A Theory of International Unions with Exits*

Michał L. Kobielarz†

This version: January 10, 2022

Abstract

The dwindling popularity of globalization and international cooperation poses the issue of exiting an international union. An individually made exit decision is inefficient, as it neglects the losses of the other members. Fiscal transfers inside the union eliminate socially inefficient exits and restore the first-best outcome. Adjusting the union-wide policy in the face of an exit threat may prevent the exit but is not Pareto-efficient. When fiscal transfers are not possible, the union benefits from introducing exit costs during the formation process. Those costs are Pareto-optimal, despite being a deadweight loss. The paper also explores the scope for post-exit cooperation between the exiting country and the union. I show that a *soft exit* is always preferred to a no-deal exit. However, the union might be reluctant to agree to a deal if it forms a precedent for the other union members. The model sheds light on Brexit and the UK-EU negotiations, but it also applies to other international unions.

Keywords: international unions, international policy coordination, exit, EU, Brexit.
JEL Classification Numbers: D71, F02, F53, H11, H87

*For their suggestions and comments I would like to thank Fabio Canova, Edoardo Ciscato, Ferre De Graeve, Geert Dhaene, Alessandro Dovis, Paweł Doligalski, Tore Ellingsen, Luca Fornaro, Bård Harstad, Wojtek Paczos, Krzysztof Pytka, Morten Ravn, Martin Schneider, Jaume Ventura, and the participants of the Una Europa Seminar on Global Governance.

†KU Leuven; Naamsestraat 69, 3000 Leuven, Belgium; michal.kobielarz@gmail.com; Declarations of interest: none.

1 Introduction

After a long period of progressing globalization and international integration, the last decade saw a substantial reduction in the willingness of countries to cooperate internationally. The dwindling popularity of international organizations constitutes not only a reversal of the global trend for ever more international integration, but it also poses a challenge to those institutions, as they must deal with the potential exit of their members. This issue is crucial in the context of cooperation between sovereign nations, where all agreements require voluntary participation, as there is no supranational enforcement.

The most recent and prominent example of those tendencies is Brexit, i.e., the exit of the United Kingdom from the European Union. Decided by a referendum in June 2016 and a subsequent parliamentary vote, the exit process took until January 2020. It involved two general elections and many international summits and negotiations. Despite the significant efforts involved, the exit date had to be postponed three times, and the final agreement was not ready when the UK left the EU - leaving their relationship in a transitional period for over a year.

The shape of the economic and political cooperation between the UK and the EU is not the only big question remaining. At least equally important are the questions regarding the future of the European Union itself. Will more exits follow? Especially once it is clear what an exit entails and how successful the UK's post-Brexit experience is. Is Brexit going to trigger a reform of the EU? Is it going to lead to more or less integration within the EU? Those questions are essential given anti-EU and populist sentiments in several European countries.

Another example of an exit from an international union that nearly happened is the experience of Greece during the Eurozone Crisis. Between 2010 and 2015, there were multiple moments Greece was believed to exit or be expelled from the monetary union. In the end, the country never left the currency area. Instead, it received large fiscal transfers as part of international bailout agreements (Gourinchas et al. (2020)).

The erosion of popular support for international organizations is not a purely European phenomenon. Other prominent examples of such tensions include the threat of a lesser US involvement in NATO, as declared by then-president Donald Trump, the tensions between Turkey and its NATO allies, the threat of a US withdrawal from the NAFTA, the failure of the TTIP in a final stage of the negotiations process, the US withdrawal from the JCPOA, or the recent experience of international trade wars.[1] All of the above call for reconsidering

[1]The abbreviations refer to the North Atlantic Treaty Organization, the North American Free Trade

the setup of international unions, or international organizations more broadly, in the face of potential exits.

To shed light on these questions, I extend the seminal model of international unions by Alesina et al. (2005) with the possibility of exit. The union consists of a group of countries with heterogeneous preferences. By forming an international union, the countries decide to coordinate the provision of public goods, which generates international spillovers within the union. The coordination comes at the price of full conformity with the union policy. In my extension, a preference shock might make it no longer optimal for one of the countries to participate in the union and push it to exit the union. I examine the consequences of an exit decision and the policies aimed at preventing the exit.

I first establish that an individually made exit decision is suboptimal from the union's point of view, as the exiting country takes only its own welfare into account. Following this point, I explore three policies that can potentially prevent an inefficient exit. The first one is fiscal transfers within the union. Whenever a member country finds it individually optimal to exit, the other members might offer it fiscal transfers in exchange for remaining in the union. Whenever the exit is socially inefficient, fiscal transfers can be implemented in a Pareto-efficient way, with no country paying more than it gains by preventing the exit. Such an implementation eliminates the inefficiency of the exit decision and achieves the first-best scenario.

The second intervention I consider is an adjustment of the common policy. The union decides whether to change the level of public good spending closer to the preferences of the exiting country. Such a policy may prevent an exit and be supported by a majority of union members. Despite having the popular support of union members, the policy is not Pareto-efficient, nor does it restore the first-best outcome. The main intuition behind this result is that a policy change needs to be supported by the median country but might make countries on the opposite side of the preferences spectrum worse off.

The previous two policies are ex-post interventions, as the union implements them after one of the countries expresses the will to exit. An alternative approach is the ex-ante introduction of an exit clause with a stipulation of exit costs. Such costs reduce the individual incentives of countries to exit, potentially bringing it closer to the first-best set. However, once the country exits, the costs paid are a deadweight loss.

Despite the inefficient nature of exit costs, the union is likely to introduce them, as

Agreement, the Transatlantic Trade and Investment Partnership, and the Joint Comprehensive Plan of Action (Iranian nuclear deal), respectively.

they make most countries better off. In fact, all union members might ex ante support the introduction of exit costs under specific parameterizations. In such a case, positive exit costs are Pareto-optimal. This result relies on the fact that union members might value more the benefit of committing their partners to remain in the union than they value their own flexibility to exit it.

