the Lisbon Treaty, and more recently, European Green Deal, and Digital Services and Market Acts. Treaty of Lisbon increased integration considerably by strengthening the European Parliament and by establishing a long-term President of the European Council. The European Green Deal is a binding agreement that establishes restrictions on $CO2$ emissions and Digital Markets Act establishes common regulations for digital services, two areas of increased cooperation where policy making is historically reserved for sovereign governments. Another notable initiative in this regard is the Permanent Structured Cooperation (PESCO), established on December 11, 2017, which allows willing member states to deepen defense cooperation through joint projects, thereby enhancing Europe's security and defense capabilities while fostering closer integration among participating countries. Moreover, these integration efforts (in which all countries have participated) are also accompanied with deeper and flexible integration in many areas, such as the Schengen Agreement, European Monetary Union and Unitary Patent Agreement.

C. Omitted Proofs for Technical Lemmas

C.1. Proof of Lemma 2

Let $BR_i(t_{-i}, \gamma_i, b) = \arg\max_{t'_i \in [b, t_{max}]} u_i(t'_i, t_{-i}, \gamma)$, which is singleton as u_i is a strictly concave function of t_i. Let $BR(t, b)$ denote the profile of best response strategies obtained from $BR_i(t, \gamma_i, b)$ for all $i \in N$. Starting from $\bar{t}_0 = (t_{max}, ..., t_{max})$ and repeatedly applying best response correspondences, we obtain following sequence, $\overline{T}_k(b) \equiv BR^{k-1}(\bar{t}_{k-1}, b)$.

Note that $BR(t, b)$ is isotone in t and $\lim_{k\to\infty} BR^k \to \overline{T}^*(b)$. As t_{max} does not depend on b and $BR(t, b)$ is isotone, for $\tilde{b} \leq b$, we have, for all k, $\overline{T}_k(b) \geq \overline{T}_k(\tilde{b})$. Thus we have:

$$\overline{T}^*(t^b) = \lim_{k\to\infty} BR_i^k(b) \geq \lim_{k\to\infty} BR_i^k(\tilde{b}) = \overline{T}^*(\tilde{b}) \tag{34}$$

which yields the result.

C.2. Proof of Lemma 3

For a contradiction, suppose that $\overline{T}_i(b) > \overline{T}_m(b) \geq b$. Observe that this also implies $\overline{T}_{-m}(b) > \overline{T}_{-i}(b)$. As u is strictly concave, we have $u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_i) > u(\overline{T}_m(b), \overline{T}_{-i}(b), \gamma_i)$. Then, as $\overline{T}_i(b) > \overline{T}_m(b)$, by increasing differences, $u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_m) > u(\overline{T}_m(b), \overline{T}_{-i}(b), \gamma_m)$. As $\overline{T}_{-m}(b) > \overline{T}_{-i}(b)$ and $\overline{T}_i(b) > \overline{T}_m(b)$. And by increasing differences, $u(\overline{T}_i(b), \overline{T}_{-m}(b), \gamma_m) > u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m)$. But then as $\overline{T}_i(b) > b$, we have that $\overline{T}_{-m}(\overline{T}_i(b)) \geq \overline{T}_{-m}(b)$, which implies $u(\overline{T}_i(b), \overline{T}_{-m}(\overline{T}_i(b)), \gamma_m) > u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m)$. This contradicts that b is the optimal bound for the median country. Thus, $\overline{T}_i(b) = \overline{T}_m(b)$.

To prove that $\overline{T}_m(b) = b$, suppose that $\overline{T}_m(b) > b$. Again, from strict concavity, $u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m) > u(b, \overline{T}_{-m}(b), \gamma_m)$. As $\overline{T}_m(b) > b$, $\overline{T}_{-m}(\overline{T}_m(b)) \geq \overline{T}_{-m}(b)$. Thus $u(\overline{T}_m(b), \overline{T}_{-m}(\overline{T}_m(b)), \gamma_m) > u(b, \overline{T}_{-m}(b), \gamma_m)$. This contradicts that b is the favorite policy of the median country. Thus, $\overline{T}_i(b) = \overline{T}_m(b)$, which finishes the proof.

C.3. Proof of Lemma 6

Assume for a contradiction $t_j^* < t_k^*$. As u is strictly concave in first argument and t_k^* is chosen by country k, we have that $u(t_k^*, t_j^*, t_{-jk}^*, \gamma_k) - u(t_j^*, t_j^*, t_{-jk}^*, \gamma_k) > 0$. As $\gamma_k < \gamma_j$ and $t_k^* > t_j^*$, this implies that $u(t_k^*, t_j^*, t_{-jk}^*, \gamma_j) - u(t_j^*, t_j^*, t_{-jk}^*, \gamma_j) > 0$. Next, as $t_k^* > t_j^*$, by increasing differences, we have $u(t_k^*, t_k^*, t_{-jk}^*, \gamma_j) - u(t_j^*, t_k^*, t_{-jk}^*, \gamma_j) > 0$, which contradicts the choice of t_j^* by country j.

C.4. Proof of Lemma 7

By construction, t_1 is the highest integration level country 1 can choose in any equilibrium, and is a best-response when all other countries choose weakly higher integration levels. Let $i > 1$ be given. We will consider two cases, $t_i > t_{i-1}$ and $t_i = t_{i-1}$. If $t_i > t_{i-1}$, take any $t' < t_i$. As $t_i > t_{i-1}$, we have $(|U| - i)\tilde{u}'(t', \gamma_i) > c'(t', \gamma_i)$, which means that increasing integration level above t' is strictly beneficial to i as there are at least $(|U| - i)$ countries who integrate more than t'. Moreover, from concavity of $\tilde{u}$ and strict concavity of c, for any $t'' > t_i$ $(|U| - i)\tilde{u}'(t'', \gamma_i) < c'(t'', \gamma_i)$. As there are at most $(|U| - i)$ countries who integrate above t'', increasing integration level above t_i strictly decreases the utility of i. Thus, t_i is a best response to t_{-i}. If $t_i = t_{i-1}$, let j denote the lowest type country who chooses $t_i = t_j$. Then we have $(|U| - j)\tilde{u}'(t_i, \gamma_j) = c'(t_i, \gamma_j)$, which implies $(|U| - j)\tilde{u}'(t_i, \gamma_i) > c'(t_i, \gamma_i)$. Take any $t' < t_i$. As there are at least $(|U| - i)$ countries who integrate more than t', increasing integration level above t' is strictly beneficial to i. Next, as $t_i = t_{i-1}$, we have $(|U| - i)\tilde{u}'(t_i, \gamma_i) \leq c'(t_i, \gamma_i)$, which implies, for any $t'' > t_i$ we have $(|U| - i)\tilde{u}'(t'', \gamma_i) < c'(t'', \gamma_i)$. As there are at most $(|U| - i)$ countries who integrate above t'', increasing integration level above t_i strictly decreases the utility of i. Thus, t_i is a best response to t_{-i}.

