International Unions and Integration

Oğuzhan Çelebi* Elias Papaioannou†

MIT LBS

December 8, 2022

Abstract

We consider a model of international unions in which countries have heterogeneous preferences for integration and their integration decisions are strategic complements. We analyze equilibrium under several integration protocols that differ in the flexibility countries have in choosing how much to integrate. Unlike previous models with strategic substitutes, our results are in line with the evolution of the European Union (EU), where enlargement and flexible integration coincide with enhanced integration and are often spearheaded by the "core" countries. Moreover, when non-members (candidates, exiting countries, and other nations) can partially integrate with the union, as in practice, restrictions on their integration determine the union's size and scope and are necessary for fostering cooperation. Motivated by Brexit and the rise of euroskepticism, we allow countries to leave the union and demonstrate how restrictions on the integration of leaving countries make the union robust to changes in members' preferences.

*MIT Department of Economics, 50 Memorial Drive, Cambridge, MA 02142. Email: ocelebi@mit.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, shaping national and international policies; their domains range from trade and banking to agriculture, industrial policy, product standardization, and climate change. While international organizations' objectives differ, a common goal is to foster cooperation and integration between members. The European Union (EU) is, arguably, the most prominent international organization, as its policy domains have expanded hugely 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*, common trade policy with non-members, a common agriculture policy, legislative and regulatory harmonization in financial services, product markets standardization, among others.[1] Importantly, European integration deepened, covering increasingly wider policy domain, as the EU welcomed new members: the United Kingdom (UK) in the mid-1970s, countries from Southern Europe in the 1980s, Scandinavia in the mid-1990s, and Eastern Europe in the 2000s. BREXIT and the increased tensions with Hungary illustrate that the Union may be fragile. However, the 2010-15 economic crisis in the European periphery, the pandemic, and the war in Ukraine have led many policymakers, commentators, and the public to argue for deepening and expanding integration in new domains, like health and security.[2] Besides, EU's incremental approach towards integration and membership expansion serves as a model for regional integration worldwide.[3]

This paper aims to understand how international unions facilitate integration, with a particular focus on the EU. We contribute to the earlier literature on international unions that study a "public good" setting where countries' actions are strategic substitutes in three main respects. First, our model stressing strategic complementarities is more tightly aligned with the main focus of international unions, which is fostering cooperation rather than coordinating infrastructure investments. Second, our framework can explain some stylized facts about the experience of the EU regarding countries' preferences over enlargement and

[1]See Spolaore (2013), Sapir (2011), and Eichengreen (2006) for extensive literature reviews on the functions and transformation of the EU, Alesina and Perotti (2004) for an evaluation of European integration and Gilbert (2020) for an historical overview.

[2]In an influential talk at the European Parliament on May 2, 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, welfare state policies, and labor issues) as well as increased attention to enlargement towards Western Balkans. French President Emmanuel Macron' 2022 Presidential Elections campaign focused on deeper EU ties on defense, calling for 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.

the flexibility of integration protocols. Third, extending the model by allowing non-members to integrate with the union, we gain insights into the relationship between the EU and non-members, candidates and exiting countries. Despite its simplicity, our setup sheds light, jointly, on some core issues of international unions: formation, enlargement, and relations and integration with non-member and exiting countries, which is an increasingly important topic after BREXIT and the rising Euroskepticism across Europe.

We develop a model of international integration where countries with heterogeneous preferences endogenously decide their *integration levels*, either forming a union with a common integration policy across members (*rigid union*), a union with a minimum threshold like the EU where further integration is possible (*flexible union*) or integrating independently without any explicit enforcement (*non-union integration*). Unlike the earlier literature, countries' actions represent integration and are strategic complements rather than strategic substitutes (*e.g.*, investments in a public good). Thus, as the union's size (number of countries) and depth (the extent of integration) increase, members have a stronger incentive to integrate more to reap up scale and market size effects. We believe this is more in line with the experience of the EU, its emphasis fostering a single market for goods, services, capital, and labor, standardization of regulations and safety protocols, financial sector legislative harmonization, and legal convergence in an increasingly larger domain. Besides, the EU budget is (still) small compared to the combined budget of the member countries (around 2% of the EU public spending) and classic public goods, like education, health, policing, and defense are provided mostly at the national level.

Not only our departure from a public investment setting is a more realistic representation of international unions, but with the introduction of strategic complementarities, our framework explains the experience of the EU where enlargement, deepening, and flexibility of integration have moved in tandem (*e.g.*, Eurozone and Schengen, which are not mandatory but allow for deeper integration). Therefore, our framework is in line with some salient features of European integration that earlier theoretical explorations could not easily explain. First, flexible integration and enlargement to the periphery are often spearheaded by the "core" countries such as Germany, France and the Netherlands. Second, enlargement to periphery has been accompanied by higher, not lower, integration across the union. We show that both of these facts are in line with the predictions of our model. With strategic complementarities, non-mandatory higher integration from the "core" countries and enlargement to the periphery can increase the depth of the union; as such, "core" countries with strong preferences for integration may support the union's expansion. Theoretical explorations where countries actions are strategic substitutes, such as investments to a common public good, yield opposing results: a reduction of integration across the union after enlargement to periphery, which creates opposition to union expansion by core countries, which do not square well with the EU evolution.

In addition, we introduce integration with non-members to the study of international unions, an issue that earlier works have abstracted from. Allowing non-member countries to partially integrate with the union is not only realistic, as the EU has many "enhanced" integration agreements with candidate countries and non-members, but also allows studying exit.[4] This allows understanding the trade-offs of the relations between the EU and candidate countries (e.g., Kosovo, Serbia, Albania), other non-member nations (e.g., Switzerland, Norway), and exiting countries, like the United Kingdom. Our framework reveals that *restrictions* on the integration of non-members and exiting countries are chief determinants of the size and the scope of the union. Besides, such restrictions on the integration of the union with non-members are necessary both for the union's effectiveness to foster integration and robustness against changes in countries' preferences.

1.1. Results Preview

Integration Methods/Protocols. We begin our analysis developing a model where countries' payoffs depend on their preferences for integration (types), their own and other countries' integration levels. We study three protocols: *rigid unions*, where all members integrate to the same level, set by majority voting; *flexible unions*, where, via majority voting, the union sets a lower bound for integration, but countries may integrate further; and *non-union integration*, where countries decide their actions without any constraints and enforcement mechanisms. Using tools from supermodular games and voting theory, we characterize the equilibrium policy, determined by the country with the median type, and equilibrium integration levels across the three protocols.

Comparison of Integration Methods. The theoretical framework uncovers the trade-offs between the three protocols. The distinction between a rigid union and non-union integration regards the enforcement power of the former and the flexibility of the latter. Depending on countries' preferences, each protocol may be strictly preferred to the other by every single country. Flexible union, however, balances the two countervailing forces; a majority of countries prefer it to the rigid union and non-union integration. The analysis of countries' preferences highlights the two main effects of a flexible union compared to rigid

[4] In our setting, non-members do not "invest" less in a common public good, bargaining how much of the benefits they will accrue. Much more realistically with the various accession protocols specified by the EU Treaty, the union and non-members choose to what extent (*e.g.*, in which domains) integration takes place (for example on trade, foreign investment, product standardization, etc).

union. First, while in rigid union the integration levels of all countries equal the union policy, under flexible union some countries integrate further, and therefore, fewer countries' integration directly depends on the union policy. This force pushes the median to pick a lower integration policy for the union. Second, due to complementarities, higher integration from countries with stronger preferences for integration leads *all* countries (including the median) to prefer higher integration. When the first effect dominates, flexible union policy is lower than rigid union policy and low-type countries prefer flexible arrangements. When the second effect dominates, flexibility results in a deeper integration, preferred by higher type members. The second possibility, absent from earlier studies (discussed below), appears to be in line with the evolution of the EU. Many countries of the "core", like Germany, France, Luxembourg, Belgium, and the Netherlands, have promoted flexible union, taking steps towards further integration in different domains (*e.g.,* monetary unification, banking union, integration of legal standards). At the same time, the deeper integration of the "core" has pushed more peripheral members, with ex-ante lower preferences for integration, to follow suit, integrating across more domains.

Formation and Enlargement. We then endogenize the composition of the union to study formation and enlargement. In every equilibrium of the formation game, all countries with types higher than the median become members. Turning to enlargement, we study a setting where (lower type) candidates apply for membership and (higher type) members vote over accession (as in practice of the EU). While enlargement results in a median country with a lower type (compared to the initial union), it is quite possible that the equilibrium policy *increases*, as integration is more beneficial when more countries integrate. While this pattern appears in line with the EU's history, where enlargement went handy with deeper integration, it does not emerge in public goods settings where countries' actions are substitutes. Therefore, our framework explains two chief regularities of the EU's history. First, countries with high -rather than low- integration preferences (such as Germany, France, and the Netherlands) have supported the EUs' various enlargement rounds. Second, post-enlargement, the EU has moved towards *more*, rather than less, integration.

Non-member Integration. We then extend our framework allowing non-member countries to integrate with the union to study the multiple-type agreements that international unions have with non-members, an issue that earlier research has abstracted from. Inspired by the EU, we consider a setting where non-members can integrate until an upper bound. We show that whenever the bound is not restrictive (*i.e.,* non-members are allowed to integrate as much as the union members), the union unravels, having no effect on fostering integration. Therefore, restrictions on non-members are chief determinants of the size and depth of the union. Besides, the efficiency of the upper bound restriction depends on the incentives of the low-type members, who would prefer to integrate as non-members without such restrictions.

Enlargement, Exit, and Non-member Integration We then revisit enlargement, allowing for the integration of potential members outside the union. A new channel through which non-member integration restrictions affect the union composition emerges when we also introduce enlargement. When considering candidates from the periphery, higher type countries compare the benefits of integrating with more countries to the possibility of a lower union policy induced by the lower median country after enlargement. If non-member integration is not very restrictive, these countries have a stronger incentive to reject candidates, nudging them to integrate as non-members, as they can still reap up the benefits of integration without giving voting power. Thus, integration must be restricted to satisfy incentive constraints of both the candidates and the initial members. Lastly, as an application of non-member integration and motivated by BREXIT, we study a setting where following a preference shock, members can exit the union while maintaining some degree of integration. We demonstrate that restrictions on the integration of former members with the union are necessary to make the union robust to preference shocks. Our theoretical result, thus, explains why exiting countries cannot "cherry-pick" the new integration with the union, as some proponents of BREXIT argued before the referendum. Besides, our set-up yields a natural exit protocol, in line with the Treaty of the European Union's article 50 and the heated negotiations between the EU and the UK after the June 2016 referendum. Moreover, countries preferring tighter exit restrictions are the ones with high preferences for integration, as these countries benefit and therefore are set to lose the most following an exit. This result is in line with the tough stance of core EU countries, like France, the Netherlands, and Germany on the negotiations with the UK, post the BREXIT referendum.

1.2. Related Literature

Our study relates to works examining the formation, functions, and enlargement of international unions, mostly focusing 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 et al. (1972)).[5] As Alesina, Angeloni, and Etro (2005), we

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

consider equally sized countries that solely differ on their preferences for integration. But, we modify, extend and reformulate their framework in multiple directions. First, we consider a setting where countries choose integration levels, which are strategic complements, rather than investments in a public good. This setting, we believe, is more realistic as many EU policies entail legislative-harmonization policies in product and capital markets, standardization, and common trade policies, subject to market size effects and scale; besides, the EU budget is small, and EU's role on standard public goods provision is small. Second, we allow non-member countries to integrate with the union, as it is the case.[6] Third, allowing jointly for non-member integration and enlargement provides insights on the arrangements and deals that the EU has with accession countries, like Turkey, and exiting from the union nations, like the UK.