All of the considered policies have real-world counterparts. Greece received fiscal transfers to remain in the Eurozone. It also faced an exit penalty or cost, in the form of the risk of expulsion from the EU, if it wanted to exit the EMU. The UK shifted EU policies for years before leaving the union, while the Brexit negotiations have shown that exiting the EU cannot happen without a substantial economic cost.

Finally, the paper explores the scope for a post-exit relationship between the exiting country and the union. In particular, I propose a class of solutions in-between membership and exit. With this class of deals, there is always an agreement preferred to a *no-deal exit*. However, if the agreement constitutes a precedent, i.e., it is automatically available to all other union members, then the union might decide to sign a less cooperative agreement or not sign one at all. Hence, the precedent might render the optimal deal unacceptable to the union due to the risk of further members exiting the union.

The analysis of potential post-exit deals sheds new light on the Brexit negotiations between the UK and the EU. In particular, analyzing the talks through the lenses of my model implies that the EU was never interested in signing a soft deal with the UK.

Related literature. This paper contributes to the literature on international unions and international cooperation. The approach in the literature is inspired by the work on federalism by Oates (1972). Alesina et al. (2005) study the formation and size of international unions in the face of the trade-off between costs and benefits of coordination and integration. As unions consist of heterogeneous countries, they are bound to include different views on the speed and depth of integration, as in the case of the European Union. Harstad (2006) study the question of allowing a subset of countries in the union to cooperate more tightly or form an inner club. This phenomenon is also discussed by Berglof et al. (2008), who refer to it as a *club-in-the-club*. International unions are a key element of globalization. Still, the incentives to form or join a union also depend on the state of globalization, which is shown for an economic union in Gancia et al. (2020), and more broadly for empires and international unions in Gancia et al. (2021).

In this tradition, the international union in my paper can be interpreted not only as an economic union, but also more broadly as an international trade agreement (as in, e.g., Grossman et al. (2021), or Maggi and Ossa (2021)), international treaty (such as in, e.g., Battaglini and Harstad (2016), or Battaglini and Harstad (2020)), or international political and military alliance. I contribute to this broad literature by explicitly studying the possibility of exit from an international union, the policies that may prevent it, and the post-exit relationship between the exiting country and the union.

My work is not the first to study exit from international unions. The small, but growing literature on this topic, has mostly concentrated on monetary unions. Fuchs and Lippi (2006) study how the effect of a possible exit from a monetary union affects optimal monetary policy. Eijffinger et al. (2018) show that the risk of exit can generate contagion in a monetary union in the presence of uncertainty. Kobielarz (2021) analyzes how a monetary union might engage in bailouts to prevent a country from exiting. A rare example of a paper unrelated to monetary unions is the work of Maggi and Staiger (2015), who consider a milder form of exit, i.e.. the renegotiation of trade agreements, and how it affects the optimal setup of such contracts. I contribute to this literature by studying exit in a simple but general framework and discussing the policies available to the union explicitly.

Finally, the recent Brexit experience triggered the emergence of empirical and quantitative studies on the effects of Brexit and possible Brexit deals on the UK and the EU economies.[2] This paper complements this literature by providing a simple theoretical framework for analyzing the relationship between the UK and the EU. It also sheds additional light on the difficulties of negotiating a post-exit deal and the seeming suboptimality of the signed agreement.

The remainder of the paper proceeds as follows. Section 2 presents a simple model of an international union in which an exit could happen. Section 3 discusses policies that might prevent the exit. Section 4 explores the scope for a post-exit relationship between the exiting country and the union. Section 5 concludes.

2 A simple model

Consider an economy consisting of a group of $N+1$ equally sized countries that form an international union, $\mathcal{U}$. They all receive an exogenous endowment y, which they divide between private consumption and public good spending, the latter denoted by g_i. The countries are identical, except for their preference for the public good, α_i. We can write the

[2]Prominent examples of this growing literature include Broadbent et al. (2019), Born et al. (2019), and McGrattan and Waddle (2020).

per-period utility function of a representative individual in country i as

$$U_i\left(g_i, n; \alpha_i\right) = y - g_i + \alpha_i H\left(\left(1 + \beta(n-1)\right) g_i\right), \tag{1}$$

where n is the number of countries that country i is in a union with. The public good utility function $H(\cdot)$ is non-negative, increasing, and concave.

For an autonomous country, $n = 1$, the utility function boils down to

$$U_i^A\left(\alpha_i\right) \equiv U_i\left(g_i, 1; \alpha_i\right) = y - g_i + \alpha_i H\left(g_i\right), \tag{2}$$

where country i chooses the optimal level of public spending to equalize the marginal utility of private and public goods. This implies that g_i satisfies

$$H_g\left(g_i\right) = \frac{1}{\alpha_i}. \tag{3}$$

The case of an autonomous country is relevant as the outside option to union membership.

Let us concentrate on the economy when the union is already set up and restrict our attention to union members. The union's existence implies that the distribution of public good preferences is such that it is ex ante optimal for all countries to participate in the union.[3]

The union $\mathcal{U}$ is a rigid union, i.e., the member countries need to implement the same level of public good spending. Let us denote the utility of country i when it is a member country of a union consisting of $N+1$ members as $U_i^{U,N+1}$. Then

$$U_i^{U,N+1}\left(\alpha_i, \alpha_m\right) \equiv U_i\left(g_m^{N+1}(\alpha_m), N+1; \alpha_i\right) = y - g_m^{N+1} + \alpha_i H\left(\left(1 + \beta N\right) g_m^{N+1}\right), \tag{4}$$

where $g_m^{N+1}(\alpha_m)$ is the level of public good spending imposed by the union. All members vote on their preferred g, which is then binding for all of them. Since preferences are single-peaked in g_i, the level of spending chosen is the one selected by the member country with the median preferences, α_m.

$$\alpha_m H_g\left((1 + \beta N) g_m^{N+1}\right) = \frac{1}{1 + \beta N} \tag{5}$$

The superscript of g_m^{N+1} indicates that the median voter's optimal level of spending also

[3]I abstract from the process of union creation and potential third countries that might want to join the union. Both issues are discussed thoroughly by Alesina et al. (2005).

depends on the number of countries in the union.