To prove that t is the highest equilibrium, suppose that t' is another equilibrium where $t'_i > t_i$ for some i. Let j denote the lowest type country with $t'_j > t_j$. From definition of t, we know that $(|U| - j)\tilde{u}'(t_j, \gamma_j) \leq c'(t_j, \gamma_j)$, which implies

$$(|U| - j)\tilde{u}'(t'_j, \gamma_j) < c'(t'_j, \gamma_j) \tag{35}$$

As j is the lowest type that increases its integration under t', all lower types choose a lower integration level than t'_j. Then Equation 35 implies that j is strictly better off by decreasing its integration level, which proves that t' is not an equilibrium.

D. Additional Results and Extensions

D.1. Non-Union Integration

In this section, we compare non-union integration with the two main integration protocols we have studied.

Rigid Union vs Non-Union Integration. We first compare non-union integration to rigid union, starting with an example to illustrate the mechanisms at play.

Example 3. *There are five countries, $U = \{1, 2, 3, 4, 5\}$, with types $\gamma_1 = 2.5$ $\gamma_2 = 2.499$, $\gamma_3 = 1$ $\gamma_4 = 0.95$ $\gamma_5 = 0.9$. The utility function is:*

$$u(t_i, t_{-i}, \gamma_i) = \gamma_i \sum_{j \in U \setminus \{i\}} t_i t_j - \frac{t_i^3}{\gamma_i} \tag{36}$$

The following table gives integration levels and payoffs of countries under different protocols.[32]

	Integration Levels				Payoffs			
	{1,2}	3	4	5	{1,2}	3	4	5
Rigid Union	2.67	2.67	2.67	2.67	63.52	9.48	7.06	4.53
Non-union Integration	4.74	2.11	2.02	1.92	85.24	18.91	17.26	15.40
Flexible Union	5.49	3.01	3.01	3.01	133.10	23.97	19.98	15.82

Table 2: **Integration Levels and Payoffs under Different Integration Methods.**

In rigid union, the most preferred integration of the median, country 3, is implemented. In non-union, all countries freely choose their integration levels. Table 2, rows (1) and (2) shows the equilibrium actions and payoffs. While all countries choose the same level of integration under rigid union, non-union integration allows higher type countries to integrate more and low type countries to integrate less. This is better for all countries since higher type countries enjoy the higher integration from each other, not possible under rigid policies, while lower type countries are also better off integrating less.

The example, therefore, reveals that how the flexibility of non-union integration can be beneficial to all countries. In contrast, non-union integration lacks the commitment power of a (rigid) union, which can increase integration across the union. As the following proposition shows, when countries have similar types integration preferences, the commitment power of a rigid union dominates the flexibility of non-union integration. [See also Example 7 in the Appendix D.8.]

Proposition 12. *Suppose that u_i is continuously differentiable and strictly increasing in t_{-i}. Then for each γ_m, there exists an $\epsilon > 0$ such that rigid union is preferred to non-union integration by all countries whenever $\gamma_i \in (\gamma_m - \epsilon, \gamma_m + \epsilon)$ for all $i \in U$.*

Proof. Suppose that $\gamma_i = \gamma_m$ for all i, denote this type profile by $\tilde{\gamma}$. Let t^* denote the integration levels under the highest non-union integration equilibrium. Note that $t_i^* = t_j^*$ for all j. Since u_i is differentiable, we have that $\frac{\partial u_i(t_i, t_{-i}^*, \gamma)}{\partial t_i} = 0$. Define $\hat{u}(x, \gamma_m) = u(x, x, \ldots, x, \gamma_m)$.

[32] All values are rounded to 2 decimal places for all numerical examples.

Since u is strictly increasing in t_{-i}, $\frac{\partial \hat{u}(t^*,\gamma)}{\partial t^*} > 0$, which means that there exists ϵ such that when $r = t^* + \epsilon$, $u_i(r, r, \ldots, r, \gamma) > u_i(t^*, t^*, \ldots, t^*, \gamma)$. Since r^* is the most preferred integration level of all countries with homogeneous types, $u_i(r^*, r^*, \ldots, r^*, \gamma) \geq u_i(r, r, \ldots, r, \gamma)$ and the utility under rigid union equilibrium policy r^* is higher compared to non-union equilibrium. Let U_R denote the payoff of a country under rigid union equilibrium and U_N denote the payoff under non-union integration when types are given by $\tilde{\gamma}$. We have showed that $U_R > U_N$. Let $U_R^i(\gamma')$ denote the payoff of i under rigid union equilibrium when types are given by γ'.

Lemma 10. *For each $\delta > 0$, there exists an ϵ such that $|U_R - U_R^i(\gamma')| < \delta$ if $\gamma_i' \in (\gamma_m - \epsilon, \gamma_m + \epsilon)$ for all i and $\gamma_m' = \gamma_m$.*

Proof. Observe that the rigid union equilibrium policy under γ' is r^* since the median country has the same type. The result then follows from the continuity of u_i in γ. □

Let $U_N^i(\gamma')$ denote the non-union integration payoff of country i at γ'. Next, we prove the following lemma

Lemma 11. *For each δ, there exists ϵ such that $U_N^i(\gamma') < U_N + \delta$ whenever $\gamma_i \in [\gamma_m - \epsilon, \gamma_m + \epsilon]$*

Proof. We first prove the following claim.

Claim 7. *For each $\delta > 0$, there exists ϵ such that $|t^*(\gamma) - t^*| < \delta$ whenever $\gamma_i \in (\gamma_m - \epsilon, \gamma_m + \epsilon)$.*

Proof. If this result is not true, then there exists a sequence of type profiles γ^n such that $\gamma^n \to \tilde{\gamma}$ but $\lim_{n\to\infty} t^*(\gamma_n) \geq t^* + \epsilon_1$ for some $\epsilon_1 > 0$. This shows that the largest equilibrium under $\tilde{\gamma}$ is strictly smaller than $\lim_{n\to\infty} t^*(\gamma_n)$. But this is a contradiction to the upper-hemi continuity of the nash equilibrium (which is satisfied due to continuity of the utility function and compactness of the action space) correspondence. This proves the result. □

The lemma then follows from the above claim as the utility is continuous in t and γ. □

Observe that i prefers rigid union to non-union if

$$U_R^i(\gamma') - U_N^i(\gamma') = \left(U_R^i(\gamma') - U_R\right) + (U_R - U_N) + \left(U_N - U_N^i(\gamma')\right) > 0 \qquad (37)$$

Given $\delta = |U_R - U_N|/3$, from in lemmas 10 and 11 there exists an $\epsilon > 0$ such that the absolute value of the sum of first and third terms is smaller than the second term, which is positive, which proves the result. □

Flexible Union vs Non-Union Integration. Next, we turn to the comparison of flexible union with non-union integration. When $b = 0$, the two protocols coincide. However, when b is greater than the actions of the low-type countries, flexible integration increases the integration of these countries. As actions are complements, integration is higher under flexible union compared to non-union integration. Since the preferences of the median determine the equilibrium policy (Proposition 1), the median and all countries with higher types prefer the higher integration of the the flexible union. In Example 3, moving from non-union integration to flexible union, the median country chooses an integration bound that is higher than its non-union integration level. This brings lower type countries with it, while also causing an increase in the integration of higher type countries due to complementarities. The following proposition shows that this is a general result.

Proposition 13. *All countries choose a (weakly) higher integration level under flexible union as compared to non-union integration and a majority of countries prefers flexible union to non-union integration.*

Proof. First, note that flexible union with a trivial bound ($b = 0$) is same as non-union integration. Let $\bar{t}^*$ denote the largest non-union integration equilibrium.