Our paper also relates to works that explore theoretically certain aspects of the EU, and international unions, more generally. Harstad (2006) analyzes whether a union should allow members to form "inner clubs" to enhance cooperation. He shows that flexible cooperation is beneficial if country heterogeneity is large and externalities small. Berglof, Burkart, Friebel, and Paltseva (2008) also look at two-tier unions, showing in a framework where cooperation requires unanimity that the threat of an "inner club" by higher type members spurs contributions from lower type members, strengthening the union's cohesion. Although we take a different approach, our framework also yields that members prefer rigid union to more flexible, non-union integration when countries' preferences/types are similar [Proposition 2]. Kobielarz (2022) extends the Alesina, Angeloni, and Etro (2005)'s model to study exit allowing for transfers, adjustment of the union policy, and exit costs. Fiscal transfers can prevent inefficient exits; when transfers are unavailable, post-exit arrangements where the former member partially contributes to the public goods but receives limited spillovers may improve efficiency. Our results on exit are complementary, showing that similar economic mechanisms are present in a setting with different assumptions (strategic complements vs. substitutes) and interpretation (integration vs. investment in a public good).[7]

Our result on the necessity of restriction on non-member integration and characterization

treatment on the theory of the size of nations.

[6]Thinking about integration with non-members is chief, as the EU has various trade and investment treaties with many countries. In addition, there are five candidates (Albania, Turkey, Serbia, Montenegro, and North Macedonia), currently integrating with the EU on various domains, trade, regulation of financial markets, legal (Copenhagen Criteria). And a few years ago, most current EU members implemented a plethora of reforms, converging to the EU, when there were candidates.

[7]In both models, a worse post-exit relationship helps prevent exit and a better post-exit relationship is welfare-improving. However, in our model, post-exit relationship is governed through a restriction on the integration level former member can choose (*i.e.,* a restriction on actions) which determines to what extent the exiting nations benefits from integration. In Kobielarz (2022), former member can invest as much as it wants, but receives an exogenously-determined and smaller share of spillovers after exit.

of efficient restrictions [Propositions 9 and 10] echo similar mechanisms in Bolton and Roland (1996, 1997), who build a model with two countries deciding whether to separate or not. There are efficiency gains from unification, but also political conflict costs from differences in income and preferences towards redistribution. Bolton and Roland (1997) conclusion is that the efficiency loss from independence due to trade and factor movements barriers cement the Union.[8] Abramson and Shayo (2022) consider a two country model with core and periphery in which countries have different endogenous identities that affect their preferences over policies. They study the interplay between international integration and identity politics and how the robustness of the union depends on the social identities of member countries.

Our paper builds on the large literature on the theory of clubs, initiated by Buchanan (1965). The closest is Roberts (1999), who develops a dynamic voting problem with an endogenous electorate under strategic complementarities (see also Roberts (2015)).[9] We take a less abstract viewpoint, as our objective is to characterize various forms of international integration, understand the trade-offs of their formation, expansion, and even shrinkage, allowing realistically for partial integration of the union with non-members.

Outline. Section 2 sets up our theoretical framework and characterizes equilibria under rigid union, flexible union, and integration without a union. Section 3 compares the equilibrium policies and integration levels under the different methods, exploring which policies are preferred by what types of countries. Section 4 endogenizes the union to analyze its formation and enlargement. Section 5 extends the model allowing for non-member integration. We then study enlargement and exit under this richer setup. In Section 6 we summarize and discuss avenues for future research.

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

[9] Even though the payoff of members depends only on the club size and not actions, Roberts (1999) studies a setting similar to our rigid union. Our focus on international unions alongside with our core assumption of strategic complementarities can be viewed as a bridge between Alesina, Angeloni, and Etro (2005) and Roberts (1999), with the important addition of allowing for integration of non-members with the union and enlargement.

2. Theoretical Framework

This section develops the theoretical model that allows studying flexible unions, rigid unions, and non-union integration. We commence with the model set-up (Section 2.1). Then we define each integration method (Section 2.2), and characterize the equilibria (Section 2.3).

2.1. Model Set-Up

$U = \{1, 2, ..., |U|\}$ 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] The action of country i is denoted by $t_i \in \mathbb{R}_+$; $t \in \mathbb{R}_+^{|U|}$ denotes the action profile. Utility of country i is given by $u_i(t, \gamma_i) \equiv u(t_i, t_{-i}, \gamma_i)$, where $u : \mathbb{R}_+^{|U|+1} \to \mathbb{R}$ and is increasing in the country's preference for integration, γ_i. The actions can be thought as investment/spending decisions on a public good, as in earlier research or perhaps more realistically as efforts towards common policies in goods and services trade, movement of labor across borders, product market standardization, homogenizing laws, regulations, and acts on banking and capital markets. We now present our main assumption on the payoffs.

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

(i) u satisfies increasing differences in γ_i, t_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 if $t_j < t_i$ for all j.

The first part ensures that integration decisions are strategic complements: the payoff of increasing integration for a country is increasing in its type and the integration levels of other countries. This is the key difference between our model and public goods models where countries' actions are interpreted as their investments to a public good (e.g., Alesina, Angeloni, and Etro (2005)). The second part makes sure that integration is beneficial: each country prefers other countries to integrate more. Next assumption makes sure the model has a non-trivial equilibrium in which countries have positive integration levels.

Assumption 2. *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 $t_i > \hat{t}$ and a $\tilde{t} > 0$ such that $u(t, \gamma_i) > 0$ if $t_i = \tilde{t}$ for all $i \in U$.

(iii) u is strictly concave in t_i.

[10] We do not take a stance of the origins of integration preferences. They could, for example, reflect the desire of elites 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 citizen's ideology, values, and beliefs.

The first part of the assumption normalizes the payoff of no integration to zero and makes sure that cooperation is necessary for a country to reap the benefits of integration; for example, countries need to amend domestic legislation to harmonize product standards, remove tariffs and trade barriers that protect local industries, change labor legislation promoting free-movement, and pass new laws on union domains. The second part ensures that the benefits of integration are initially higher than the costs, but are eventually dominated by them, guaranteeing the existence of a non-trivial equilibrium.[11] The third part ensures that each country has a unique best-response given integration of other countries.

2.2. Integration Protocols

We distinguish and analyze three methods/protocols of integration *(i) non-union integration*; *(ii) rigid union*; and *(iii) flexible union*. Below we define these protocols.

Non-union integration is the simplest and most flexible form of cooperation; countries independently and simultaneously choose their integration levels. There is no explicit bargaining, negotiation, or centralized enforcement. An action profile t^* is a *non-union integration equilibrium* if for all $i \in U$

$$t_i^* \in \arg\max_{t_i} u_i(t_i, t_{-i}^*, \gamma_i) \tag{1}$$

In a *rigid union*, couturiers decide on the common integration level by majority voting.[12] The *rigid union equilibrium policy* is an integration level $r^* \in \mathbb{R}_+$ that is preferred to any r' by a majority of countries, in other words, r^* satisfies the *Condorcet Criterion*.[13]

In a *flexible union*, which resembles the EU the most, there is a minimum threshold of integration for all members, but countries can take steps for deeper integration. The *flexible union equilibrium policy* is a lower bound b, determined by majority voting. Given b, countries may voluntarily choose any action in $[b, \infty)$. A vector of integration levels $T(b)$ is a flexible union equilibrium under threshold b if

$$T_i(b) \in \arg\max_{t_i \geq b} u_i(t_i, T_{-i}(b), \gamma_i) \quad \forall i. \tag{2}$$

There might be multiple equilibria under b. However, in Proposition 1, we show that there is a highest equilibrium, which is pareto dominant. Assuming the highest (lowest) equilibrium will be played, an integration level b^* denotes a flexible union equilibrium policy

[11] Alternatively, we can bound the policy space, $t_i \in [0, 1]$, interpreting the maximum, $t_i = 1$, as becoming a single country. All results go through in this setting.

[12] 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.

[13] If a country is indifferent between two policies, we assume that the country breaks ties in favor of the lower 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 issues for the union. In the EU case, for example, a single market with common external tariffs and a union with free movement of goods, services, capital, and labor. Flexible union also allows further integration, like joining a currency area with common monetary policy. Further integration, while not mandated, allows some countries to integrate more than others in some domains, for example defense or welfare policies, or deepen their integration in core areas, for example setting up single regulatory agencies for product markets or banking. Nowadays, the EU is moving in this direction, as some countries are considering further harmonization in defense, immigration, banking, and social welfare.

2.3. Equilibrium

We start by characterizing equilibria under the three integration protocols, as this allows comparing the equilibrium policies and countries' preferences over these methods (Section 3). For the rest of the paper, we use m to denote the country with the median type, γ_m (if $|U|$ is even, $m = |U|/2$).

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

1. *In non-union integration, there are highest and lowest equilibria, $\overline{t}^*(\gamma)$ and $\underline{t}^*(\gamma)$, that are increasing in γ_i for all i. $\overline{t}^*(\gamma)$ is the pareto dominant equilibrium.*
2. *In rigid union, the most preferred policy of the median country is the Condorcet winner:* $r^* = \arg\max_r u_m(r, \ldots, r, \gamma_m)$
3. *In 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 extremal equilibrium selection, the preferred policy of the median is the Condorcet winner. That is,* $b^* = \arg\max_b u(\overline{T}(b, \gamma), \gamma_m)$

The characterization of extremal equilibria in non-union integration (part 1) is a direct consequence of increasing differences and follows from standard arguments in supermodular games (*e.g.*, Topkis (1979)).[14] In the rigid union (part 2), countries' preferences over the union policy satisfy single crossing: countries with higher preference prefer higher integration levels. The characterization of the Condorcet winner follows from the median voter theorem with single-crossing preferences (Gans and Smart, 1996). In the flexible union (part 3), the characterization of equilibrium (for a given b) is analogous to the non-union integration.

[14] See also Vives (1990) and Milgrom and Roberts (1990) for applications in oligopoly competition, arms' races, and search models.

However, in a flexible union, changes in b affect the endogenous equilibrium integration and increasing differences does not directly imply that the preferences of countries satisfy single crossing. However, we show that increasing differences still puts enough structure to the preferences and a majority votes for the policy the median prefers. For the rest of the paper, we concentrate on the highest equilibrium for non-union integration and flexible union.

3. Comparison of Integration Methods

Using Proposition 1, we compare the three integration methods. We commence comparing rigid and flexible integration methods from the perspectives of countries with different types. Most importantly, Proposition 4 demonstrates how flexible protocols (i) can increase the union-wide integration and (ii) be preferred by the higher-type countries, explaining an important regularity in the adoption of flexible integration policies such as Schengen and Eurozone.

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

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

The following table gives integration levels and payoffs of countries under different protocols.[15]

	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 1: **Integration Levels and Payoffs under Different Integration Methods.**

In rigid union, the most preferred integration of the median, country 3, is implemented; note that the median country has a low preference for integration. In non-union, all countries

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

freely choose their integration levels. Table 1, 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 6 in the Appendix A.]