The timing of decisions within the model is the following. First, all countries vote on the preferred spending level, g. Then, a preference shock ε is realized, i.e. the size of the shock and the country affected are made publicly known. Next, the union may take action to prevent the exit. After the unions decision, the shock-hit country decides on exiting the union. Finally, the other countries are also allowed to choose between leaving and staying.

The probability that country i is hit by a preference shock is p_i^{ε}, where $\sum_{i \in \mathcal{U}} p_i^{\varepsilon} \leq 1$.[4] The shock is drawn from a country-specific uniform distribution $\varepsilon_i \sim U[-\bar{\varepsilon_i}, \bar{\varepsilon_i}]$. Let us denote the public-good preference post-shock as $\tilde{\alpha}_i \equiv \alpha_i + \varepsilon_i$. To abstract from α_i turning negative, let us assume that $\forall_{i \in \mathcal{U}} 0 < \bar{\varepsilon_i} < \alpha_i$.

2.1 Exit decision

As long as the preference shock is not realized, no country has an incentive to exit, as they have joined the union voluntarily. After the preference shock is realized, as long as the union does not adjust its policy, the only country that might have an incentive to exit is the country hit by the shock. Only after the shock-hit country exits, is it possible that other countries might want to exit as well.[5] Therefore, I first concentrate on the decision of the member country hit by the shock.

From now onward, I denote the country hit by the shock as j. If country j exits the union, it has the right to choose its level of spending, g_j, but forgoes the spillovers from the public good spending of the other countries. At the same time, the other countries no longer benefit from the spillovers of g_j.

A country j will exit the union whenever the utility it can achieve as an autonomous country exceeds the utility it achieves in the union

$$U_j^A(\tilde{\alpha}_j) > U_j^{U,N+1}(\tilde{\alpha}_j, \alpha_m), \tag{6}$$

where $\tilde{\alpha}_j = \alpha_j + \varepsilon_j$ is the preference of country j after being hit by the shock. By definition, the opposite condition is satisfied by every country participating in the union, so also country j with a preference parameter α_j. However, it might reverse after country j is hit by a

[4] I exclude the possibility of shocks hitting multiple countries at once for tractability.

[5] The argument here abstracts from the union taking actions or adjusting its policy as an aftermath of the preference shock. Such decisions might push members who were not directly hit by the shock to consider an exit decision. I analyze such a possibility explicitly in the next sections when I consider different actions the union might take in response to the shock.

preference shock.

Proposition 1 *An individual exit decision is socially inefficient.*

Proof The net social benefit of country j staying in the union is equal to the sum of the net benefit experienced by country j - defined implicitly in condition (6) - and the net benefits of the remaining union members. An exit decision is socially optimal when this net benefit is negative, i.e.,

$$U_j^{U,N+1}(\tilde{\alpha}_j, \alpha_m) - U_j^A(\tilde{\alpha}_j) + \sum_{i \neq j, i \in \mathcal{U}} \left[U_i^{U,N+1}(\alpha_i, \alpha_m) - U_i^{U,N}\left(\alpha_i, \alpha_m^{'}\right) \right] < 0, \tag{7}$$

where $\alpha_m^{'}$ refers to the post-exit median voter. A comparison of conditions (6) and (7) reveals that the individual decision neglects the social consequences of exit, $\sum_{i \neq j, i \in \mathcal{U}} \left[U_i^{U,N+1}(\alpha_i, \alpha_m) - U_i^{U,N}\left(\alpha_i, \alpha_m^{'}\right) \right]$. ∎

Proposition 1 highlights the spillovers from a union exit. A socially optimal exit would also consider the lost benefits of the remaining members. In this sense, country j is too eager to exit, and hence, exits happen too frequently if they are decided unilaterally. Proposition 1 sets the stage for the remainder of the paper.

3 Preventing an exit

In this section, I consider some of the possible policy responses that the union can implement to limit the inefficiency of unilateral exits. Let us first define a union with a dense core, which will allow me to isolate the different effects of exit.

Definition A union is referred to as a union with a dense core, if there exists a $k > 0$ such that k members to the left of the median, and k members to the right of the median, share the preferences of the median member, i.e.

$$\alpha_{m-k} = \ldots = \alpha_m = \ldots = \alpha_{m+k}.$$

In a union with a dense core, an exit does not change the median member preferences, despite the identity of the median country changing. For the remainder of the paper, I assume that union $\mathcal{U}$ has a dense core. This assumption allows me to concentrate on the lost spillovers and abstract from strategic considerations surrounding a change of the decision-maker.[6]

3.1 Transfers

Let us first keep the assumption of the rigid union with a public spending level chosen before the shock hits. In this environment, I consider the possibility of fiscal transfers within the union. In particular, I study the case in which country j finds it optimal to exit the union, and the remaining member countries may decide to offer transfers to country j conditional on it forgoing the exit.

Let us denote a transfer from country i to country j as $\tau_{i,j}$. The transfers are successful if country j is better off inside the union with transfers than outside. A modified version of inequality (6) captures this condition

$$U_j^A(\tilde{\alpha}_j) \leq U_j^{U,N+1}(\tilde{\alpha}_j, \alpha_m) + \sum_{i \neq j, i \in \mathcal{U}} \tau_{i,j}. \tag{8}$$

Country j still needs to spend g_m^{N+1} on the public good. However, now it receives transfers from its partners as compensation for the high contribution.