Lemma 12. $\overline{T}(0, \gamma) = \bar{t}^*(\gamma)$

Proof. Follows immediately from the definition of equilibria in both cases. □

Let b denote the equilibrium policy for the flexible union. Next lemma shows that countries that choose higher levels integration levels compared to their integration level under non-union integration prefer flexible union to non-union integration.

Lemma 13. *Let $t_i^*(\gamma) \geq b$. Then i prefers flexible union with bound b to non-union integration.*

Proof. Since $\overline{T}(b, \gamma)$ is increasing in b, we have that $\overline{T}_{-i}(b, \gamma) \geq t_{-i}^*(\gamma)$. Therefore, $u(t_i^*, \overline{T}_{-i}(t_i^*, \gamma), \gamma_i) \geq u(t_i^*, t_{-i}^*(\gamma), \gamma_i)$. Since i can choose t_i^* in flexible union with $b \leq t_i^*$, the result follows. □

Corollary 3. *The median country prefers flexible union to non-union integration.*

Proof. First, observe that flexible union with bound equal to $b = t_m^*$ is preferable for the median to non-union integration by Lemma 13. As the flexible union bound is the most preferred bound for the median country (by Proposition 1), the result follows. □

Next, we show that all countries with types higher than the median also prefer flexible union. To simplify notation, we suppress γ in $\overline{T}(b, \gamma)$. Following lemma finishes the proof of the proposition.

Lemma 14. *If $\gamma_i > \gamma_m$, then i prefers flexible union to non-union integration.*

Proof. There are two cases, either $t_i^* \geq b$ or $t_i^* < b$. First case is immediate from Lemma 13. To prove the second case, assume $t_i^* < b$. Note that due to increasing differences in γ_i and t_i, $t_i^* \geq t_m^*$. Moreover, $t_i^* \geq t_m^*$ implies that $t_{-m}^* \geq t_{-i}^*$. There are two cases, $\overline{T}_i(b) > b$ and $\overline{T}_i(b) = b$. We first prove the first case, $\overline{T}_i(b) > b$:

$$u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_i) > u(t_i^*, \overline{T}_{-i}(b), \gamma_i) \geq u(t_i^*, t_{-i}^*, \gamma_i) \tag{38}$$

where first inequality follows from the strict concavity of u in its first argument and second from the fact that $\overline{T}_{-i}(b) \geq t_{-i}^*$. If $\overline{T}_i(b) = b$, then as $\gamma_i > \gamma_m$, $T_m(b) = b$. Then we have

$$u(b, \overline{T}_{-m}(b), \gamma_m) - u(t_m^*, t_{-m}^*, \gamma_m) \geq 0 \tag{39}$$

Then $\overline{T}_i(b) = b$ implies $\overline{T}_m(b) = b$. These two imply that $\overline{T}_{-m}(b) = \overline{T}_{-i}(b)$. Moreover, note that

$$u(t_m^*, t_{-m}^*, \gamma_m) \geq u(t_i^*, t_{-m}^*, \gamma_m) \geq u(t_i^*, t_{-i}^*, \gamma_m) \tag{40}$$

where first inequality follows from optimality of t_m^* and second from $t_{-m}^* \geq t_{-i}^*$. Combining these, we obtain:

$$u(b, \overline{T}_{-i}(b), \gamma_m) - u(t_i^*, t_{-i}^*, \gamma_m) \geq 0 \tag{41}$$

As $b > t_{-i}^*$ and $\overline{T}_{-i}(b) \geq t_i^*$, by increasing differences we get:

$$u(b, \overline{T}_{-i}(b), \gamma_i) - u(t_i^*, t_{-i}^*, \gamma_i) \geq 0 \tag{42}$$

which proves the result. □

□

D.2. Union Formation

In this section, we consider the initial formation of a union. Consider a finite set of countries, denoted by $N = \{1, 2, \ldots, |N|\}$. We study a union formation game where countries decide whether or not to form a union, vote over the union policy, and decide on their integration. $U \subseteq N$ denotes union members. We analyze the Subgame Perfect Equilibrium (SPE) of the following union formation game.

1. Countries decide to become a member or not.
2. Members decide the equilibrium policy b^* (r^*) of the flexible (rigid) union with majority voting.
3. Countries choose their actions.

- If $i \in U$ [member country], then i chooses an integration level $t_i \in [b, \infty)$ in flexible union and integrate at r^* at rigid union.
- If $i \notin U$ [non-member], then i chooses an integration level $t_i \in [0, \overline{o}]$.

As in many settings with strategic complementarities, the union formation game features multiple equilibria. We start with an example to illustrate the multiplicity of equilibria and the mechanisms at play.

Example 4. *There are six countries, $N = \{1, 2, 3, 4, 5, 6\}$, with the following preferences over integration: $\gamma_6 = 1.7$, $\gamma_5 = 1.6$, $\gamma_4 = 1.4$, $\gamma_3 = 1.39$, $\gamma_2 = 1.21$ and $\gamma_1 = 1.2$. The utility function is*

$$\hat{u}(t_i, t_{-i}, \gamma_i) = \sum_{j \neq i} (t_i t_j)^{\frac{\gamma_i}{2}} - t_i^2 \tag{43}$$

In the rigid union with k members and equilibrium policy r, the payoff of each country is

$$u^k(r, \gamma_i) = (k-1) r^{\gamma_i} - r^2 \tag{44}$$

Maximizing $u^k(r, \gamma_m)$ with respect to r, we find that the equilibrium policy with k members and median country type γ_m is given by

$$r^*(k, \gamma_m) = \left(\frac{(k-1)\gamma_m}{2} \right)^{\frac{1}{2-\gamma_m}} \tag{45}$$

Table 3 gives the equilibrium policies and payoffs for different rigid unions.

Union	Equilibrium Policy	u_5	u_4	u_3	u_2	u_1
$\{5, 6\}$	0.57	0.08	0	0	0	0
$\{4, 5, 6\}$	3.23	2.62	-0.12	0	0	0
$\{3, 4, 5, 6\}$	3.44	9.84	5.83	4.87	0	0
$\{2, 3, 4, 5, 6\}$	5.56	31.36	13.26	12.51	0.96	0
$\{1, 3, 4, 5, 6\}$	5.56	31.36	13.26	12.51	0	0.42
$\{1, 2, 3, 4, 5, 6\}$	7.70	71.80	27.82	26.06	-0.22	-1.42

Table 3: **Equilibrium Integration Levels and (relevant) Payoffs, Example 4**.