Proposition 2. *Suppose that u_i is continuously differentiable and strictly increasing in t_{-i}. Then for each γ_m, there exists an ϵ 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 $i \in U$.*

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 1, 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 3. *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.*

Flexible Union vs Rigid Union In a rigid union, the median country (effectively) chooses r^* by considering the benefits of a higher policy, the *direct* increase in the actions of others, and costs that come with increasing its action. Consider r^* as a potential flexible union policy. First, under a flexible union, some higher type countries may choose deeper integration and integrate more than r^*. As actions are complements, this, in turn, causes the median country to prefer a higher policy, pushing towards deeper integration. Second, in flexible union, the equilibrium actions of higher type countries that integrate more than r^* depend *indirectly* on the equilibrium policy. It is possible that higher type countries choose high actions regardless of flexible union policy determined by the median. When this indirect effect is weak, a higher policy has less of an effect on the equilibrium integration levels under flexible union compared to rigid union; this lowers the benefit of a higher policy and pushes the median country to pick a lower equilibrium policy. However, it is also possible that this indirect effect is strong and moves the preference of the median towards higher integration. This is what happens under the parametrization of Example 1 (see Table 1, row (3)); flexibility causes the median to prefer a higher union policy, raising the integration levels and payoffs for all countries. The second *indirect* channel may cause the equilibrium policy to be higher or lower under flexible union. Therefore, flexible union may result in either higher or lower union policy compared to rigid union. Both $b^* \geq r^*$ and $b^* < r^*$ are possible.

When $b^* \geq r^*$, all countries' actions are higher under flexible union. The median country can always choose r^* as the flexible union policy and therefore prefers flexible to rigid union. Moreover, when $b^* \geq r^*$, any country with higher than the median type also prefers the higher integration that comes with flexible union. Conversely, when $b^* < r^*$, all countries with lower type than the median prefer flexible union, since they have lower preferences for integration than the median and can integrate at b^* instead of r^*.

Proposition 4. *The comparison between b^* (flexible union minimum threshold) and r^* (common rigid-union integration) is ambiguous. A majority of countries prefers flexible union to rigid union.*

- *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) median country and all countries with lower types prefer flexible union.*

Comparison. Strategic Substitutes vs Strategic Complements. It is instructive to compare our results with the case where countries' actions are strategic substitutes. First, with strategic substitutes, in a flexible union, higher type countries choose actions higher than the median. This causes the median country to choose a *lower* integration level, as effectively low-type countries and the median "free-rides" on the investments of higher-type members. Second, as in our strategic complements setting, the direct effect is weaker under a flexible union since fewer countries' decisions depend directly on the equilibrium policy. Third, under strategic substitutes, a higher policy level leads to *lower* actions from the higher type countries, further lowering the incentives of the median to choose a higher policy. All three channels reduce the equilibrium policy and therefore, $b^* < r^*$; flexibility (always) causes low-type countries to free ride on the contributions of higher type countries. Hence, low, rather than high-type countries, prefer flexible union, while rigid union is preferred by the higher type countries.[16]

Conversely, when policies are complements, while a majority always prefers flexible to rigid union, it is not *ex-ante* clear which countries select more flexible arrangements. If integration is higher under flexible union, then countries with higher types prefer it, as they can (and will) integrate further while maintaining the benefits from the integration of low-type countries. Proposition 4, we believe, is consistent with the dynamics of the EU, as many countries of the "core" (like Germany and France) have promoted flexible union, allowing many (Eastern and Southern) European countries to join, while integrating themselves more (via monetary unification, for example). These developments cannot be easily explained by public goods models with strategic substitutes. In contrast, if the policy level is lower under flexible union, then countries with lower types prefer it. This (more direct) result - that echoes the Alesina, Angeloni, and Etro (2005) - follows from the fact that low integration type countries prefer a flexible union with a low bound, as they pay a cost from integrating to much-higher-than-desired level.

4. International Unions. Formation and Enlargement

In this section, we expand our theoretical model to study two major issues regarding international unions, formation (4.1) and enlargement (4.2).

4.1. Union Formation

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.

[16] Alesina, Angeloni, and Etro (2005) provide a formalization of this result (their Proposition 4), 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 or 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."*.

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. Members choose their actions $t_i \in [b, \infty)$ in flexible union and integrate at r^* in rigid union. Integration levels of non-members are 0. Payoffs are realized.

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 2. *There are six countries, $N = \{1, 2, 3, 4, 5, 6\}$, with the following preferences over integration: $\gamma_6 = 1.7$, $\gamma_5 = 1.6$, $\gamma_4 = 1.41$, $\gamma_3 = 1.4$, $\gamma_2 = 1.23$ and $\gamma_1 = 1.22$. The utility function is:*

$$\hat{u}(t_i, t_{-i}, \gamma_i) = \sum_{j \neq i} (t_i t_j)^{\frac{\gamma_i}{2}} - t_i^2 \tag{4}$$

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

$$u^k(r, \gamma_i) = (k-1) r^{\gamma_i} - r^2 \tag{5}$$

The following equation gives the equilibrium policy with k members and median country type γ_m:

$$r^*(k, \gamma_m) = \left(\frac{(k-1)\gamma_m}{2} \right)^{\frac{1}{2-\gamma_m}} \tag{6}$$

Table 2 gives the equilibrium policies and payoffs for different rigid unions.

Union	Equilibrium Policy	u_5	u_4	u_3	u_2	u_1
$\{5, 6\}$	0.56	> 0	0	0	0	0
$\{4, 5, 6\}$	3.28		-0.12	0	0	0
$\{3, 4, 5, 6\}$	3.56			> 0	0	0
$\{2, 3, 4, 5, 6\}$	5.80				1.14	0
$\{1, 3, 4, 5, 6\}$	5.80				0	0.53
$\{1, 2, 3, 4, 5, 6\}$	8.46				-2.45	-3.91

Table 2: **Equilibrium Integration Levels and (relevant) Payoffs, Example 2**.

We now explore which of these are indeed equilibria. We start by considering a two member union of countries $\{5, 6\}$, row (1). The equilibrium policy is 0.56, 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* 3.28*, due to the complementary nature of integration. However, for country* 4*, the payoff from joining is negative,* -0.12*; the payoff is even lower for lower type countries. Therefore,* $U = \{5, 6\}$ *is an equilibrium union; there is no equilibrium with* 3 *members.*[17]

Next, we consider a four country union of $\{3, 4, 5, 6\}$ *(row 3). The median country is* 4*; equilibrium policy is* 3.56*, 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* 5.80*, 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* 8.46*. 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.*[18]

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.[19]

[17]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.

[18]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).

[19]This result suggests that EU's 12 members incentive to integrate further increased, when the EU ex-

Therefore, in any equilibrium union, all countries with types above the median country are members. The following proposition formalizes this result.

Proposition 5. *If U is an equilibrium union with median country m and $\gamma_i \geq \gamma_m$, then $i \in U$.*

This results is in line with the initial formation of the EU, where the core countries with higher preferences of integration have initiated the process.

4.2. Enlargement

We now generalize the model to allow for enlargement. There are two types of countries: $I \subset N$ denotes the initial members, before potential enlargement; $C = N \setminus I$ denotes candidates. We study the following extensive-form game:

1. Candidate countries decide whether or not 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.
4. Union members decide the equilibrium policy b^* (r^*) of the flexible (rigid) union with majority voting.
5. Members choose their actions $t_i \in [b, \infty)$ in flexible union and integrate at r^* in rigid union. Integration levels of non-members are 0. Payoffs are realized.

We focus on enlargement with unanimity, since this is the paradigm that the EU, and other international unions, has followed since inception. Before giving the propositions, we discuss a couple of examples to illustrate the main mechanisms.

Example 3. *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 $\hat{u}$ from Example 2.*

Union	Equilibrium Policy	u_2	u_1
$\{2, 3, 4, 5\}$	5.06	2.50	0
$\{1, 2, 3, 4, 5\}$	9.00	1.97	0.16

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

panded 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.

Without enlargement, the median country is 3 *and the equilibrium policy is* $r^* \approx 5.06$*, giving country* 2 *a payoff of* 2.5*. If country* 1 *joins, the median country is unchanged. However, the equilibrium integration increases to* $r^* = 9$*. Even under this higher integration level,* 1 *prefers to join. However, due to its low preference for integration, the payoff of country* 2 *drops to* 1.96 *and the candidacy of country* 1 *will be rejected by country* 2*.*

Example 4. *There are five members,* $I = \{2, 3, 4, 5, 6\}$*, and two candidates,* $C = \{1, 0\}$*. The countries' types are:* $\gamma_6 = 1.8$ $\gamma_5 = 1.6$ $\gamma_4 = 1.4$*,* $\gamma_3 = 1.24$ $\gamma_2 = 1.15$*, and* $\gamma_1 = 1.1$ $\gamma_0 = 1$*. The utility function is* $\hat{u}$ *from Example 2.*

Union	Equilibrium Policy	u_6	u_2	u_1	u_0
$\{2, 3, 4, 5, 6\}$	5.56	56.87	0.42	0	0
$\{1, 2, 3, 4, 5, 6\}$	4.43	53.27	4.89	6.07	0
$\{0, 1, 2, 3, 4, 5, 6\}$	5.63	103.00	16.02	8.45	2.07

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

Without enlargement, the median country is 4 *and the rigid union equilibrium policy* $r^* \approx 5.56$*, which gives country* 6 *a payoff of* 56.87*. First, suppose that country* 0 *is not a candidate. If country* 1 *joins, then country* 3 *becomes the decisive median, which reduces the equilibrium policy to* $r^* \approx 4.43$*. Even with more countries, the high-type country* 6 *is worse off (with a payoff of* 53.27*) due to the lower integration policy. Therefore, it rejects the candidacy of* 1*. Second, suppose that both* 0 *and* 1 *are candidates. If both countries are admitted, the median country type drops from* 1.4 *to* 1.24*. However, the equilibrium policy increases from* 5.56 *to* 5.63 *due to more countries and complementarities. Moreover, under the larger union, both the high-type country* 6 *and the low-type country* 2 *are better off; this implies that countries* 3*,* 4 *and* 5 *are also better off. Consequently, the union will admit the joint candidacy of* 0 *and* 1*. Thus, the larger union that includes all countries is an equilibrium.*[20] *However, no enlargement is another equilibrium that obtains when neither of* 1 *and* 0 *applies. This is an equilibrium as if either* 1 *or* 0 *deviates, their candidacy will be rejected, so there is no profitable deviation for both countries.*

So the equilibrium policy depends on the enlargement. As more countries integrate, the union's policy changes, reflecting two distinct mechanisms. First, as more countries integrate, all countries prefer higher integration; this mechanism pushes up the union's equilibrium policy. Second, the median country changes. If new members have lower preferences

[20] This example shows that admitting at the same time more countries make balance the tradeoffs across countries with quite different preferences for integration, consistent with various EU enlargement rounds when the union admitted many countries (e.g., 2004 round in Eastern Europe, 1995 round with three countries).

for integration (as is currently the case with the EU and accession countries in the Balkans), the median's lower preference for integration brings down the union's equilibrium 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. These theoretical results stand in clear contrast with the outcome under strategic substitutes. In that setting, admitting lower type countries *always decreases* the equilibrium policy and all countries approve enlargement when the change in the median country is limited (Alesina, Angeloni, and Etro (2005) Proposition 2). In Example 3, the median country is the same under both union compositions, but country 2 rejects the enlargement. The discussion and examples highlight two potential roadblocks towards enlargement:

1. High-type members blocking entry as they do not want a reduced level of integration. This is the (direct) mechanism often discussed in media and policy circles when the EU expanded in the South in the 1980s, Eastern Europe and the Balkans in 2004 and 2007. Besides, many policymakers and commentators make this point, when discussing future EU enlargement with Albania, Serbia, Kosovo, North Macedonia, and Turkey, nowadays. [This mechanism is present in example 4.]
2. Low-type members blocking entry as they do not want increased integration after enlargement. This novel mechanism, present in example 3, is not much-discussed. Nonetheless, this mechanism is 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 at the Balkans and the East.