When evaluating the fiscal transfers, it is essential to ensure that all countries within the union are willing to make the transfers. Hence, every union member must be at least as well off paying the transfer as being in a union of N countries without transfers. All of the above concerns are captured and formalized in the following proposition.

Proposition 2 *Whenever country j finds it optimal to exit the union, there exists a set of fiscal transfers $\{\tau_{i,j}\}_{i \neq j, i \in \mathcal{U}}$, which prevents country j from exiting and is (weakly) preferred by all other union members over the exit of j, if and only if, the exit of country j is socially inefficient.*

Proof In the appendix.

Proposition 2 highlights the power of fiscal transfers within an international union. First, the fact that the transfers are voluntary leads to them being Pareto-optimal, as no country is

[6] The dense core assumption is likely to be non-restrictive in reality. Often at the core of international unions is a group of countries with similar preferences, even if the union as a whole is heterogeneous. This assumption is more likely to be violated if we consider an international union consisting of countries that differ substantially in size and one of the large countries wants to exit, e.g., a potential exit of the US from NATO.

worse off by implementing the transfers. The result is even stronger, as coordinated voluntary fiscal transfers can prevent an exit whenever it is socially inefficient. Hence, a union with transfers can achieve the first-best outcome and shift the exit condition to inequality (7).[7]

Proposition 2 also guarantees that fiscal transfers are only applied when the exit is inefficient. This result is related to the Pareto-efficiency and voluntary nature of the transfers. If the exit is efficient, then the exiting country gains more than the other countries lose. Hence, the former cannot be sufficiently compensated by the latter for staying in the union.

There are at least two examples of international unions in which transfers are used or were used in the past to prevent countries from exiting under a rigid policy rule. The first is the implicit fiscal transfers, which were part of the bailout packages during the Eurozone Crisis. Several Eurozone members hit hard by the crisis received financial support. The assistance was more generous than the terms offered by the markets or the IMF, which Gourinchas et al. (2020) interpret as the countries receiving implicit fiscal transfers. The largest beneficiary of the assistance programs was Greece, which received transfers exceeding 40% of its 2010 GDP. Those transfers were a compensation for remaining in the monetary union and forgoing the possibility of a competitive devaluation.

The second example is related to the budget of the European Union, over three-fourths of which come from country contributions proportional to their Gross National Income and parts of their Value Added Tax collections. In the period 2014-2020, five EU members benefited from reduced contribution rates.[8] The rules governing the budget contributions resemble the uniform policy of my model, whereas the reduced contributions are, in fact, implicit transfers from the union back to those members. The reductions are officially motivated by the fact that the contributions of the wealthiest countries would be excessive. In the UK, significant budget contributions would make EU membership unpopular, preventing the UK from participating in the EU. An exit from the EU may not be as likely for the remaining beneficiaries, but a veto on the common EU budget is. Such a veto would be an exit from a critical part of the EU.

[7] There are many scenarios in which more than one set of fiscal transfers is possible, and they only differ in the distribution of the surplus from preventing the exit. Here, I abstract from issues of choosing a scheme and strategic free-riding by union members.

[8] The UK received a share of its net contributions back. Denmark, the Netherlands, and Sweden benefited from reductions in their gross GNI contributions. Germany, the Netherlands, and Sweden enjoyed reduced VAT call rates (0.15% instead of 0.30%).

3.2 Policy adjustment

Fiscal transfers can be a very efficient way of compensating country j for staying. However, they might also be politically difficult to implement, as international unions lack the authority to force members to carry out the transfers. The choice of a sharing rule for the transfers adds an extra layer of negotiations and bargaining, especially if the individual gains are unobservable. An alternative way of preventing a union member from exiting is altering the union policy to be acceptable for the country.

In terms of the model, I allow the union to adjust its public good spending level g after learning the shock's value. Country j observes the new spending level, $\tilde{g}_{m,j}$, and decides whether it still prefers to exit or remains in the union.

I assume that the union decides about the policy change by majority voting.[9] In particular, the union might choose a level of spending that was not supported by most countries before the shock. The majority supports the spending level now to keep country j in the union.

The following proposition formally states the conditions for a policy adjustment to prevent the exit and be accepted by the union.

Proposition 3 *Any public spending level, $\tilde{g}_{m,j}$ can prevent the exit of country j and be accepted by the remaining members of the union $\mathcal{U}$ if it satisfies the following conditions:*

i. Country j prefers to stay in the union under policy $\tilde{g}_{m,j}$ rather than to exit,

$$U_j\left(\tilde{g}_{m,j}, N+1; \tilde{\alpha}_j\right) \geq U_j^A\left(\tilde{\alpha}_j\right), \tag{9}$$

ii. The median country prefers the policy adjustment $\tilde{g}_{m,j}$ over the exit of country j,

$$U_m\left(\tilde{g}_{m,j}, N+1; \alpha_m\right) \geq U_m^{U,N}\left(\alpha_m, \alpha_m\right), \tag{10}$$

iii. No union member prefers exiting the union over staying under policy $\tilde{g}_{m,j}$,

$$\forall_{i \neq j, i \in \mathcal{U}} U_i\left(\tilde{g}_{m,j}, N+1; \alpha_i\right) \geq U_i^A\left(\alpha_i\right). \tag{11}$$

Proof In the appendix

[9]The new policy is chosen by the median voter, even though the median voter theorem does not apply here. This result is formalized in Lemma 7 in the appendix.

Proposition 3 puts forward the conditions for an exit-preventing policy adjustment. However, it does not determine whether such an adjustment exists, nor how large the set of possible policy adjustments is. The size of the set depends on the parametrization of the model and the distribution of preferences. Figure 1 illustrates examples of two possible scenarios.[10]

Panel a) of figure 1 displays a case where country j and the median country can agree on a policy adjustment. Any policy between $\underline{g}_m$ and $\bar{g}_j$ is preferred by both countries over an exit. In this example, all other countries in the union accept the policy change as well.