We now explore which of these unions are indeed equilibria. We start by considering a two member union of countries $\{5, 6\}$, row (1). The equilibrium policy is 0.57, giving both members positive payoff. To determine whether this is indeed an equilibrium, we check whether any of the remaining four countries prefers to deviate and join the union. In a three country union with 5 as the median country (row 2), the equilibrium policy jumps to 3.23, due to the complementary nature of integration. However, for country 4, the payoff from *joining is negative,* -0.12*; moreover, the payoff of lower type countries would be even lower. Therefore,* $U = \{5, 6\}$ *is an equilibrium union; there is no equilibrium with* 3 *members.*[33]

Next, we consider a four country union of $\{3, 4, 5, 6\}$ *(row 3). The median country is* 4*; equilibrium policy is* 3.44*, giving all members positive payoff. Compared to a three member union of* $\{4, 5, 6\}$*, the equilibrium policy increases even though the median has lower type, due to the complementarity of countries' actions. To examine whether this is an equilibrium or not, we consider the incentives of the two non-members, low-type countries* 1 *and* 2*. In a potential five country union, the median is still country* 4*, but the equilibrium policy increases to* 5.56*, as the larger union and complementarity nudge for deeper integration. Moreover, country* 1 *and* 2 *obtain a positive payoff if they enter the* 5 *country union. Therefore, the four member union* $\{3, 4, 5, 6\}$ *is not an equilibrium.*

Finally, we explore whether the 5 *country unions are equilibria, considering countries' payoffs from the six country union. The median is country* 3*; and equilibrium policy is* 7.70*. This increase in integration level results in a negative payoff for the low-type countries* 1 *and* 2*, who do not join a six country union. Therefore, the six country union is not an equilibrium, while the two five country unions,* $\{1, 3, 4, 5, 6\}$ *and* $\{2, 3, 4, 5, 6\}$ *are. Moreover, in* $\{1, 3, 4, 5, 6\}$*,* 2 *is not a member although* $\gamma_2 > \gamma_1$*. Thus, countries with lower integration types may opt out even when a country with lower preferences for integration enters, as they anticipate that the entry of a new member will endogenously raise the equilibrium integration policy.*[34]

Nonetheless, we can partially characterize the equilibria exploring some properties. First, due to strategic complementarities, if we compare two unions with the same median, the larger union will have a higher integration policy. Second, if a country with above-median type joins a union, there are two effects, working on the same direction of deeper integration. To start with, the preferences for integration of the median country (weakly) increases, pushing the equilibrium policy towards higher integration. Besides, the union becomes larger, which also increases equilibrium integration, as countries' payoffs from integration rise.[35] Therefore, in any equilibrium union, all countries with types above the median country are members. The following proposition formalizes this result.

[33] This simple example reveals an additional result of our model: country 4 would prefer to join the union, if the equilibrium policy remained the same after joining in. However, due to complementarities, its entry pushes the high-type countries 5 and 6 to set a higher integration and prevents country 4 from joining.

[34] This shows that the equilibrium union may not be contiguous, which is the case when the actions are strategic substitutes; see Proposition 1 of Alesina, Angeloni, and Etro (2005).

[35] This result suggests that EU's 12 members incentive to integrate further increased, when the EU expanded in 1995, admitting Austria, Finland, and Sweden, which for mostly geo-political reasons were not members. The subsequent policies, for example, to integrate capital and product markets, for example with the Financial Services Action Plan (FSAP) are in line with the model's prediction.

Proposition 14. *If U is an equilibrium union with median country m and $\gamma_i \geq \gamma_m$, then $i \in U$.*

Proof. Assume for a contradiction U is an equilibrium with median m, $\gamma_i > \gamma_m$ and $i \notin U$. Consider $\tilde{U} = U \bigcup \{i\}$ and let $\tilde{m}$ denote the median country at $\tilde{U}$. Since $\gamma_i > \gamma_m$, $\gamma_i \geq \gamma_{\tilde{m}}$.

We first consider the rigid union case. Let r denote the rigid union equilibrium policy under $\tilde{U}$. As the policy is chosen by the median, Part 2 of Assumption 3 guarantees that $u(t^r, \gamma_{\tilde{m}}) > 0$. Then $u(t^r, \gamma_i) > 0$ since $u(\cdot, \gamma)$ is increasing in γ. Thus, joining the union gives a strictly positive payoff to i and U is not an equilibrium.

Next, consider the flexible union case. Let b denote the flexible union equilibrium policy under $\tilde{U}$. By Lemma 3, $\overline{T}_{\tilde{m}}(b) = b$. We consider two cases, $\overline{T}_i(b) = b$ and $\overline{T}_i(b) > b$. If $\overline{T}_i(b) = b$ (which implies $\overline{T}_i(b) = \overline{T}_{\tilde{m}}(b)$ and $\overline{T}_{-i}(b) = \overline{T}_{-\tilde{m}}(b)$), then

$$u(b, \overline{T}_{-i}(b), \gamma_i) = u(b, \overline{T}_{-\tilde{m}}(b), \gamma_i) \geq u(b, \overline{T}_{-\tilde{m}}(b), \gamma_{\tilde{m}}) > 0 \tag{46}$$

where first equality holds by $\overline{T}_{-i}(b) = \overline{T}_{-\tilde{m}}(b)$, second inequality holds by increasing differences and $\gamma_i \geq \gamma_{\tilde{m}}$ and the final inequality holds from the fact that $\tilde{m}$ is the median country under $\tilde{U}$.

Finally, assume that $\overline{T}_i(b) > b$. Then from strict concavity of $u(t_i, \cdot)$

$$u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_i) > u(b, \overline{T}_{-i}(b), \gamma_i) \tag{47}$$

Moreover, $u(0, \overline{T}_i(b), \gamma_i) = 0$ and $0 < b < \overline{T}_i(b)$. Then the result follows from strict concavity of of $u(t_i, \cdot)$. □

Proposition 5 shows that formation of an initial union starts with countries with high preferences for integration and refines earlier results that show countries with similar preferences form the initial union.[36] Moreover, it is more in line with the initial formation of the EU, where the core countries with higher preferences of integration have initiated the process.

D.3. Enlargement to Periphery: Examples and Discussion

Example 5. *There are four initial members, $I = \{2, 3, 4, 5\}$ and one candidate, $C = \{1\}$, with types: $\gamma_5 = 1.7$ $\gamma_4 = 1.6$, $\gamma_3 = 1.5$, $\gamma_2 = 1.38$ and $\gamma_1 = 1.37$. The utility function is*

$$\hat{u}(t_i, t_{-i}, \gamma_i) = \sum_{j \neq i} (t_i t_j)^{\frac{\gamma_i}{2}} - t_i^2 \tag{48}$$

Under $\hat{u}$, the benefit i receives from integrating with j depends on the product on their integration levels $(t_i t_j)$, and i's type, γ_i. The total benefit of integration is obtained by summing

[36] In particular, Proposition 1 in Alesina, Angeloni, and Etro (2005).

across all countries, and subtracting the costs associated with integration, t_i^2.