Example 4 also demonstrates the possibility of equilibrium multiplicity, which is a direct consequence of strategic complementarity. While it is always challenging moving from the abstract model to the complex reality of European politics, we note that most enlargement rounds entailed many countries joining in at the same time. Spain and Portugal in 1986, Sweden, Finland, and Austria (in 1996), which during the Cold War (until 1990) maintained a"neutral position". During the Eastern Enlargement of 2004, the EU admitted eight Transition countries (the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, and Slovenia), plus Malta and Cyprus. And in 2007, the EU admitted jointly Bulgaria and Romania. Arguably all these countries were quite dissimilar to the existing members at the time, with low integration types in our framework. Yet the considerable increase in members, coupled with strategic complementarity, made even high integration members better off.

Efficiency and Enlargement 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. To make a prediction about the union, 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 (coal and steel) and over time the union admitted new members, while integration deepened considerably. Moreover, in Proposition 5 we show that during the union formation phase all countries that have 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 6, 7, 8 and 11).

Assumption 3. *Initial union members are more numerous and have higher types than candidates. That is, (i)* $|C| < |I|$ *and (ii) if* $i \in I$ *and* $j \in C$*, then* $\gamma_i > \gamma_j$.

The following proposition shows that under Assumption 3, there is a set of payoff-equivalent initial union optimal equilibria.

Proposition 6. *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.*

Proposition 6 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.

Before considering what determines enlargement in this context, we introduce some notation. For the rigid union case, the payoff of a country depends on the equilibrium integration level r^*, the number of members $|U|$, and the country's type. Let r_S denote the integration vector where S countries integrate at r and others do not integrate. $u^R(r, \gamma, S) \equiv u(r_S, \gamma)$ denotes the payoff of a country with type γ in an S member union and union policy r.

As Example 4 demonstrates, enlargement has two main effects. First, integration increases due to complementarities. Second, the median country has a lower type, which reduces integration. Whether integration increases or decreases then depends on the difference between the median country in the initial union and the final union. If the type of the median country in the final union is high enough, then integration increases after enlargement that is preferred by the high-type countries. The following proposition formalizes this discussion.[21]

[21]We prove Propositions 7 and 8 for the rigid union case as under flexible union, the effect of a higher union policy on the integration levels of higher type countries is not necessarily monotone in union size.

Proposition 7. *Consider rigid union and assume $u^R(r, \gamma, S)$ is continuously differentiable in r and satisfies strictly increasing differences in integration level r and union size S. Let γ_m and γ'_m denote the types of the median countries under 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, whenever integration increases after enlargement, initial members with above-median types prefer enlargement.*

Comparison. Strategic Substitutes vs Strategic Complements. It is instructive to compare this result with Alesina, Angeloni, and Etro (2005) [their Proposition 2]. Under strategic substitutes, enlargement to the periphery is *always* accompanied by lower integration across the union and enlargement happens if and only if the change in the median is small enough. However, in our setting integration may increase even after enlargement towards the periphery due to complementarities. The change in the median country does not determine enlargement, but determines whether the integration increases or not, which in turn determines whether candidates are admitted or not. We believe this is in line with EU's history where enlargement went handy with deeper integration, often supported by core countries. In particular, if the union policy after enlargement stays the same, then all countries are strictly better off, as they are integrating at the same level but now more countries are integrating. If the utility function is continuous, a sufficient condition for enlargement is a small change in the union policy after enlargement:

Proposition 8. *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.*

5. Non-member Integration

In this section, we allow non-members to integrate with the union, as the EU and most other international unions integrate on various policy areas with non-members. We start with providing some motivation based on the evolution of European integration. Second, extending the model for non-member integration, we study countries' benefits from joining the union or staying out but with some (endogenous) degree of integration with the union. Third, we re-examine enlargement with non-member integration to study in a more realistic setting the trade-offs of the ongoing debate on EU's potential expansion in the Balkans, Turkey, and elsewhere. Fourth, we leverage the model with non-member integration to study exit, a topical issue given UK's decision to leave the EU and the rise of euro-skepticism across the continent.

5.1. Motivation

Assuming that any country not a part of the union does not integrate was a useful starting point, as it features in much of earlier research (e.g., (Alesina, Angeloni, and Etro, 2005; Roberts, 1999)), allowing comparability with earlier contributions. However, international unions do allow for some integration with non-member countries. The EU has treaties with non-member countries covering an increasingly large set of domains. For example, Turkey and the EU have a *Customs Union*, an enhanced trade agreement allowing imports-exports to flow across the border freely. Norway, a non-member, has access to EU's single market, with some exceptions on agriculture, fishing, and food. Besides, candidate countries, like Serbia, Albania, Montenegro, the Republic of North Macedonia, and Turkey integrate with the EU on 35 domains, *accession chapters*, covering company law, public procurement, energy, taxation, financial services, consumer and health protection, among others. In each area of the *acquis*, candidate countries are *"required to adapt their administrative and institutional infrastructures and to bring their national legislation into line with EU legislation in these areas."*[22]

Integrating with non-member countries is potentially beneficial for members and non-members for (at least) two reasons. First, given Proposition 5, when a union enlarges towards its periphery, the median country has a lower type. As non-member countries do not have voting rights, integrating with them as non-members may help keep integration high without moving the decisive median. Therefore, non-member integration has a potentially desirable feature from the viewpoint of members. This may explain why the EU core countries have historically favored expanding the number and depth of deals with non-member countries. For example, Germany has been historically a proponent of deepening ties with Turkey. Second, countries whose preferences are far from the union's policies may choose to integrate as non-members. Opting out of the union, but without totally excluded may be optimal, balancing the costs and benefits of integration under (relatively) low preferences for integration. Reasonable examples include, we believe, Switzerland and Norway, which have been integrating with the EU in various domains, but without joining, as they prefer keeping national policy-making in some core areas.

5.2. Non-member Integration Restrictions

5.2.1. Set-up

For the rest of the paper, impose some further assumptions on the payoffs.

[22]These areas are conceptually much closer to complementary policies rather than investments in a common public good.

Assumption 4. *$u(t_i, t_{-i}, \gamma_i) = \sum_{j \neq i} \tilde{u}(\min\{t_i, t_j\}, \gamma_i) - c(t_i, \gamma_i)$ where $\tilde{u}$ and c satisfy differentiability and supermodularity conditions.*[23]

The first part of Assumption 4 states that the payoff of integration is additively separable across countries. The payoff of country i from integrating with country j depends on the integration levels of the two countries, t_i and t_j, while the total payoff of i is the sum of its integration with all countries. The additive separability of payoffs makes the analysis tractable, allowing to study the dynamic union formation and enlargement model with non-member integration, without sacrificing much generalizability. The second part of the assumption suggests that benefit of country i from integration with country j depends on $\min\{t_i, t_j\}$, the extent of common integration (aligned with our interpretation of t_i as the integration levels).

$N = \{1, 2, \ldots, |N|\}$ denotes the finite set of countries and $U \subseteq N$ union members. Besides, $\overline{o}$ denotes the maximum integration non-member countries can choose. In our analysis, we take $\overline{o}$ as given and study its effect on the size and the policies of an equilibrium union.[24] We modify the union formation game of Section 4.1 to allow for non-members to integrate and analyze the Subgame Perfect Equilibrium (SPE) of the following game.

1. Countries (in N) decide to become a member the union or not, forming the union U.
2. Members decide the equilibrium union policy b with majority voting.
3. Countries choose their actions.
   - If $i \in U$ [member country], then i chooses an integration level $t_i \in [b, \infty)$.
   - If $i \notin U$ [non-member], then i chooses an integration level $t_i \in [0, \overline{o}]$.

5.2.2. Implications

We start by analyzing the relationship between non-members and the union. A union U 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 main function. Moreover, whenever a union is ineffective, all members are indifferent between being members or integrating as non-members. The following proposition shows that restrictions on the degree of integration among non-members is necessary for a union to be effective.

[23]We assume that $\tilde{u}$ is concave, differentiable and satisfies increasing differences, strictly increasing in both arguments and $u(t_i, t_{-i}, 0) = 0$ for all t. c is strictly convex, differentiable, satisfies strictly decreasing differences and strictly increasing in t_i and strictly decreasing in γ_i.

[24]Proposition 9 shows why a non-trivial $\overline{o}$ is necessary in this setting.

Proposition 9. *Let U be an equilibrium union with union policy b^*. If $\bar{o} \geq b^*$, then U is ineffective.*

To prove Proposition 9, we first show that when non-member integration is unrestricted, countries with integration preferences below the median country do not have an incentive to join the union. Joining increases their integration to the union policy and these countries prefer to integrate as non-members, as they are allowed to choose their most preferred integration, 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. All such unions are ineffective, as they do not increase integration compared to non-union integration.[25] Proposition 9 shows that the policies towards non-members are important determinants of the union size and scope. However, when the non-member integration level is restricted (as in the EU), then lower type countries have an incentive to join the union, integrating more. The following example illustrates this.

Example 5. *There are four countries, $N = \{1, 2, 3, 4\}$, with the following types: $\gamma_4 = 1.6$, $\gamma_3 = 1.5$ $\gamma_2 = 1.2$ and $\gamma_1 = 1.1$. The utility function is:*

$$u(t_i, t_{-i}, \gamma_i) = \sum_{j \neq i} (\min\{t_i, t_j\})^{\gamma_i} - t_i^2 \tag{7}$$

If all countries join, $U = \{4, 3, 2, 1\}$, the equilibrium integration is $r^* \approx 2.08$. Let u_1^{in} denote the payoff of the low-integration type country 1 if it enters the union. We compute this payoff as $u_1^{in} \approx 2.38$. If, however, three countries join, $U = \{4, 3, 2\}$, then equilibrium policy rises to $r^* = 2.25$. Country 1 solves the following problem to determine its non-member integration level t, if it stays out.

$$\max_{t \leq \bar{o}} 3t^{\gamma_1} - t^2 \tag{8}$$

Plugging in $\gamma_1 = 1.1$, this function is maximized at $t \approx 1.74$ and is increasing in $[0, 1.74]$. Therefore, country 1 chooses $\min\{1.74, \bar{o}\}$ as its integration level, obtaining a payoff of 2.48, which is greater than its member integration payoff, $u_1^{in} \approx 2.38$.

Let o^* denote the value that solves $3\bar{o}^{1.1} - \bar{o}^2 = u_1^{in}$; in other words, o^* is the non-member integration restriction that makes country 1 indifferent between joining and integrating as a

[25] Our result on the necessity of non-member integration restriction echoes the discussion 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 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."*

non-member, where $o^* \approx 1.40$. Whenever $\overline{o} > o^*$, country 1 strictly prefers integrating as a non-member to joining. Thus the four-member union $U = \{1, 2, 3, 4\}$ is not an equilibrium. But, when non-member integration is restricted (for example, $\overline{o} < 1.40$), country 1 prefers to join and the four-member union becomes an equilibrium.

The example illustrates the impact of imposing restrictions on non-member integration. A more restrictive non-member integration bound results in a larger union with less integration within members, while a less restrictive non-member integration bound allows for more integration with non-members, but results in a smaller union.