Panel b) explicitly goes beyond analyzing countries j and m and highlights the importance of the third condition in Proposition 3. In this example, any policy adjustment in the range $[\underline{g}_m; \bar{g}_j]$ can prevent the exit of country j and is acceptable to the median member. Nevertheless, there is a country, which I refer to as h, which finds the policy adjustments unattractive. In particular, country h would rather exit the union than stay if a policy to the left of $\underline{g}_h$ is implemented. This implies that the set of acceptable policies is limited to $[\underline{g}_h; \bar{g}_j]$. The preferences of country h may be such that no exit-preventing policy adjustment is possible.[11]

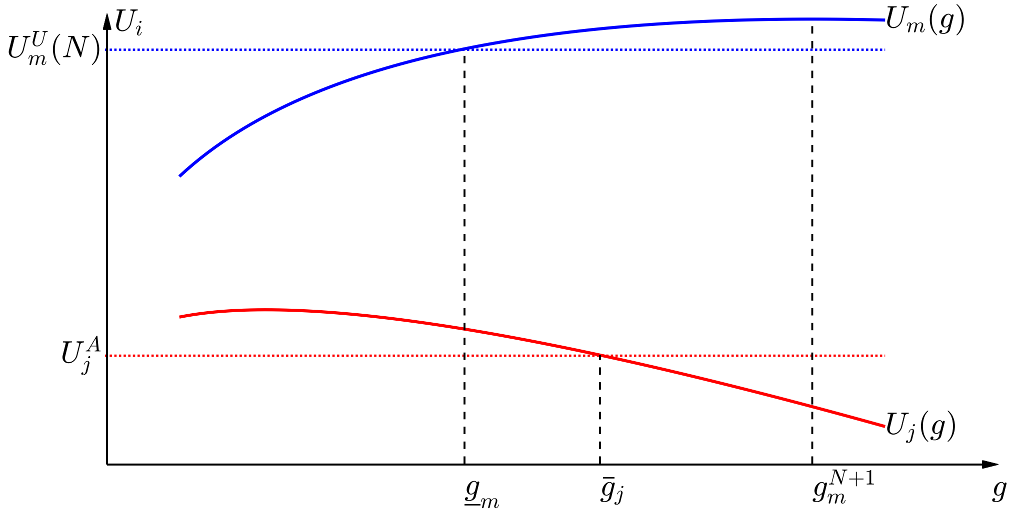
There are two critical differences between propositions 2 and 3. The policy adjustment is not conditional on social efficiency. The median member can implement a policy adjustment even in some cases where the exit is socially efficient. There are also cases where a policy adjustment cannot prevent a socially inefficient exit.

The second difference is the Pareto-efficiency of the two policies. Whereas the proof of Proposition 2 sketches a transfer scheme that guarantees that transfers are a Pareto-improvement, the policy adjustment is designed such that it benefits the median and the exiting member countries. Most member countries are also better off, as the policy needs to obtain popular support in the union. However, in many cases, the welfare improvement enjoyed by those countries happens at the expense of other members.

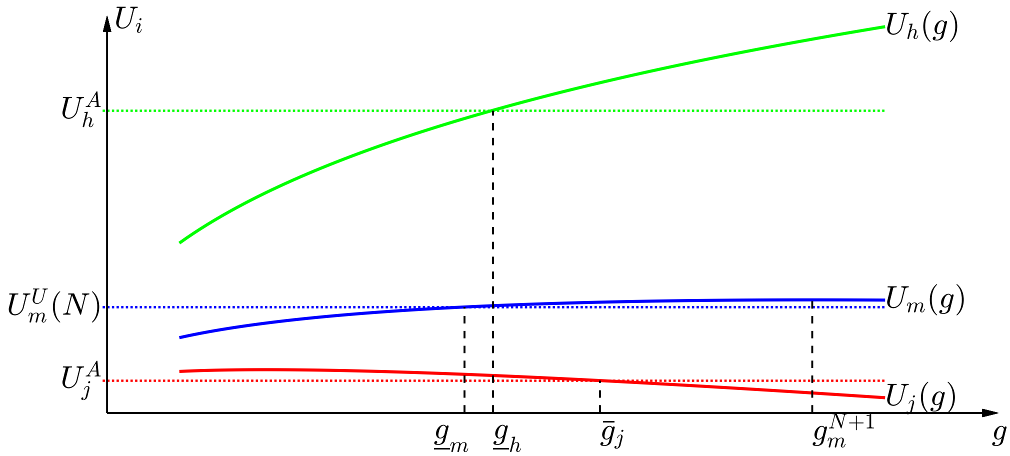
One example of such a welfare shift is the case in which the exiting country j is the country with the lowest preference for public good, i.e. $\forall_{i \neq j, i \in \mathcal{U}} \alpha_i \geq \alpha_j$. In such a case, country j wants to exit because g_m^{N+1} is too high for its preferences. Therefore, if the union wants to keep the country in, then $\tilde{g}_{m,j}$ has to be lower than the initial spending level g_m^{N+1}. If $(1+\beta N)\tilde{g}_{m,j} < (1+\beta(N-1))\, g_m^{N+1}$ then the policy adjustment might decrease the utility of union members with a sufficiently high preference for the public good. In this case, the welfare improvement enjoyed by country j and countries with $\alpha \leq \alpha_m$ comes at the expense

[10]The examples are solved numerically; hence, the figures present actual special cases of the model. The parametrizations of all numerical examples are included in Appendix B.

[11]This happens when $\underline{g}_h > \bar{g}_j$.



(a) Set of possible policy adjustments



(b) Set limited by high preference member

Figure 1: Exit-preventing policy adjustments.

of the high α union members.