Union	Equilibrium Policy	u_2	u_1
$\{2,3,4,5\}$	5.06	2.50	0
$\{1,2,3,4,5\}$	9.00	1.96	0.16

Table 4: **Equilibrium Integration Levels and (relevant) Payoffs, Example 5**

Without enlargement, the median country is 3*, and the equilibrium policy is* $r^* \approx 5.06$*, giving the lowest type member, country* 2*, a payoff of* 2.5*. If country* 1 *joins, the median country is unchanged. However, the equilibrium integration becomes* $r^* = 9$*. Even under this higher integration level, (low-type) country* 1 *prefers to join. However, due to the country* 2*'s low preference for integration, its payoff drops to* 1.96*. So, while the other countries' payoffs increase, enlargement will be blocked by country* 2*.*

Example 6. *There are five initial members,* $I = \{3,4,5,6,7\}$*, and two candidates,* $C = \{2,1\}$*. The countries' types are:* $\gamma_7 = 1.8$ $\gamma_6 = 1.6$ $\gamma_5 = 1.4$*,* $\gamma_4 = 1.24$ $\gamma_3 = 1.15$*, and* $\gamma_2 = 1.1$ $\gamma_1 = 1$*. The utility function is* $\hat{u}$ *from Example 5.*

Union	Equilibrium Policy	u_7	u_3	u_2	u_1
$\{3,4,5,6,7\}$	5.56	56.87	0.42	0	0
$\{2,3,4,5,6,7\}$	4.43	53.26	4.89	6.07	0
$\{1,2,3,4,5,6,7\}$	5.63	102.99	16.02	8.45	2.07

Table 5: **Equilibrium Integration Levels and (relevant) Payoffs, Example 6**

Without enlargement, the median country is 5*, and the rigid union equilibrium policy* $r^* \approx 5.56$*, giving country* 7*, with the highest preferences for integration, a payoff of* 56.87*. First, suppose that only country* 2 *applies. If country* 2 *joins, country* 4 *becomes the decisive median, which reduces the equilibrium policy to* $r^* \approx 4.43$*. Even though the union is larger, the high-type country* 7 *is worse off (with a payoff of* 53.27*) due to the lower integration policy, decided now by a lower-type median. Therefore, country* 7 *rejects the candidacy of* 2*. Second, suppose that countries* 1 *and* 2 *are considered jointly for admission. If both countries are admitted, the type of the median drops from* 1.4 *to* 1.24*. However, a larger union causes the equilibrium policy to increase from* 5.56 *to* 5.63*. Under the larger union, both the high-type country* 7 *and the low-type country* 3 *are better off, implying that countries* 4*,* 5*, and* 6 *are also better off. Consequently, the union will admit jointly countries* 1 *and* 2*. Thus, the* *seven-member union is an equilibrium. No enlargement is another equilibrium when neither* 2 *nor* 1 *applies, as even if one applies, their candidacy will be rejected.*

These examples illustrate two distinct mechanisms of enlargement. First, as more countries integrate, all prefer higher integration, pushing up the union's equilibrium policy. Second, the median country changes. If new members have lower preferences for integration (as historically and currently the case), the median's lower preference for integration may bring down the union's policy. As a result, the equilibrium policy for the enlarged union is ambiguous, depending on the preferences (and number) of initial members and the candidates. Example 6 shows that admitting at the same time more countries may balance the tradeoffs across countries with different preferences for integration, consistent with various EU enlargement rounds when the union admitted many countries.[37]

D.4. Additional Results on the Structure of Non-member Integration Equilibrium

In this section, we present some further results for the highest non-member integration equilibrium characterized in Appendix A.6. These results are useful in proving the results in Appendices D.5 and D.6.

Claim 8. *Fix U with lowest type member j and let $V = U \setminus \{j\}$. If $i \in \underline{N}$ at U, then i has same integration level at both U and V.*

Proof. Under both unions, i and all lower type countries chooses their integration level in the same order and in the same way. □

Claim 9. *Fix a union U with median m and union policy $b > \bar{o}$. Changing non-member integration bound from $\bar{o}$ to $o' \in (\bar{o}, b)$ does not affect the union policy. Moreover, all members and non-members who integrate less than $\bar{o}$ choose the same integration level in both cases.*

Proof. As $b > \bar{o}$,

$$(|U| - 1)\tilde{u}'(b, \gamma_m) = c'(b, \gamma_m) \tag{49}$$

[37]While it is always challenging to move from the abstract model to the complex reality of European politics, we believe this is in line with the EU's history, where enlargement went hand in hand with deeper integration, often spearheaded by core countries. Most enlargement rounds entailed many countries joining in at the same time. Spain and Portugal in 1986, Sweden, Finland, and Austria in 1996. During the Eastern Enlargement of 2004, the EU admitted eight Transition countries (the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, and Slovenia), Malta, and Cyprus; and three years later, also admitted Bulgaria and Romania. Arguably, these countries were dissimilar to the members at the time, corresponding to low integration types in our framework. Yet the considerable increase in members, coupled with strategic complementarity, made even high-integration members better off post-enlargement.

From concavity of $\tilde{u}$ and strict concavity of c, for all $o' \in (\overline{o}, b)$, $(|U|-1)\tilde{u}'(o', \gamma_m) > c'(o', \gamma_m)$. Thus, the construction of equilibrium is still at Case 1, which means that the equilibrium policy is still determined by Equation 49. The fact that all members and non-members who integrate less than $\overline{o}$ choose the same integration level is immediate from the construction of the equilibrium. □

Claim 10. *Fix a union U with equilibrium policy $b > \overline{o}$. Suppose that a non-member i joins and the union becomes $\hat{U} = U \cup \{i\}$ with union policy $\hat{b} \geq b$. Then all (member and non-member) countries choose a weakly higher integration level under $\hat{U}$.*

Proof. As $\hat{b} \geq b > \overline{o}$, the equilibrium under $\hat{U}$ is from the first case. Then the equilibrium integration level of all non-members is either (i) determined by the same equation (for non-members with lower types than i and members with higher types than i, if they are integrating above $\tilde{b}$) or (ii) determined by an equation that supposes one more country has a weakly higher integration level than them (for non-members with higher types than i and members with lower types than i). In both cases, each country chooses a weakly higher integration level, which proves the claim. □

Claim 11. *Suppose that U is an equilibrium union with $b > \overline{o}$ and integration level vector t^U. If $j \in U$, $\gamma_j < \gamma_m$, then $(|U|-1)\tilde{u}'(\overline{o}, \gamma_j) \geq c'(\overline{o}, \gamma_j)$.*

Proof. For a contradiction, suppose that

$$(|U|-1)\tilde{u}'(\overline{o}, \gamma_j) < c'(\overline{o}, \gamma_j) \tag{50}$$

Let $t_j^* \equiv BR(t^U, \gamma_j)$ denote the best-response of j to t^U. Equation 50 implies that $t_j^* < \overline{o}$.

From strict concavity of u and $t_j^* \leq \overline{o}$, if all other countries choose their integration levels under t^U and j chooses t_j^*, then j is strictly better off (let $\hat{u}$ denote the payoff of j in this case). Moreover, from the construction of equilibrium, if the union was $\hat{U} = U \setminus \{j\}$, (i) all non-members with lower types than j choose the integration level the choose under t^U and (ii) all other countries choose an integration level higher than t_j^*. Thus, in this case j obtains a payoff of $\hat{u}$, which is strictly larger than its payoff at U, which contradicts that U is an equilibrium. □

Claim 12. *Suppose that U is a union with equilibrium policy b and integration levels t^U, $i \notin U$ and $t_i^U = \overline{o}$. If the equilibrium policy under $\hat{U} = U \cup \{i\}$, $\hat{b}$, greater than $\overline{o}$, then the integration levels of all non-members under $\hat{U}$ is equal to their levels under t^U.*

Proof. As $\hat{b} \geq \overline{o}$, the equilibrium is computed from Case 1. Then in both cases, all non-members with lower types than i determine their integration levels the same way. Moreover, if $j \notin U$ and $\gamma_j > \gamma_i$, then $t_j^U = \overline{o}$ as $t_i^U = \overline{o}$. Thus, under $\hat{U}$, j also chooses $\overline{o}$ as $\hat{b} \geq \overline{o}$ and j takes into account higher integration level of i while determining its integration level. □

D.5. Efficient Non-member Integration Restrictions

Since the non-member integration restriction $\bar{o}$ is crucial for the union's effectiveness, we analyze its efficient determination. Given a non-member integration restriction $\bar{o}$ and equilibrium union U, the *membership incentive constraint binds* if the member with the lowest type is indifferent between joining and integrating as a non-member. Moreover, *non-member integration is restricted* if a non-member would prefer to increase its integration above $\bar{o}$, but cannot do so due to restrictions. The following proposition characterizes the efficient levels of non-member integration.