5.2.3. Efficient Non-member Integration Restrictions

Since the non-member integration restriction $\overline{o}$ is crucial for the effectiveness of the union, we analyze its efficient determination. Given a non-member integration restriction $\overline{o}$ 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* is a non-member would prefer to increase its integration level above $\overline{o}$. The following proposition characterizes the efficient levels of non-member integration.

Proposition 10. *Let U be an equilibrium union where membership incentive constraint does not bind. Then there exists $\overline{o}' > \overline{o}$ such that U is an equilibrium union under $\overline{o}'$. In this equilibrium, all countries are weakly better off and if non-member integration is restricted at $\overline{o}$, then all members are strictly better off.*

Proposition 10 shows that it is without loss of optimality to restrict attention to bounds that make the lowest type non-member indifferent between becoming a member or integrating as a non-member. Whenever membership incentive constraint does not 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 their membership status.

5.3. Enlargement with Non-member Integration

We now extend our analysis considering jointly enlargement and non-member integration. Formally, we study the SPE of the following game where I denotes the set of initial members and C denotes the set of candidates.

1. Candidate countries decide whether or not 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.
4. Union members decide the equilibrium policy b^* with majority voting.
5. Countries choose their actions.
   - If $i \in U$ [member], then i chooses an integration level $t_i \in [b, \infty)$.
   - If $i \notin U$ [non-member], then i chooses an integration level $t_i \in [0, \overline{o}]$.
6. Payoffs are realized

Enlargement is similar to union formation for the incentives of candidates about joining. If non-member integration is not restricted, then candidates with lower types than the initial members, prefer integrating as non-members to joining the union. Thus, enlargement does not happen if non-member integration is not restricted. However, in this (richer) setting, there is another mechanism limiting enlargement. As we discussed in Section 4.2, initial members may be resistant to admitting new members as they may be worse off after enlargement. 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, when the non-member integration bound is high, if high-type initial members believe that after enlargement the union policy will decrease, then they would be more willing to reject the candidates, keeping union policy high by not giving them voting power, while at the same time benefiting from integrating with them as non-members.[26]

Therefore, there are two bounds, $\overline{o}_C$ and $\overline{o}_I$ that depend on the preferences of candidates and initial members. The bound determined by the preferences of the candidates, $\overline{o}_C$, is similar to the union formation case; its assures that non-member integration is restricted so that candidates prefer to join the union. The bound determined by the preferences of the initial members, $\overline{o}_I$, makes sure that non-member integration is restricted so that the initial members have incentive to admit the candidates.

Proposition 11. *Consider an initial union I and possible enlargement $U = I \cup C$ with equilibrium policy b_U.*

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

[26]Example 7 in Appendix A demonstrates how non-member integration restrictions may be necessary for high-type initial members to accept candidates.

5.4. Exit

We leverage the flexibility of our framework with non-member integration to study exit, an increasingly topical issue after BREXIT and the often adversarial relationship of the EU with Hungary. To study exit, we allow for a preference shock, as in many countries government's and people's view towards the EU and nationalism have changed over time.[27]

5.4.1. Set-up

There is a union, $U = \{1, \ldots, |U|\}$, with $|U| > 2$. We assume that U is an equilibrium union in the union formation game described in Section 4.1 with union policy b. This makes sure that all members prefer being a union member to leaving the union under their initial types. With probability ϵ, there is a preference shock that reduces the type of a (randomly determined) country. This can be related to an event that affects the preferences of the citizens, or election of a new leader who is against integration. After the preference shock, the country can decide to exit the union and integrate up to $\bar{e}$, which denotes the upper bound of integration for countries leaving the union. Moreover, if the country leaves, all remaining members receive a cost $\kappa \geq 0$. This cost represents the time and effort needed to negotiate the exit, legislative amendments and dealing with workers from the leaving country. Formally, we analyze the SPE of the following game:

1. With probability ϵ, a random country gets a shock that reduces its type (preferences for integration) to $\gamma_l \leq \gamma_1$.
2. The country facing the preference shock decides whether or not to exit the union. If they stay as a member, they choose an integration level above b. If they exit, they choose an integration level below $\bar{e}$ and all remaining members get an extra disutility of $-\kappa$ (where $\kappa \geq 0$).
3. Payoffs are realized.

As after a shock, the type of the shocked country is γ_l, the payoff from staying in is given by $U_{in} = (|U| - 1)u(b, \gamma_l) - c(b, \gamma_l)$, while payoff from leaving the union is given by $U_{exit} = \max_{t \in [0, \bar{e}]}(|U| - 1)u(t, \gamma_l) - c(t, \gamma_l)$. In any SPE of this game, the country that receives the preference shock stays in if and only if $U_{in} \geq U_{exit}$.[28] A union U is *robust* under $\bar{e}$ if the preference shock does not lead the country to leave the union.

[27]See Guriev and Papaioannou (2022) for an overview of works on populism, euro-skepticism, and attitudes against globalization, and Guiso, Herrera, and Morelli (2016) on the role of cultural differences on EU's institutional integration.

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

5.4.2. Implications

The following proposition summarizes the core results linking union stability, the preference shock's size, and restrictions on integration with non-members.

Proposition 12. *The following statements are true:*

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 12 shows that the union is robust when the upper bound on the integration of countries exiting the union is restricted to be 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 exiting (candidate) country is significantly worse to that of members, say restricting access to the single market and requiring licenses (passports) to conduct cross-border banking, then the union becomes more cohesive (part (2)). As the magnitude of the shock increases (which corresponds to a lower γ_l), more restrictive non-member integration policies are necessary to keep the union intact (part (3)). 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. Conversely, exit is not restricted when $\overline{e} > e(\gamma_l)$.

Next, we consider countries' preferences towards exit, analyzing when 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 exit restriction ($\overline{e} = e(\gamma_l)$) is greater than without. A robust union with exit restriction results in a larger union with deeper integration, while without restrictions there is exit after a preference shock. On the one hand, as higher integration from other countries is beneficial to a country and exit restrictions keep integration higher after a shock by preventing exit, they are beneficial to countries when they do not get the preference shock. On the other hand, when countries get the shock, exit restrictions are harmful since they prevent exiting and keeping a desired level of integration. As the benefit of more integration is higher for high-type countries, the following result emerges:

Proposition 13. *If country i prefers exit restriction and $\gamma_j > \gamma_i$, then country j also prefers exit restriction.*

Moreover, the following corollary characterizes when exit restriction is pareto-improving and is adopted under majority voting.

Corollary 1. *There is a cut-off country $k(\gamma_l, \kappa)$, such that all countries with higher types 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 exit, then the benefit of a robust union is 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 greater, the benefit of a robust union for the remaining countries do not change, while the cost of exit restriction is higher for the country that gets the preference shock. The following proposition formalizes these points.

Proposition 14. *Fewer countries prefer exit restriction if the preference shock is greater or the cost of exit is smaller.*

6. Conclusion

We develop a model of international integration where countries with heterogeneous preferences decide to integrate, either joining a rigid or a 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 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. First, we study the trade-offs of each integration method, comparing equilibrium policies. Second, we examine union formation and enlargement. Unlike previous models that focus on public goods games where actions are strategic substitutes, our model can explain how flexible integration and enlargement to the periphery are spearheaded by the "core" countries and how these changes are accompanied by higher integration across the union. Our model also incorporates non-member integration to the study of international unions, allowing us to think about the incentives, pros and cons of union's tendency to have special arrangements with non-members. We show that restrictions on the integration of non-members are necessary for a union to be effective and determine the size and scope of the union. Besides, modeling non-member integration allows thinking about exit, a topical issue given BREXIT and the rise of anti-EU parties across the continent. We demonstrate that placing restrictions on the post-exit integration level with leaving countries, even severe, are necessary to maintain the robustness of the union. Our framework, therefore, implies that euroskeptic politicians pushing for their countries' exit from the EU but at the same time arguing that the leaving nations will fully maintain the integration benefits (with some "special deal") are unrealistic.

Our analysis and conceptual framework offer a useful baseline to think about the trade-offs of international integration with unions. 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. Second, one could explore how in a flexible union, different tiers with further lower bounds may increase or slow down integration.[29] Third, extending the model to include countries with different sizes may yield additional insights on enlargement and relations with non-members. Fourth, adding dynamics will shed light on union stability. Fifth, allowing for investments (by the union) on countries' preferences for integration may be useful studying how integration may promote its own constituency and opponents. Sixth, one could consider alternative voting rules for enlargement or determination of integration protocols, such as qualified majority. Finally, allowing for within-country heterogeneity on preferences for integration and integration promoting winners and losers may bring insights on euro-skepticism and how to tackle it.

References

Abramson, B., and M. Shayo (2022): "International integration and social identity," *Journal of International Economics*, 137, 103577.

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

Alesina, A., and R. Perotti (2004): "The European Union: a politically incorrect view," *Journal of Economic Perspectives*, 18(4), 27–48.

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

[29]Berglof, Burkart, Friebel, and Paltseva (2008) analyzes how threat clubs within clubs increase low-type members' contributions.

——— (2005): *The size of nations*. Mit Press.

Alesina, A., E. Spolaore, and R. Wacziarg (2000): "Economic integration and political disintegration," *american economic Review*, 90(5), 1276–1296.

Berglof, E., M. Burkart, G. Friebel, and E. Paltseva (2008): "Widening and deepening: reforming the European Union," *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.

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," *Journal of Public Economics*, 89(7), 1305–1331.

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

Eichengreen, B. (2006): "The parallel-currency approach to Asian monetary integration," *American Economic Review*, 96(2), 432–436.

Gancia, G., G. A. Ponzetto, and J. Ventura (2020): "A theory of economic unions," *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," *Journal of International Economics*, 99, S97–S113.

Guriev, S., and E. Papaioannou (2022): "The political economy of populism," *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.

Kobielarz, M. L. (2022): "A Theory of International Unions with Exits," *Working Paper*.

Milgrom, P., and J. Roberts (1990): "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica: Journal of the Econometric Society*, pp. 1255–1277.

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

Oates, W. E., et al. (1972): "Fiscal federalism," *Books*.

Roberts, K. (2015): "Dynamic voting in clubs," *Research in Economics*, 69(3), 320–335.

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," *Journal of Economic Literature*, 49(4), 1200–1229.

Spolaore, E. (2013): "What is European integration really about? A political guide for economists," *Journal of Economic Perspectives*, 27(3), 125–44.

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.

Vives, X. (1990): "Nash equilibrium with strategic complementarities," *Journal of Mathematical Economics*, 19(3), 305–321.