Policy shifts in international unions are difficult to observe, as many negotiations happen behind closed doors, and the final press releases do not necessarily reflect the whole process. The usual subtlety applied in those situations makes the explicit US push to increase military spending of all NATO members the more uncommon. During the presidential campaign of 2016, then-candidate Donald Trump threatened that the US would not act to defend NATO allies who do not contribute enough to the union. Despite softening his critique on NATO as president, he repeated that any US military support is conditional on sufficient contributions (among others at the 2018 NATO summit in Brussels). He later even threatened that the

US might exit NATO if other members did not increase their payments.[12]

A more subtle example of policy adjustments within the union to accommodate a member at risk of exiting comes from the EU. From the UK joining the European Economic Community in 1973 until the Brexit referendum of 2016, the UK enjoyed disproportional political power in the EU.[13]

3.3 Exit costs

So far, I have assumed that any country may just freely exit the union after the shock hits. Depending on the type of international union, this might not always be a reasonable assumption. An exit might be costly in terms of efforts and administrative procedures or involve a penalty. One example might be the European Economic and Monetary Union (EMU), which lacks an official exit clause, making it administratively cumbersome to exit. The EMU could also penalize an exiting member by requiring it also to leave the EU.[14]

Whether it is an administrative cost to establishing the exit procedure or a penalty agreed upon by the union, exit costs affect the exit decision of a union member. Denoting the exit costs as C, we can write the modified exit condition for country j as

$$U_j^A\left(\tilde{\alpha}_j\right) - C > U_j^{U,N+1}\left(\tilde{\alpha}_j, \alpha_m\right), \tag{12}$$

where I assume that C is expressed in utility units.[15]

Comparing the above inequality with the condition for a socially optimal exit in equation (7), it is clear that for any country j, we can mimic the socially optimal exit decision if C_j is set equal to the net loss experienced by the other union members.

$$\begin{aligned} C_j &\equiv \sum_{i \neq j, i \in \mathcal{U}} \left[U_i^{U,N+1}\left(\alpha_i, \alpha_m\right) - U_i^{U,N}\left(\alpha_i, \alpha_m\right)\right] \\ &= \left[H\left((1+\beta N) g_m^{N+1}\right) - H\left((1+\beta(N-1)) g_m^N\right)\right] \sum_{i \neq j, i \in \mathcal{U}} \alpha_i - \left(g_m^{N+1} - g_m^N\right) N. \end{aligned} \tag{13}$$

The sum of the preference parameters depends on j. Hence, C_j would need to be set

[12]Benitez (2019) provides a detailed analysis of Trump's policy towards NATO.

[13]The EEC is a predecessor of the European Union.

[14]From a legal point of view, the current treaties implicitly condition an EMU exit on an EU exit. Athanassiou (2009) concludes that "a member state's exit from EMU, without a parallel withdrawal from the EU, would be legally inconceivable."

[15]Since the utility of the private good is linear, the current formulation is equivalent to C being expressed in units of the private good.

differently for each country to obtain a set of socially optimal exit decisions for all countries. When C is chosen equal across all members, then we cannot implement the set of optimal exit decisions.[16]

Imagine that the union members vote on the introduction of exit costs before realizing the preference shock. Then, each country tries to select a level of exit costs that maximizes its expected utility.

$$
\begin{aligned}
&\mathbb{E}\left[U_j|C\right] = \\
&= p_j^{\varepsilon} P\left(U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right) < U_j^A\left(\tilde{\alpha}_j\right) - C\right) \mathbb{E}\left[U_j^A\left(\tilde{\alpha}_j\right) - C | U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right) < U_j^A\left(\tilde{\alpha}_j\right) - C\right] \\
&\quad + p_j^{\varepsilon} P\left(U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right) \geq U_j^A\left(\tilde{\alpha}_j\right) - C\right) \times \\
&\quad \times \mathbb{E}\left[U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right) | U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right) \geq U_j^A\left(\tilde{\alpha}_j\right) - C\right] \\
&\quad + \sum_{i\neq j, i\in\mathcal{U}} p_i^{\varepsilon} P\left(U_i^{U,N+1}\left(\tilde{\alpha}_i,\alpha_m\right) < U_i^A\left(\tilde{\alpha}_i\right) - C\right) \max\left\{U_j^{U,N}\left(\alpha_j,\alpha_m\right), U_j^A\left(\alpha_j\right)\right\} \\
&\quad + \sum_{i\neq j, i\in\mathcal{U}} p_i^{\varepsilon} P\left(U_i^{U,N+1}\left(\tilde{\alpha}_i,\alpha_m\right) \geq U_i^A\left(\tilde{\alpha}_i\right) - C\right) U_j^{U,N+1}\left(\alpha_j,\alpha_m\right) \\
&\quad + \left(1 - \sum_{i\in\mathcal{U}} p_i^{\varepsilon}\right) U_j^{U,N+1}\left(\alpha_j,\alpha_m\right).
\end{aligned}
$$

To facilitate further analysis, let me first introduce some additional notation. I denote the threshold level of preferences $\underline{\alpha}^C$ as the preference parameter of a hypothetical union member j who is marginally indifferent between staying in the union and exiting it, under exit costs C, i.e., $\underline{\alpha}^C$ is such that

$$U_j^A\left(\underline{\alpha}^C\right) - C = U_j^{U,N+1}\left(\underline{\alpha}^C,\alpha_m\right). \tag{14}$$

In particular, $\underline{\alpha}$ is going to be the threshold value for $C = 0$.