Proposition 15. *Let U be an equilibrium union where the membership incentive constraint does not bind. Then there exists $\bar{o}' > \bar{o}$ such that U is an equilibrium union under $\bar{o}'$. In this equilibrium, all members are weakly better off. If non-member integration is restricted at $\bar{o}$, then all members are strictly better off.*

Proof. Assume U is an equilibrium with $b > \bar{o}$ and median country m. We now extend Proposition 14 to non-member integration setting.

Lemma 15. *If $\gamma_i > \gamma_m$, then $i \in U$.*

Proof. For a contradiction, assume $i \notin U$. We will show that i can strictly increase its payoff by joining the union. As $b > \bar{o}$, we have $(|U| - 1)\tilde{u}'(b, \gamma_m) - c'(b, \gamma_m) = 0$. Suppose that i joins the union and let $\hat{U} = U \cup \{i\}$. Let $\tilde{m}$ denote the median under $\hat{U}$. Observe that as $\gamma_i > \gamma_m$, we have $\gamma_i \geq \gamma_{\tilde{m}} \geq \gamma_m$ and at least one of the inequalities is strict. Let $\tilde{b}$ denote the equilibrium policy under $\hat{U}$. As $|\hat{U}| > |U|$, $\gamma_{\tilde{m}} \geq \gamma_m$ and u satisfies increasing differences, we have $(|\hat{U}| - 1)\tilde{u}'(\tilde{b}, \gamma_{\tilde{m}}) - c'(\tilde{b}, \gamma_{\tilde{m}}) = 0$. where $\tilde{b} > b$ as $\hat{U}$ has weakly higher median and strictly more members. As $\gamma_i \geq \gamma_{\tilde{m}}$,

$$(|\hat{U}| - 1)\tilde{u}'(\tilde{b}, \gamma_i) - c'(\tilde{b}, \gamma_i) \geq 0 \tag{51}$$

We now show that i gets a strictly higher payoff under $\hat{U}$ compared to U. As $\tilde{b} > \bar{o}$, from strict concavity of u and Equation 51, lowering the integration level of i from $\tilde{b}$ to $\bar{o}$ (or any other lower value) while keeping all other countries integration levels constant makes i strictly worse off. As $\bar{o} < b \leq \tilde{b}$, from Claim 10, all other countries choose a weakly lower equilibrium integration level under U, which makes i weakly worse off. Therefore, i gets a strictly higher utility if it joins the union and U is not an equilibrium, proving the result. □

Observe that this lemma shows that if U is an equilibrium union, the set of non-members with types above the median country characterized in Section A.6, $\overline{N}$, is empty.

Lemma 16. *Suppose that $b > \bar{o}$ and membership incentive constraint does not bind. Then there exists $o' > \bar{o}$ such that U is still equilibrium.*

Proof. First, as $\tilde{u}$ is continuous, if $o' - \overline{o}$ is small enough (in particular, of $o' < b$), then by Claim 9 equilibrium policy under o' is same as the equilibrium policy under $\overline{o}$ and only non-members who integrate at $\overline{o}$ change their integration levels. As these countries cannot choose any level more than o', if $|\overline{o} - o'|$ is small enough, continuity of $\tilde{u}$ implies membership incentive constraints are satisfied under o'. Therefore, all members prefer to stay members under o'. Thus, to show that U is still an equilibrium, we need to show any $i \notin U$ does not want to join if non-member integration bound is o'.

Let $i \notin U$. From Lemma 15, $\gamma_i < \gamma_m$. Let $\hat{U} = U \cup \{i\}$ where $\hat{m}$ denotes the median at $\hat{U}$ and $\hat{b}$ denote the equilibrium policy under $\hat{U}$. We will now show that i obtains a weakly higher payoff under U compared to $\hat{U}$. We first prove the following claim.

Claim 13. *$\hat{b} > \overline{o}$.*

Proof. First, suppose that the median does not change, *i.e.,* $\hat{m} = m$, then $\hat{b} > b > \overline{o}$ as $\hat{b}$ solves $(|U|)\tilde{u}'(\hat{b}, \gamma_m) - c'(\hat{b}, \gamma_m) = 0$, while b solves $|U| - 1)\tilde{u}'(b, \gamma_m) - c'(b, \gamma_m) = 0$. Second, suppose that the median changes. As $\gamma_i < \gamma_m$, $\gamma_i \leq \gamma_{\tilde{m}}$. By Claim 11, we have $(|U| - 1)\tilde{u}'(\overline{o}, \gamma_{\tilde{m}}) \geq c'(\overline{o}, \gamma_{\tilde{m}})$, which implies $|U|\tilde{u}'(\overline{o}, \gamma_{\hat{m}}) > c'(\overline{o}, \gamma_{\hat{m}})$, which proves that $\hat{b} > \overline{o}$. ☐

Let t_i^U denote the integration level vector under U. We will prove the result in two cases.

Case 1: $t_i^U < \overline{o}$. Take $o' \in (\overline{o}, \min\{b, \hat{b}\})$ where $|\overline{o} - o'|$ is small enough that all members prefer to stay members under o'. Claim 9 implies that i and all countries with lower types choose the same integration levels, while all other choose integration levels weakly higher than i. Then by definition of t_i^U , i obtains the highest payoff it can get when all lower type countries choose their integration levels as t_i^U (which is equal to $t_i^{\hat{U}}$) and all higher type countries choose integration levels weakly higher than t_i^U. Moreover, this utility value is unique from strict concavity. Thus, if i joins the union and increases integration level to $\hat{b}$, i is strictly worse off.