Appendices

A. Omitted Examples

Example 6. *There are two countries* 1 *and* 2 *where* $\gamma_1 = 1$ *and* $\gamma_2 = 1.5$

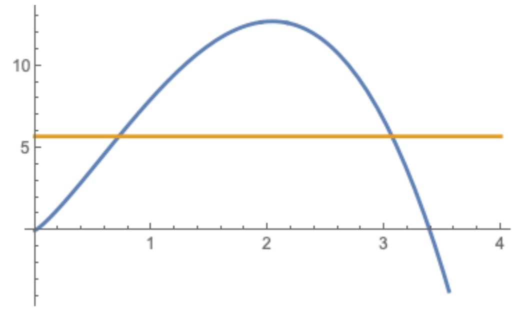

Fig. 1. **Example 7. Joining or Integrating as Non-Members. Country** 2**.** The horizontal axis denotes non-member integration level and the vertical axis denotes the integration payoff of members (Orange Line) and non-members integration payoff (Blue Line) for Country 2.

$$u_i(t_i, t_j, \gamma_i) = \gamma_i t_i t_j - \frac{1}{3} t_i^3 \tag{9}$$

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 7. *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:*

$$u(t_i, t_{-i}, \gamma_i) = \sum_{j \neq i} (\min\{t_i, t_j\})^{\gamma_i} - t_i^3 \tag{10}$$

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*

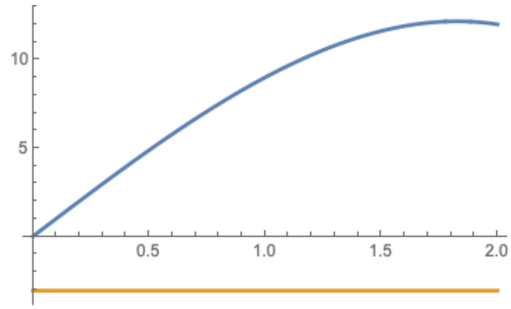

Fig. 2. **Example 7. Joining or Integrating as Non-Members. Country 1.** Horizontal axis denotes non-member integration level and vertical axis denotes membership payoff (Orange Line) and non-member integration payoff (Blue Line) for Country 1.

o gives a payoff of $9(\overline{o}^{\gamma_2}) - \overline{o}^3$. *Figure 1 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 5, illustrating how non-member integration is important for equilibrium union.*

Next, examine the policy trade-off for country 1 *plotting in Figure 2 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. 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 3. 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*

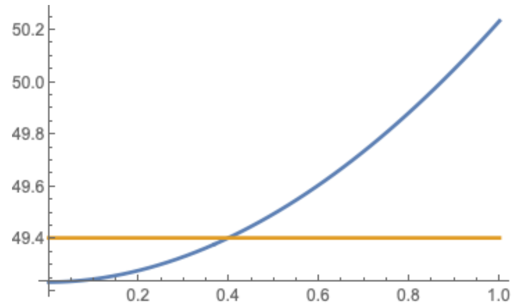

Fig. 3. **Example 7. Trade-off of High-Type Member Country** Horizontal axis denotes non-member integration level of country 2 and vertical axis denotes the payoff of country 11 when 2 becomes a member (Orange Line) and integrates as a non-member (Blue Line).

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, may be restricted.*

B. Proofs

B.1. Proof of Proposition 1

B.1.1. Part 1

From part 2 of Assumption 2, 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 $[0, t_{max}]^N$ as the strategy space eliminates only strictly dominated strategies and does not change the set of equilibria. Moreover, $[0, t_{max}]^N$ 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, it is easy to see that $\bar{t}^*(\gamma)$ is the pareto dominant equilibrium. Let $\hat{t}(\gamma)$ denote another equilibrium. For all i, we have $u(\hat{t}_i(\gamma), \hat{t}_{-i}(\gamma), \gamma_i) \leq u(\hat{t}_i(\gamma), \bar{t}^*_{-i}(\gamma), \gamma_i) \leq u(\bar{t}^*_i(\gamma), \bar{t}^*_{-i}(\gamma), \gamma_i)$, where the first inequality follows from the fact that $\bar{t}(\gamma)$ is the highest equilibrium (i.e., $\bar{t}^*_{-i}(\gamma) \geq \hat{t}_{-i}(\gamma), \gamma_i)$ and second follows from the fact that $\bar{t}^*_i(\gamma)$ is a best response to $\bar{t}^*_{-i}(\gamma)$.

B.1.2. Part 2

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.[30] If 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

$$u(t^{r'}, \gamma_i) - u(r', t^r_{-i}, \gamma_i) \geq u(t^{r'}, \gamma_j) - u(r', t^r_{-j}, \gamma_j) \tag{11}$$

$$u(r', t^r_{-i}, \gamma_i) - u(t^r, \gamma_i) \geq u(r', t^r_{-j}, \gamma_j) - u(t^r, \gamma_j) \tag{12}$$

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, $\gamma_{m'} > \gamma_m$ implies 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^*

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

B.1.3. Part 3

The characterization of equilibria under a given b follows from restricting the action space in the proof of Part 1 to $[b, t_{max}]^N$. To prove the fact that the most preferred policy of the median country is the Condorcet winner, we show that the structure of preferences of countries under rigid union is still true under flexible union. For notational simplicity, we suppress γ in $\overline{T}(b, \gamma)$. We first show that the equilibrium integration level $\overline{T}(b)$ is increasing in b.

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

Proof. Define t_{max} as in proof of Part 1 and 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:

$$\overline{T}_k(b) \equiv BR^{k-1}(\bar{t}_{k-1}, b) \tag{13}$$

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:

$$\overline{T}^*(t^b) = \lim_{k\to\infty} BR_i^k(b) \geq \lim_{k\to\infty} BR_i^k(\tilde{b}) = \overline{T}^*(\tilde{b}) \tag{14}$$

which yields the result. □

We now show 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. 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

$$u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_i) > u(\overline{T}_m(b), \overline{T}_{-i}(b), \gamma_i) \tag{15}$$

Then, as $\overline{T}_i(b) > \overline{T}_m(b)$, by increasing differences,

$$u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_m) > u(\overline{T}_m(b), \overline{T}_{-i}(b), \gamma_m) \tag{16}$$

As $\overline{T}_{-m}(b) > \overline{T}_{-i}(b)$ and $\overline{T}_i(b) > \overline{T}_m(b)$, by increasing differences

$$u(\overline{T}_i(b), \overline{T}_{-m}(b), \gamma_m) > u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m) \tag{17}$$

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

$$u(\overline{T}_i(b), \overline{T}_{-m}(\overline{T}_i(b)), \gamma_m) > u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m) \tag{18}$$

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,

$$u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m) > u(b, \overline{T}_{-m}(b), \gamma_m) \tag{19}$$

As $\overline{T}_m(b) > b$, $\overline{T}_{-m}(\overline{T}_m(b)) \geq \overline{T}_{-m}(b)$. Thus

$$u(\overline{T}_m(b), \overline{T}_{-m}(\overline{T}_m(b)), \gamma_m) > u(b, \overline{T}_{-m}(b), \gamma_m) \tag{20}$$

Again, this contradicts that b is the favourite policy of the median country. Thus, $\overline{T}_i(b) = \overline{T}_m(b)$, which finishes the proof. □

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 $\overline{T}_j(b) > \overline{T}_j(b')$

$$u(\overline{T}_j(b), \overline{T}_{-j}(b), \gamma_j) > u(\overline{T}_j(b'), \overline{T}_{-j}(b), \gamma_j) \geq u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_j) \tag{21}$$

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

$$u(\overline{T}_m(b), \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_m(b'), \overline{T}_{-m}(b'), \gamma_m) > 0 \tag{22}$$

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. $\square$

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

$$u(\overline{T}_j(b), \overline{T}_{-j}(b), \gamma_j) > u(\overline{T}_j(b'), \overline{T}_{-j}(b), \gamma_j) \geq u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_j) \tag{23}$$

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

$$u(b, \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_m(b'), \overline{T}_{-m}(b'), \gamma_m) > 0 \tag{24}$$

From optimality of $\overline{T}_m(b')$, this implies

$$u(b, \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_j(b'), \overline{T}_{-m}(b'), \gamma_m) > 0 \tag{25}$$

As $\overline{T}_{-m}(b) \geq \overline{T}_{-j}(b')$,

$$u(b, \overline{T}_{-m}(b), \gamma_m) - u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_m) > 0 \tag{26}$$

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 $\gamma_j > \gamma_m$,

$$u(b, \overline{T}_{-m}(b), \gamma_j) - u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_j) > 0 \tag{27}$$

As $\overline{T}_m(b) = \overline{T}_j(b)$, we have $\overline{T}_{-m}(b) = \overline{T}_{-j}(b)$, which implies

$$u(b, \overline{T}_{-j}(b), \gamma_j) - u(\overline{T}_j(b'), \overline{T}_{-j}(b'), \gamma_j) > 0 \tag{28}$$

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:

$$u(\overline{T}_m(t_k), \overline{T}_{-m}(t_k), \gamma_m) \geq u(\overline{T}_m(t_{k-1}), \overline{T}_{-m}(t_{k-1}), \gamma_m) \text{ for all } k \tag{29}$$

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:

$$u(\overline{T}(b), \gamma_m) - u(\hat{t}, \overline{T}_{-m}(\hat{t}), \gamma_m) \geq 0 \tag{30}$$

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

$$u(b, \overline{T}_{-j}(b), \gamma_m) - u(\hat{t}, \overline{T}_{-j}(\hat{t}), \gamma_m) \geq 0 \tag{31}$$

From optimality of $\hat{t}$ and that $\overline{T}_j(\hat{t}) = \overline{T}_m(\hat{t})$, we have:

$$u(b, \overline{T}_{-j}(b), \gamma_m) - u(b', \overline{T}_{-j}(\hat{t}), \gamma_m) \geq 0 \tag{32}$$

As $b' > b$ and $\overline{T}_{-j}(\hat{t}) \geq \overline{T}_{-j}(b)$, from $\gamma_j < \gamma_m$ and increasing differences we have:

$$u(b, \overline{T}_{-j}(b), \gamma_j) - u(b', \overline{T}_{-j}(\hat{t}), \gamma_j) \geq 0 \tag{33}$$

As $\hat{t} \geq b'$, we have $\overline{T}_{-j}(\hat{t}) \geq \overline{T}_{-j}(b')$, thus:

$$u(b, \overline{T}_{-j}(b), \gamma_j) - u(b', \overline{T}_{-j}(b'), \gamma_j) \geq 0 \tag{34}$$

This inequality, together with $b' > b$ shows that j prefers b to b', which proves the result. $\square$

B.2. Proof of Proposition 2

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 $\hat{u}(x, \gamma_m) = u(x, x, \ldots, x, \gamma_m)$. 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 $U^i_R(\gamma')$ denote the payoff of i under rigid union equilibrium when types are given by γ'.

Lemma 5. *For each $\delta > 0$, there exists an ϵ such that $|U_R - U^i_R(\gamma')| < \delta$ 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. $\square$

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

Lemma 6. *For each δ, there exists ϵ such that $U^i_N(\gamma') < U_N + \delta$ whenever $\gamma_i \in [\gamma_m - \epsilon, \gamma_m + \epsilon]$*

Proof. We first prove the following claim.

Claim 2. *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. $\square$

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

Observe that i prefers rigid union to non-union if

$$U_R^i(\gamma') - U_N^i(\gamma') = \big(U_R^i(\gamma') - U_R\big) + (U_R - U_N) + \big(U_N - U_N^i(\gamma')\big) > 0 \tag{35}$$

Given $\delta = |U_R - U_N|/3$, from in lemmas 5 and 6 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.

B.3. Proof of Proposition 3

First, note that flexible union with a trivial bound ($b = 0$) is same as non-union integration.