Let me also define a distribution function for the preference parameter of union member j conditional on the union member being hit by a preference shock

$$F_j\left(\alpha\right) \equiv P\left(\tilde{\alpha}_j \leq \alpha\right). \tag{15}$$

The introduction of exit costs has three effects on country j, which are best visible when

[16] A noteworthy exception is a symmetric union, i.e., a union consisting of homogeneous countries. In the case of such a union, there exists a level of exit costs C^* that can mimic the socially efficient exit decision.

comparing the expected utility under $C = C_j$ with the expected utility under no exit costs

$$\begin{aligned}
&\mathbb{E}\left[U_j|C=C_j\right]-\mathbb{E}\left[U_j|C=0\right]= \\
&=p_j^{\varepsilon}\left[\int_{\underline{\alpha}^{C_j}}^{\underline{\alpha}}\left[U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right)-U_j^{A}\left(\tilde{\alpha}_j\right)\right]f_j\left(\tilde{\alpha}_j\right)d\tilde{\alpha}_j\right]-p_j^{\varepsilon}F_j\left(\underline{\alpha}^{C_j}\right)C_j \\
&\quad+\sum_{k\neq j,k\in\mathcal{U}}p_k^{\varepsilon}\left[F_k\left(\underline{\alpha}\right)-F_k\left(\underline{\alpha}^{C_j}\right)\right]\left[U_j^{U,N+1}\left(\alpha_j,\alpha_m\right)-U_j^{U,N}\left(\alpha_j,\alpha_m\right)\right], \qquad (16)
\end{aligned}$$

where $f_j(\cdot)$ is the density function of $\tilde{\alpha}_j \equiv \alpha_j + \varepsilon_j$, analogous to the cumulative distribution function defined in equation (15).

Firstly, higher exit costs reduce the probability of country j exiting, which is visible in the first element on the right-hand side of equation (16). Whenever $C > 0$, if $\tilde{\alpha}_j$ takes a value in the range $\left[\underline{\alpha}^{C_j}; \underline{\alpha}\right]$, then country j remains inside the union despite the lower utility it enjoys there. I refer to this effect as the *burden of commitment*. Secondly, country j has to pay the exit costs, conditional upon exiting. Since the payment of the costs does not any other country, this is the *deadweight loss of commitment*. Finally, higher exit costs makes all other members less likely to exit, which increases the expected utility of country j. I refer to the last effect as the *value of committing others*.

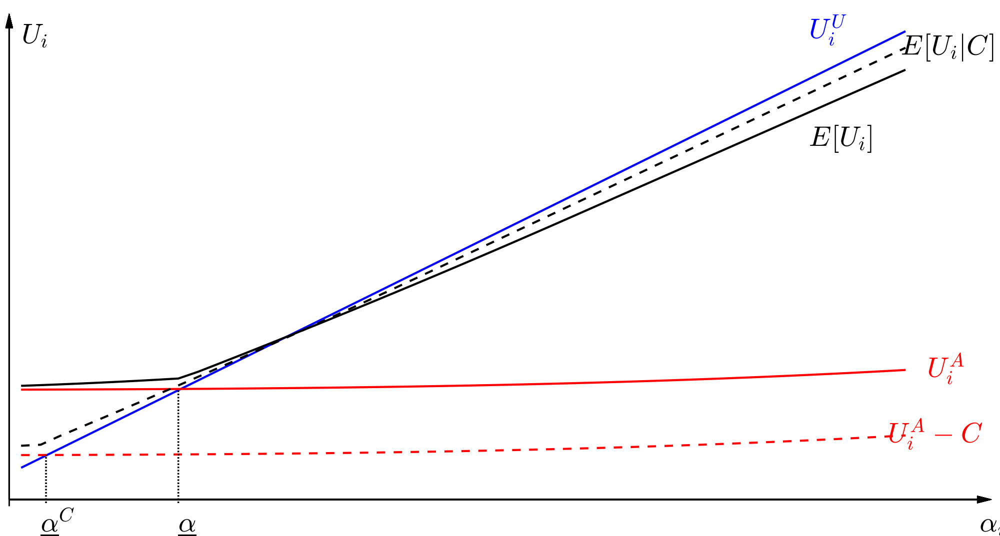


Figure 2: The utility of country i as a function of its α_i. Comparison of expected utility without exit costs (black solid line) and with exit costs C (black dashed line).

Figure 2 illustrates the expected utility of a union member as a function of its α. The solid lines represent the case of no exit costs. The blue and red lines are the utility of being in the union and independent, respectively. The black curve draws the expected utility, which is a combination of the two utilities. The value of exiting is particularly influential around $\underline{\alpha}$, the value of α at which country i is indifferent between leaving and staying in the union. In this region, the expected utility is above the blue line representing $U_i^{U,N+1}$. For high values of α, the black line lies below the blue line, as the expected utility considers the risk of other union members exiting. The higher a country values the public good, the more it loses from the exit of partner countries.

The dashed lines represent the utilities after introducing exit costs C. Exit costs shift the utility of exiting and change the expected utility of country i. The dashed black line represents the modified expected utility. The difference in expected utility results from the three effects discussed earlier. The *burden of commitment* is most visible in the region between $\underline{\alpha}^C$ and $\underline{\alpha}$, where the former is the exit threshold with exit costs C. The *deadweight loss of commitment*, on the other hand, is most visible to the left of $\underline{\alpha}^C$, where exits are very likely. Finally, the *value of committing others* is responsible for shifting the expected utility of country i under the exit costs above the solid line for higher values of α. In this part of the state space, country i benefits from its partners being less likely to leave more than it loses from potentially facing the exit costs.

Figure 2 provides the first indication that a member country's gain from higher exit costs C increases with its preference for the public good. This feature of the model is captured by Lemma 4. To formalize the intuition, however, I need to limit the dimensions of heterogeneity by assuming that countries face the same probability of being hit by a shock, $\forall_{i \in \mathcal{U}} p_i^{\varepsilon} = p^{\varepsilon}$, and the shocks are drawn from the same distribution, $\forall_{i \in \mathcal{U}} \bar{\varepsilon}_i = \bar{\varepsilon}$.