Case 2: $t_i^U = \overline{o}$. we consider two sub-cases. First, suppose that $|U|\tilde{u}'(\overline{o}, \gamma_i) - c'(\overline{o}, \gamma_i) \leq 0$. In this case, if i joins, by Claim 12, all non-members choose the same integration levels, while the integration level of i increase from $\overline{o}$ to $\hat{b}$. Thus, from strict concavity of u, the payoff of i decreases strictly if i joins the union under $\overline{o}$. Consider $o' = \overline{o} + \epsilon$ with $\epsilon > 0$ for some small ϵ (in particular, small enough to make sure $o' < \min\{b, \hat{b}\}$). Claim 9 implies that, apart from countries who choose $\overline{o}$ as the integration level, change from $\overline{o}$ to o' has no effect on the integration levels of countries. As $\tilde{u}$ is continuous, there exists a small enough ϵ such that i strictly prefers integrating as a non-member under $o' = \overline{o} + \epsilon$. Next, suppose that

$$|U|\tilde{u}'(\overline{o}, \gamma_i) - c'(\overline{o}, \gamma_i) > 0 \tag{52}$$

Note that this implies that for small enough ϵ, under o', i chooses o' as its integration level. Moreover, from Claim 9 observe that the payoff difference between becoming a member and integrating as a non-member for i under $\overline{o}$ is given by

$$(|U|\tilde{u}(\overline{o},\gamma_i) - c(\overline{o},\gamma_i)) - \left(|U|\tilde{u}(\hat{b},\gamma_i) - c(\hat{b},\gamma_i)\right) \geq 0 \tag{53}$$

For small enough if $\epsilon = o' - \overline{o}$ is small enough, the same difference under o' is given by

$$(|U|\tilde{u}(o',\gamma_i) - c(o',\gamma_i)) - \left(|U|\tilde{u}(\hat{b},\gamma_i) - c(\hat{b},\gamma_i)\right) \tag{54}$$

Equation 52 implies that if ϵ is small enough, value given in equation 54 is larger than the one given in 53, and thus i prefers to integrate as a non-member under $\overline{o}$. This finishes the proof of the lemma. □

From Claim 9, all non-members weakly increase their integration under o' compared to $\overline{o}$, while the union policy and integration levels of members stay the same. Moreover, whenever there are non-members who prefer to integrate more than $\overline{o}$, they do so under o', which increases the payoffs of all members and finishes the proof of the result. □

Proposition 15 establishes that when considering union formation, the main determinant of the union's size is the incentives of the low-type potential members; it is without loss of optimality to restrict attention to bounds that make the lowest type non-member indifferent between joining or integrating outside the union. Moreover, whenever the membership incentive constraint doesn't bind, there is an inefficient restriction on the integration of non-members, and allowing non-members to integrate more is beneficial for all countries, regardless of membership status.

D.6. Enlargement Alongside Deeper Integration with Non-member Integration

We extend Proposition 3 to the setting with non-member integration described in Section 4.3.

Proposition 16. *Suppose that Assumption 5 holds and let γ_m and γ'_m denote the types of the median countries in I and $I \cup C$. There exists a $\hat{\gamma} < \gamma_m$ such that integration increases after admission of C to the union if and only if $\gamma'_m \geq \hat{\gamma}$. Moreover, initial members with above the median country types prefer enlargement whenever integration increases after enlargement.*

Proof. Let t denote the equilibrium integration level, and b denote the union policy before enlargement. First, note that as all candidates have lower types than the initial members, all such countries integrate below b, regardless of $\overline{o}$. Then b is given by the unique solution to the following equation: $b = t$ where $(|I|-1)\tilde{u}'(t,\gamma_m) - c'(t,\gamma_m) = 0$. This implies that $(|I \cup C|-1)\tilde{u}'(b,\gamma_m) - c'(b,\gamma_m) > 0$. As $\tilde{u}'(t,\gamma_m)$ is increasing in its second argument and $c'(t,\gamma_m)$ is decreasing in its second argument and both functions are continuous, there exists $\hat{\gamma} < \gamma_m$ such that $(|I \cup C|-1)\tilde{u}'(t,\gamma'_m) - c'(t,\gamma'_m) < 0$ if $\gamma'_m < \hat{\gamma}$ and $(|I \cup C|-1)\tilde{u}'(t,\gamma'_m) - c'(t,\gamma'_m) \geq 0$ otherwise, with equality at $\gamma'_m = \hat{\gamma}$, which proves the first part.

To prove the second part, let j denote initial member with above median type and $b' \geq b$ denote the equilibrium integration after enlargement. By Claim 10, all countries choose a weakly higher integration level after enlargement. First, if j was integrating above b' initially, then clearly j is better off after enlargement as it can still choose the same integration level and now all other countries are choosing weakly higher integration levels.

Second, suppose that j's integration level at the initial union was in $[b, b')$, denoted by t_j. From definition of b', we have that

$$(|I \cup C|-1)\tilde{u}(b',\gamma'_m) - c(b',\gamma'_m) \geq (|I \cup C|-1)\tilde{u}(b,\gamma'_m) - c(b,\gamma'_m) \tag{55}$$

As $\gamma_j > \gamma'_m$, we have that

$$(|I \cup C|-1)\tilde{u}(b',\gamma_j) - c(b',\gamma_j) \geq (|I \cup C|-1)\tilde{u}(t_j,\gamma_j) - c(t_j,\gamma_j) \tag{56}$$

Moreover, the second term is strictly larger than j's payoff before enlargement as non-members integrate below $t_j \geq b$, which completes the proof.

□

D.7. Initial Union Optimal Equilibria

Example 6 also demonstrates the possibility of equilibrium multiplicity stemming from strategic complementarity. A question regards the efficiency of the multiple equilibria. An equilibrium is an *initial union optimal equilibrium* if all initial members obtain (weakly) higher payoffs compared to any other equilibrium. The following proposition shows that under Assumption 4, there is a set of payoff-equivalent initial union optimal equilibria.

Proposition 17. *There is a set of initial union optimal equilibria that have the same number of members, same integration levels for all initial members and are the equilibria with most members. Each initial union member obtains the same payoff in all initial union optimal equilibria.*

Proof. To prove the first part, for a contradiction assume that $U' \subset U$, U' and U are equilibria and U gives a payoff that is strictly larger than its payoff under U' to some initial member i. Note that a profitable deviation for country i is rejecting all countries in $U \setminus U'$.

Under that deviation, the outcome would be U', which is a contradiction to our assumption that U is an equilibrium.

Next, from Assumption 4 and $|I| > |C|$, it is guaranteed that the median country will be an initial union member. If $|U| = |U'|$, number of new members are same both under U and U'. As all new members have lower types than the initial members, the median country in both unions is the same, which implies that the union policy and the integration profile and the payoff of all initial members are same. Then we have the following claim.

Claim 14. *If U and U' are both equilibria and $|U| > |U'|$, then all initial members prefer U to U'.*

Proof. Choose a U'' such that $U'' \subset U$ and $|U''| = |U'|$. We have already showed that if U is an equilibria, all initial members strictly prefer U to U''. As $|U''| = |U'|$, all initial members obtain the same payoff under U'' and U', which implies that they strictly prefer U to U'. □

Therefore, there exists a set of initial union optimal equilibria and all these equilibria has the same size, integration profile and initial member payoffs. □

Proposition 17 shows that the preferences of initial members over (equilibria of the) enlargement (game) are aligned. Therefore, it is reasonable to expect the initial members to agree on a union which is optimal, as the accession is an esoteric, complex, detailed and long process and the set of outcomes is a small set.

D.8. Omitted Examples

Example 7. *There are two countries 1 and 2 where $\gamma_1 = 1$ and $\gamma_2 = 1.5$*

$$u_i(t_i, t_j, \gamma_i) = \gamma_i t_i t_j - \frac{1}{3} t_i^3 \tag{57}$$

Non-union integration admits a unique equilibrium, where $t_1 \approx 1.14$ and $t_2 \approx 1.31$ with payoffs $u_1(t^*, \gamma_1) = 1$ and $u_2(t^*, \gamma_2) = 1.5$. The rigid union equilibrium policy is $r^* = 2$ and countries' payoffs are $u_1(2, 2, \gamma_1) \approx 1.33$ and $u_2(2, 2, \gamma_2) \approx 3.33$. In this example, the enforcement power of the union allows countries to increase their integration to higher levels, making both countries better off compared to non-union integration. This example illustrates that the enforcement power of unions plays an important role in fostering integration and moving from non-union integration (without commitment and enforcement) to rigid union may be Pareto-improving for all members.