Lemma 7. $\overline{T}(0, \gamma) = \overline{t}^*(\gamma)$

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

Let b denote the equilibrium policy for the flexible union. To simplify notation, for the rest of the proof, we write t^* instead of $\overline{t}^*$. 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 8. *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 2. *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 8. 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 9. *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 8. 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, $\overline{T}_i(b) > b$:

$$u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_i) > u(t_i^*, \overline{T}_{-i}(b), \gamma_i) \geq u(t_i^*, t_{-i}^*, \gamma_i) \tag{36}$$

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

$$u(b, \overline{T}_{-m}(b), \gamma_m) - u(t_m^*, t_{-m}^*, \gamma_m) \geq 0 \tag{37}$$

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

$$u(t_m^*, t_{-m}^*, \gamma_m) \geq u(t_i^*, t_{-m}^*, \gamma_m) \geq u(t_i^*, t_{-i}^*, \gamma_m) \tag{38}$$

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

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

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

$$u(b, \overline{T}_{-i}(b), \gamma_i) - u(t_i^*, t_{-i}^*, \gamma_i) \geq 0 \tag{40}$$

which proves the result. □

B.4. Proof of Proposition 4

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

Lemma 10. *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 $\overline{T}_{-i}(x) \geq t_{-i}^x$. Then

$$u(\overline{T}(x), \gamma_i) \geq u(t_i^x, \overline{T}_{-i}(x), \gamma_i) \geq u(t^x, \gamma_i) \tag{41}$$

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. □

Let r denote the rigid union equilibrium policy level and t^r denote the policy vector. From Lemma 10, $u(\overline{T}(r), \gamma_i) \geq u(t^r, \gamma_i)$ 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 10.

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

$$u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_i) > u(r, \overline{T}_{-i}(b), \gamma_i) \geq u(r, t^r_{-i}, \gamma_i) \tag{42}$$

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,

$$u(b, \overline{T}_{-m}(b), \gamma_m) > u(r, \overline{T}_{-m}(b), \gamma_m) \geq u(r, t^r_{-m}, \gamma_m) \geq 0 \tag{43}$$

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

$$u(b, \overline{T}_{-i}(b), \gamma_m) - u(r, t^r_{-i}, \gamma_m) > 0 \tag{44}$$

Since $b > r$ and $\overline{T}_{-i}(b) \geq t^r_{-i}$, by increasing differences we have

$$u(b, \overline{T}_{-i}(b), \gamma_i) - u(r, t^r_{-i}, \gamma_i) > 0 \tag{45}$$

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

$$u(b, T_{-m}(b), \gamma_m) \geq u(r, T_{-m}(r), \gamma_m) \tag{46}$$

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 46 and increasing differences imply

$$u(b, T_{-i}(b), \gamma_i) \geq u(r, T_{-i}(r), \gamma_i) \geq u(r, t^r, \gamma_i) \tag{47}$$

which proves the result for the case $r > b$.

B.5. Proof of Proposition 5

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 2 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

$$u(b, \overline{T}_{-i}(b), \gamma_i) = u(b, \overline{T}_{-\tilde{m}}(b), \gamma_i) \geq u(b, \overline{T}_{-\tilde{m}}(b), \gamma_{\tilde{m}}) > 0 \tag{48}$$

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 $u(t_i, \cdot)$

$$u(\overline{T}_i(b), \overline{T}_{-i}(b), \gamma_i) > u(b, \overline{T}_{-i}(b), \gamma_i) \tag{49}$$

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 $u(t_i, \cdot)$.

B.6. Proof of Proposition 6

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 3 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 3. *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'. $\square$

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

B.7. Proof of Proposition 7

Given I, γ_m and the integration level in the initial union r_I, we have that

$$\left.\frac{\partial u^R(r, \gamma_m, |I|)}{\partial r}\right|_{r=r_I} = 0 \tag{50}$$

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,

$$\left.\frac{\partial u^R(r, \gamma_m, |U|)}{\partial r}\right|_{r=r_I} > 0 \tag{51}$$

Moreover, from the optimality of r_I, we have

$$u^R(r_I, \gamma_m, |I|) \geq u^R(r', \gamma_m, |I|), \quad \forall r' \leq r_I \tag{52}$$

As u^R satisfies strictly increasing differences in integration level and union size and $|U| > |I|$,

$$u^R(r_I, \gamma_m, |U|) > u^R(r', \gamma_m, |U|), \quad \forall r' \leq r_I \tag{53}$$

As u^R is continuous,

$$u^R(r_I, \gamma, |U|) > u^R(r', \gamma, |U|), \quad \forall r' \leq r_I \tag{54}$$

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

$$\underline{u}(\gamma) = \max_{r < r_I} u^R(r, \gamma, |U|) \tag{55}$$

$$\overline{u}(\gamma) = \max_{r \geq r_I} u^R(r, \gamma, |U|) \tag{56}$$

Claim 4. *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 55 and 56 for γ. Let $\underline{r}'$ and $\overline{r}'$ denote an element of maximizer set of equations 55 and 56 for γ'. As $\overline{u}(\gamma) > \underline{u}(\gamma)$, we have

$$u^R(\overline{r}, \gamma, |U|) > u^R(\underline{r}, \gamma, |U|) \geq u^R(\underline{r}', \gamma, |U|) \tag{57}$$

Then by increasing differences,

$$u^R(\overline{r}, \gamma', |U|) > u^R(\underline{r}', \gamma', |U|) \tag{58}$$

From optimality of $\overline{r}'$,

$$u^R(\overline{r}', \gamma', |U|) \geq u^R(\overline{r}, \gamma', |U|) > u^R(\underline{r}', \gamma', |U|) \tag{59}$$

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

$$\begin{aligned} u^R(r_U, \gamma_i, |U|) - u^R(r_I, \gamma_i, |I|) &> u^R(r_U, \gamma_i, |U|) - u^R(r_I, \gamma_i, |U|) \\ &\geq u^R(r_U, \gamma'_m, |U|) - u^R(r_I, \gamma'_m, |U|) \\ &\geq 0 \end{aligned} \tag{60}$$

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.

B.8. Proof of Proposition 8

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 $|r_U - r_I| < \epsilon$.

B.9. 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 that 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 11. *Let t^* denote a non-union integration equilibrium and $\gamma_j > \gamma_k$. Then $t^*_j \geq t^*_k$.*

Proof. 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

$$u(t^*_k, t^*_j, t^*_{-jk}, \gamma_k) - u(t^*_j, t^*_j, t^*_{-jk}, \gamma_k) > 0 \tag{61}$$

As $\gamma_k < \gamma_j$ and $t^*_k > t^*_j$, this implies that

$$u(t^*_k, t^*_j, t^*_{-jk}, \gamma_j) - u(t^*_j, t^*_j, t^*_{-jk}, \gamma_j) > 0 \tag{62}$$

Next, as $t^*_k > t^*_j$, by increasing differences, we have

$$u(t^*_k, t^*_k, t^*_{-jk}, \gamma_j) - u(t^*_j, t^*_k, t^*_{-jk}, \gamma_j) > 0 \tag{63}$$

Which contradicts the choice of t^*_j by country j. $\square$

Lemma 11 suggests an iterative procedure to characterize the equilibrium integration levels in the highest equilibrium. Define the following vector of integration levels iteratively:

$$t_i = \begin{cases} \tilde{t} & \text{where } (|U-1|)\tilde{u}'(\tilde{t}, \gamma_1) = c'(\tilde{t}, \gamma_1), \text{ for } i = 1 \\ \min\{\tilde{t}, t_{-i}\} & \text{where } (|U-i|)\tilde{u}'(\tilde{t}, \gamma_i) = c'(\tilde{t}, \gamma_i), \text{ for } i > 1 \end{cases} \tag{64}$$

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.

Claim 5. *Vector t defined above is the highest non-union integration equilibrium.*

Proof. First, observe that t_1 is the highest integration level country 1 can choose, 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

$$(|U| - i)\tilde{u}'(t', \gamma_i) > c'(t', \gamma_i) \tag{65}$$

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 $t'' > t_i$

$$(|U| - i)\tilde{u}'(t'', \gamma_i) < c'(t'', \gamma_i) \tag{66}$$

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

$$(|U| - j)\tilde{u}'(t_i, \gamma_j) = c'(t_i, \gamma_j) \tag{67}$$

which implies

$$(|U| - j)\tilde{u}'(t_i, \gamma_i) > c'(t_i, \gamma_i) \tag{68}$$

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

$$(|U| - i)\tilde{u}'(t_i, \gamma_i) \leq c'(t_i, \gamma_i) \tag{69}$$

which implies, for any $t'' > t_i$

$$(|U| - i)\tilde{u}'(t'', \gamma_i) < c'(t'', \gamma_i) \tag{70}$$

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

$$(|U| - j)\tilde{u}'(t_j, \gamma_j) \leq c'(t_j, \gamma_j) \tag{71}$$

which implies

$$(|U| - j)\tilde{u}'(t'_j, \gamma_j) < c'(t'_j, \gamma_j) \tag{72}$$

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 72 implies that j is strictly better off by decreasing its integration level, which proves that t' is not an equilibrium.

$\square$

We will now characterize the equilibrium policy and integration levels for a given union U with median country m and lowest type country j. 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 $h_{\underline{n}_i} = \min\{t_{\underline{n}_i}, \overline{o}\}$. Observe that 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 $\hat{n}_i$,

$$h_{\hat{n}_i} = \begin{cases} \min\{\tilde{t}, t_{\underline{n}_{|\underline{N}|}}, \overline{o}\} & \text{where } (|U| + |\overline{N}| + |\hat{N}| - 1)\tilde{u}'(\tilde{t}, \gamma_{\hat{n}_1}) = c'(\tilde{t}, \gamma_{\hat{n}_1}), \text{ for } i = 1 \\ \min\{\tilde{t}, h_{\hat{n}_{-1}}, \overline{o}\} & \text{where } (|U| + |\overline{N}| + |\hat{N}| - i)\tilde{u}'(\tilde{t}, \gamma_{\hat{n}_i}) = c'(\tilde{t}, \gamma_{\hat{n}_i}), \text{ for } i > 1 \end{cases} \tag{73}$$

Observe that 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, we will consider two cases.

Case 1: $(|U| - 1)\tilde{u}'(\overline{o}, \gamma_m) - c'(\overline{o}, \gamma_m) \geq 0$.

This means that even if all non-members integrate below $\overline{o}$ due to restrictions, the median still sets an integration level above $\overline{o}$, which is given by the unique solution to the following equation:

$$b^* = t \text{ where } (|U| - 1)\tilde{u}'(t, \gamma_m) - c'(t, \gamma_m) = 0 \tag{74}$$

Then b^* is the equilibrium policy in U in the first case. Equation 76 implies that, for all $t' < b^*$ and $i \in \overline{N}$,

$$(|U|)\tilde{u}'(t', \gamma_i) - c'(t', \gamma_i) > 0 \tag{75}$$

As all countries in U has higher integration levels than $\overline{o}$, $h_{\overline{n}_i} = \overline{o}$ 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 76 implies

$$(|U| - 1)\tilde{u}'(t'', \gamma_i) - c'(t'', \gamma_i) < 0 \tag{76}$$

As there are at most $(|U| - 1)$ other countries with integration levels above b^*, 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

$$h_{\overline{u}_i} = \min\{h_{\overline{u}_{i-1}}, \tilde{t}\} \text{ where } (|\overline{U}| - i)\tilde{u}'(t, \gamma_{u_i}) = c'(t, \gamma_m) \tag{77}$$

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:

$$b^* = t \text{ where } (|U| + |\overline{N}| - 1)\tilde{u}'(t, \gamma_m) = c'(t, \gamma_m) \tag{78}$$

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, all countries in $\overline{N}$ 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 $U \cup \overline{N}$, 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. Thus, 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 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 6. *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 7. *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 8. *Fix a union U with median m and union policy $b > \overline{o}$. Changing non-member integration bound from $\overline{o}$ to $o' \in (\overline{o}, b)$: does not affect the union policy. Moreover, all members and non-members who integrate less than $\overline{o}$ choose the same integration level in both cases.*

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

$$(|U| - 1)\tilde{u}'(b, \gamma_m) = c'(b, \gamma_m) \tag{79}$$

From concavity of $\tilde{u}$ and strict concavity of c, for all $o' \in (\overline{o}, b)$

$$(|U| - 1)\tilde{u}'(o', \gamma_m) > c'(o', \gamma_m) \tag{80}$$

Thus, the construction of equilibrium is still at Case 1, which means that the equilibrium policy is still determined by Equation 79. 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 9. *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 10. *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*

$$(|U|-1)\tilde{u}'(\overline{o}, \gamma_j) \geq c'(\overline{o}, \gamma_j) \tag{81}$$

Proof. For a contradiction, suppose that

$$(|U|-1)\tilde{u}'(\overline{o}, \gamma_j) < c'(\overline{o}, \gamma_j) \tag{82}$$

Let $t_j^* \equiv BR(t^U, \gamma_j)$ denote the best-response of j to t^U.

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 11. *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. □

Claim 12. *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 13. *Consider a union U under two 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 $\hat{t}_i \geq t'_i$. $\square$

B.10. Proof of Proposition 9

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 $BR_j(t_{-j}^U) = b^*$. 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 $BR_j(t_{-j}^U) \geq t_j^U$, which implies that $t_j^N \geq t_j^U$. 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 6, 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.

B.11. Proof of Proposition 10

Assume U is an equilibrium with $b > \overline{o}$ and median country m. We now extend Proposition 5 to non-member integration setting.

Lemma 12. *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 $b > \overline{o}$, we have

$$(|U| - 1)\tilde{u}'(b, \gamma_m) - c'(b, \gamma_m) = 0 \tag{83}$$

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:

$$(|\hat{U}| - 1)\tilde{u}'(\tilde{b}, \gamma_{\tilde{m}}) - c'(\tilde{b}, \gamma_{\tilde{m}}) = 0 \tag{84}$$

where $\tilde{b} > b$ as $\hat{U}$ has weakly higher median and strictly more members. As $\gamma_i \geq \gamma_{\tilde{m}}$,

$$(|\hat{U}| - 1)\tilde{u}'(\tilde{b}, \gamma_i) - c'(\tilde{b}, \gamma_i) \geq 0 \tag{85}$$

We now show that i gets a strictly higher payoff under $\hat{U}$ compared to U. As $\tilde{b} > \overline{o}$, from strict concavity of u and Equation 85, lowering the integration level of i from $\tilde{b}$ to $\overline{o}$ (or any other lower value) while keeping all other countries integration levels constant makes i strictly worse off. As $\overline{o} < b \leq \tilde{b}$, from Claim 9, 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. $\square$

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 B.9, $\overline{N}$, is empty.

Lemma 13. *Suppose that $b > \overline{o}$ and membership incentive constraint does not bind. Then there exists $o' > \overline{o}$ 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 8 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 12, $\gamma_i < \gamma_m$. 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 14. *$\hat{b} > \bar{o}$.*

Proof. First, suppose that the median does not change, *i.e.,* $\hat{m} = m$, then $\hat{b} > b > \bar{o}$ as $\hat{b}$ solves

$$(|U|)\tilde{u}'(\hat{b}, \gamma_m) - c'(\hat{b}, \gamma_m) = 0 \tag{86}$$

while b solves

$$(|U| - 1)\tilde{u}'(b, \gamma_m) - c'(b, \gamma_m) = 0 \tag{87}$$

Second, suppose that the median changes. As $\gamma_i < \gamma_m$, $\gamma_i \leq \gamma_{\tilde{m}}$. By Claim 10, we have

$$(|U| - 1)\tilde{u}'(\bar{o}, \gamma_{\tilde{m}}) \geq c'(\bar{o}, \gamma_{\tilde{m}}) \tag{88}$$

which implies

$$|U|\tilde{u}'(\bar{o}, \gamma_{\hat{m}}) > c'(\bar{o}, \gamma_{\hat{m}}) \tag{89}$$

which proves that $\hat{b} > \bar{o}$. □

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

Case 1: $t_i^U < \bar{o}$. Take $o' \in (\bar{o}, \min\{b, \hat{b}\})$ where $|\bar{o} - o'|$ is small enough that all members prefer to stay members under o'. Claim 8 implies that i and all countries with lower types choose the same integration levels, while all other choose integration levels weakly higher than i, which implies, from definition of t_i^U, i obtains the highest payoff it can get in this scenario, which is unique. Thus, if i joins the union and increases integration level to $\hat{b}$, i is strictly worse off.

Case 2: $t_i^U = \bar{o}$. we consider two sub-cases. First, suppose that

$$|U|\tilde{u}'(\bar{o}, \gamma_i) - c'(\bar{o}, \gamma_i) \leq 0 \tag{90}$$

In this case, if i joins, by Claim 11, all non-members choose the same integration levels, while the integration level of i increase from $\bar{o}$ to $\hat{b}$. Thus, from strict concavity of u, the payoff of i decreases strictly if i joins the union under $\bar{o}$. Consider $o' = \bar{o} + \epsilon$ with $\epsilon > 0$ for some small ϵ (in particular, small enough to make sure $o' < \min\{b, \hat{b}\}$). Claim 8 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

$$|U|\tilde{u}'(\overline{o}, \gamma_i) - c'(\overline{o}, \gamma_i) > 0 \tag{91}$$

Note that this implies that for small enough ϵ, under o', i chooses o' as its integration level. Moreover, from Claim 8 observe that the payoff difference between becoming a member and integrating as a non-member for i under $\overline{o}$ is given by

$$(|U|\tilde{u}(\overline{o}, \gamma_i) - c(\overline{o}, \gamma_i)) - \left(|U|\tilde{u}(\hat{b}, \gamma_i) - c(\hat{b}, \gamma_i)\right) \geq 0 \tag{92}$$

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

$$(|U|\tilde{u}(o', \gamma_i) - c(o', \gamma_i)) - \left(|U|\tilde{u}(\hat{b}, \gamma_i) - c(\hat{b}, \gamma_i)\right) \tag{93}$$

Equation 91 implies that if ϵ is small enough, value given in equation 93 is larger than the one given in 92, and thus i prefers to integrate as a non-member under $\overline{o}$. This finishes the proof of the lemma. □

From Claim 8, 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.

B.12. Proof of Proposition 11

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

$$u^i_{join} = (|U| - 1)\tilde{u}(b, \gamma_i) - c(b, \gamma_i) \tag{94}$$

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

$$(|U|-1)\tilde{u}'(t_i,\gamma_i)-c'(t_i,\gamma_i)=0 \tag{95}$$

Let $t_i^*=\min\{\hat{t}_i,\overline{o}\}$. As $\gamma_i<\gamma_{\tilde{m}}$, we have

$$(|U|-1)\tilde{u}'(\hat{t}_i,\gamma_{\tilde{m}})-c'(\hat{t}_i,\gamma_{\tilde{m}})>0 \tag{96}$$

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

$$u^i_{non-member}(\overline{o})=\max_{t_i\leq\overline{o}}(|U|-1)\tilde{u}(t_i,\gamma_i)-c(t_i,\gamma_i) \tag{97}$$

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

$$(|U|-1)\tilde{u}'(b,\gamma_m)-c'(b,\gamma_m)=0 \tag{98}$$

As $\gamma_i<\gamma_m$,

$$(|U|-1)\tilde{u}'(b,\gamma_i)-c'(b,\gamma_i)<0 \tag{99}$$

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 show that

Lemma 14. *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 under o' and $\hat{o}$. Let $\hat{t}$ and t' denote the equilibrium integration levels under o' and $\hat{o}$.

From Claim 13, 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 12, 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 15. *There exists $j \in C$ such that 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

$$(|U| - 1)\tilde{u}'(\hat{b}, \gamma_{\tilde{m}}) - c'(\hat{b}, \gamma_{\tilde{m}}) = 0 \tag{100}$$

which implies, as $b' < \hat{b}$,

$$(|U| - 1)\tilde{u}'(b', \gamma_{\tilde{m}}) - c'(b', \gamma_{\tilde{m}}) < 0 \tag{101}$$

which contradicts that b' is the equilibrium policy. Thus, there exists such a $j \in C$. $\square$

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

$$u(\hat{b}, \hat{t}_{-j}, \gamma_j) > u(t'', \hat{t}_{-j}, \gamma_j) \tag{102}$$

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 102, and $\gamma_i > \gamma_j$, for all $t'' < \hat{b}$

$$u(\hat{b}, \hat{t}_{-i}, \gamma_i) > u(t'', \hat{t}_{-i}, \gamma_i) \tag{103}$$

From optimality of $\hat{b}$ for i,

$$u(\hat{b}, \hat{t}_{-i}, \gamma_i) > u(t'_i, \hat{t}_{-i}, \gamma_i) \tag{104}$$

As $t' \leq \hat{t}$, we have

$$u(\hat{b}, \hat{t}_{-i}, \gamma_i) > u(t'_i, t'_{-i}, \gamma_i) \tag{105}$$

which shows that i is strictly worse off under o' compared to $\hat{o}$.

Case 2.1: $\hat{t}_i > \hat{b}$. From concavity of u and strict concavity of c

$$u(\hat{t}_i, \hat{t}_{-i}, \gamma_i) \geq u(t'_i, \hat{t}_{-i}, \gamma_i) \tag{106}$$

with strict inequality if $t'_i < \hat{t}_i$. Moreover, as the median integrates at $\hat{b}$ under $\hat{o}$ but b' under o', even if $t'_i = \hat{t}_i$, we have the following as at least one country with lower integration that i decreases its integration.

$$u(\hat{t}_i, \hat{t}_{-i}, \gamma_i) > u(t'_i, t'_{-i}, \gamma_i) \tag{107}$$

which shows that i is strictly worse off under o' compared to $\hat{o}$. $\square$

Lemma 14 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 14, 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.

B.13. Proof of Proposition 12

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 15. *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 $\overline{e}'$. $\square$

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 15 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 15, if $\overline{e} \leq e(\gamma_l)$, the the country stays as a member after getting a shock, which finishes the proof of part 2.

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

$$((|U| - 1)\tilde{u}(b, \gamma_l) - c(b, \gamma_l)) - ((|U| - 1)\tilde{u}(e(\gamma_l), \gamma_l) - c(e(\gamma_l), \gamma_l)) = 0 \tag{108}$$

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

$$((|U| - 1)\tilde{u}(b, \gamma_l') - c(b, \gamma_l')) - ((|U| - 1)\tilde{u}(e(\gamma_l), \gamma_l') - c(e(\gamma_l), \gamma_l')) > 0 \tag{109}$$

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

B.14. Proof of Proposition 13

Fix γ_i. Let t^e denote the most preferred integration level of a country after a preference shock, which is given by

$$(|U| - 1)\tilde{u}'(t^e, \gamma_l) = c'(t^e, \gamma_l) \tag{110}$$

Observe that as $\gamma_l < \gamma_m$, $t^e < b$. We can characterize the difference in payoff of country i with and without exit restriction as

$$E(\gamma_i, \gamma_l, \kappa) \equiv \frac{\epsilon}{|U| - 1} (\tilde{u}(b, \gamma_i) - (\tilde{u}(t^e, \gamma_i) - \kappa)) + \epsilon ((\tilde{u}(b, \gamma_l) - c(b, \gamma_l)) - (\tilde{u}(t^e, \gamma_l) - c(t^e, \gamma_l))) \tag{111}$$

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 $b > t^e$, 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.

B.15. Proof of Proposition 14

The result follows from the fact that $E(\gamma_i, \gamma_l, \kappa)$ is increasing in κ and decreasing in γ_l.