Example 8. *There is an initial union of nine countries, $I = \{3, \ldots, 11\}$ and there are two candidates, $C = \{1, 2\}$. The types of countries are: $\gamma_{11} = 1.93$, $\gamma_i \in (1.8, 1.93)$ for*

$i \in \{8, 9, 10\}$, $\gamma_7 = 1.8$ $\gamma_6 = 1.6$ $\gamma_i = (1.5, 1.6)$ for $i \in \{3, 4, 5\}$, $\gamma_2 = 1.2$ and $\gamma_1 = 1$. All countries have the following the utility function:

$$u(t_i, t_{-i}, \gamma_i) = \sum_{j \neq i} (\min\{t_i, t_j\})^{\gamma_i} - t_i^3 \tag{58}$$

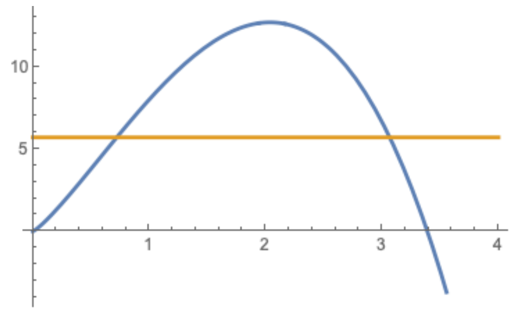

Fig. 4. **Example 8. Joining or Integrating as Non-Members. Country** 2. The horizontal axis denotes non-member integration level and the vertical axis denotes the integration payoff of members (Orange Line) and non-members integration payoff (Blue Line) for Country 2.

In the initial union, the median country is 7 and the equilibrium policy is $r^*(I) \approx 3.7$. We now consider when country 2 can join. If the country joins, country 6 becomes the decisive median; the new equilibrium policy falls, $r^*(I \cup \{2\}) \approx 3$. Although the union has added a member, the median country prefers a lower integration. We check the preferences of country 2 if it becomes a member or integrates without joining. On the one hand (net of integration with country 1, which is the same in both cases), joining gives country 2 a payoff of $9 \times 3^{\gamma_2} - 3^3 \approx 5.7$. On the other hand, integrating as an non-member at the cutoff level o gives a payoff of $9(\overline{o}^{\gamma_2}) - \overline{o}^3$. Figure 4 plots the payoffs (in the vertical axis) against the non-member integration level (horizontal line) for country 2 if it joins the union (orange line) and integrates at the union policy, $r^*(I \cup \{2\}) = 3$, and integrates as a non-member (blue line). When non-member integration is restricted below 0.7 (the point where the two curves intersect), country 2 prefers becoming a member since it results in a higher payoff, while if non-member integration restriction is weaker, then it prefers to integrate as a non-member. This echoes Example 2, illustrating how non-member integration is important for equilibrium union.

Next, examine the policy trade-off for country 1 plotting in Figure 5 its preferences joining the union and integrating as a non-member. Joining the union entails a negative payoff, due to the country's low type. Nonetheless, the country is willing to integrate as much as 1.7 as a non-member. However, given the 0.7 bound that the union imposes to make sure country 2 has incentive to join the union, country 1 cannot benefit fully from non-member integration.

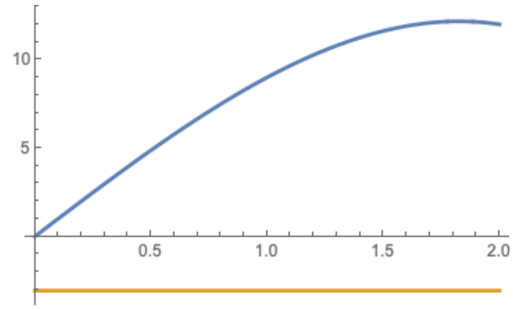

Fig. 5. **Example 8. Joining or Integrating as Non-Members. Country 1.** Horizontal axis denotes non-member integration level and vertical axis denotes membership payoff (Orange Line) and non-member integration payoff (Blue Line) for Country 1.

Moreover, the initial members also lose out from this restriction as they beneficial integration with country 1 *is prevented. However, the country is still better off as it can integrate partly with the union; and union members also benefit from country* 1 *partial integration as a non-member.*

Lastly, we explore the incentives of the high-type country 11 *on whether to allow country* 2 *to join. Country* 11 *compares a lower integration level but with a larger union against a higher integration level under a smaller union. In particular, the comparison entails the smaller union payoff (when* 2 *integrates at non-member integration bound* $\overline{o}$*)* $8\times(3.7)^{\gamma_{11}}+\overline{o}^{\gamma_{11}}-(3.7)^3$ *with the payoff if country* 2 *joins,* $9\times 3^{\gamma_1}-3^3\approx 49.5$*, plotted in Figure 6. Whenever the union allows country* 2 *to integrate more than cutoff* 0.4*, country* 11 *prefers country* 2 *to integrate as a non-member and therefore, rejects the candidacy of country* 2*. As a result, a non-member integration bound at* 0.4 *is necessary for unanimous acceptance of country* 2*'s candidacy. Note that at that bound, the incentive constraint of country* 2 *is slack. Country* 2 *would prefer to join the union even if non-members are allowed to integrate more than* 0.4*. Moreover,* 1 *wants to integrate more than* 0.4*. Thus, under enlargement, the efficiency of the non-member integration policies is not determined only by the incentives of the marginal members, but also depend on the incentives of the higher type countries who might block enlargement in favor of higher non-member integration.*

This example above shows that with enlargement, the incentives of the initial members are also important. Higher type countries may prefer to exclude lower type members joining in, so as to keep the equilibrium integration policy of the union high. Allowing for non-member integration, if anything, boosts this, as high-type countries have an even stronger incentive to reject candidates when non-member integration bound is high, since they can still benefit from their integration. Therefore, a low non-member integration bound might be important to convince such countries to allow enlargement. This is an additional channel that may cause inefficiency, as the integration of non-members, such as country 1 *in this example,*

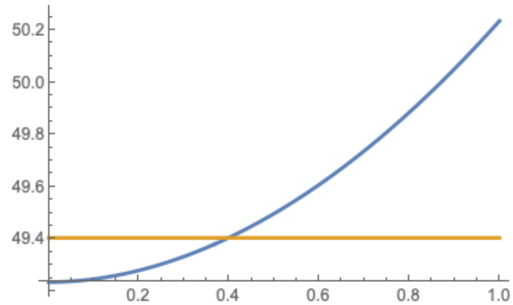

Fig. 6. **Example 8. Trade-off of High-Type Member Country** Horizontal axis denotes non-member integration level of country 2 and vertical axis denotes the payoff of country 11 when 2 becomes a member (Orange Line) and integrates as a non-member (Blue Line).

may be restricted.