Lemma 4 *In an international union $\mathcal{U}$, if the preferences of all union members satisfy $0 < \alpha_i \leq (1 + \beta N)\alpha_m$, then the gain (loss) from introducing positive exit costs increases (decreases) with the preference for public goods, i.e.,*

$$\forall_{i,j \in \mathcal{U}} \forall_{C_j > 0} \ \alpha_i > \alpha_j \Rightarrow \mathbb{E}\left[U_i | C = C_j\right] - \mathbb{E}\left[U_i | C = 0\right] \geq \mathbb{E}\left[U_j | C = C_j\right] - \mathbb{E}\left[U_j | C = 0\right].$$

Proof In the appendix.

Lemma 4 states that countries with a higher preference for public goods will benefit more from introducing exit costs.[17] The result is intuitive, as a country with a higher α is less likely to consider an exit but suffers more from the exit of its partners.[18] This lemma is key

[17]The inequality in Lemma 4 is weak to account for the case of no union member at risk of exit. Whenever at least one country is at risk of exit, the inequality becomes strong.

[18]This statement is true for values of $\alpha \leq (1 + \beta N)\alpha_m$. Above this value, a country might want to exit, as

for evaluating the political economy problem of choosing exit costs in an international union. There are two direct implications from Lemma 4 that are key for studying the choice of exit costs in an international union. First, any C that benefits the median country also benefits a majority of union members. Second, if the introduction of exit costs benefits the country with the lowest preference for public goods, then it benefits all countries in the union. The following proposition builds on those results.

Proposition 5 *In an international union $\mathcal{U}$, in which the preferences of all union members satisfy $0 < \alpha_i \leq (1+\beta N)\alpha_m$, positive deadweight exit costs are introduced whenever*

$$n_m^0 \left[U_m^{U,N+1}(\alpha_m,\alpha_m) - U_m^{U,N}(\alpha_m,\alpha_m)\right] \geq \frac{\max\{\underline{\alpha} - \alpha_m + \bar{\varepsilon}, 0\}}{\underline{\alpha}} \left[g_m^{N+1} - g(\underline{\alpha})\right], \tag{17}$$

where n_m^0 is the number of union members (excluding the median country) who are at risk of exiting under $C = 0$, i.e. union members whose preference parameters satisfy $\alpha_i \leq \underline{\alpha} + \bar{\varepsilon}$, and $g(\underline{\alpha})$ is the public good spending level chosen by an autonomous country with the preference parameter $\underline{\alpha}$.

Moreover, there exists a level of exit costs C^*, such that introducing exit costs C^* Pareto dominates no exit costs whenever

$$n_j^0 \left[U_j^{U,N+1}(\alpha_j,\alpha_m) - U_j^{U,N}(\alpha_j,\alpha_m)\right] \geq \frac{\max\{\underline{\alpha} - \alpha_j + \bar{\varepsilon}, 0\}}{\underline{\alpha}} \left[g_m^{N+1} - g(\underline{\alpha})\right], \tag{18}$$

where j is the union member with the ex-ante lowest α, and n_j^0 is the number of union members (excluding country j) who are at risk of exiting under $C = 0$, i.e. union members whose preference parameters satisfy $\alpha_i \leq \underline{\alpha} + \bar{\varepsilon}$.

Proof In the appendix.

Proposition 5 specifies a condition for deadweight exit costs to be introduced in an international union, captured by equation (17). It concentrates on the trade-off faced by the median country. On the one hand, the median country faces the risk of partner countries exiting the union. This potential loss manifests itself in the product of the number of partner countries at risk of exiting and the utility loss the median country experiences after a partner's exit.

the spillovers are not enough to compensate for the too low level of g in the union. In this section, I abstract from those cases, as I view them uninteresting.

On the other hand, the median country may also want to exit. The max operator captures two cases. The median country might either be at risk of exit, when $\underline{\alpha} > \alpha_m - \bar{\varepsilon}$, or it might never want to exit when α_m is high enough. In the latter case, the median country faces no trade-off in introducing exit costs; hence, it introduces the maximal level, such as to eliminate exits. Even if α_m is low enough for the median country to face exit risk, positive exit costs might still be beneficial for the median country. For this, the risk of exit multiplied by the fiscal cost of union membership needs to be lower than the expected losses from losing partners.

Proposition 5 also addresses the issue of Pareto-efficiency of the deadweight exit costs. Pareto-efficiency is achieved when all members are at least as well off under the exit costs C^*, as they are under no costs. Relying on Lemma 4, I can restrict attention to the country with the lowest preference for public goods and explore when the country benefits from introducing positive exit costs. Condition (18) is very similar to (17), but it is written for of country j. Again, the max operator captures two cases. Whenever $\alpha_j - \bar{\varepsilon} \geq \underline{\alpha}$, country j faces no exit risk and the problem becomes trivial. As no country in the union is at risk of exit, the introduction of exit costs does not affect the welfare of any country. The more interesting case is country j facing exit risk. Then, country j might still benefit from positive exit costs as long as there are sufficiently many countries in the union that are at risk of exiting.

4 Post-exit relationship

So far, I have assumed that any union member must choose between remaining in the union or exiting it and severing all economic ties. However, given the experience of the UK leaving the European Union, it is clear that there are intermediate possibilities. A country exiting an international union can still negotiate a close relationship with the union.

In this section, I consider a post-exit relationship between country j and the union. In particular, I propose a class of relationships, which I call γ-deals. In a deal of this class, the country agrees to extend the spillovers from its public good to the union and partially comply with the union's policy. In return, it gains partial access to the spillovers of the union's public good.

Let us denote the utility of country j from engaging in a γ-deal after exit as U_j^γ. Then,

$$U_j^\gamma \left(\tilde{\alpha}_j, \alpha_m\right) \equiv y - g_j^\gamma + \tilde{\alpha}_j H \left(g_j^\gamma + \gamma\beta N g_m^N\right), \tag{19}$$

where $g_j^\gamma \equiv g_j + \gamma \left(g_m^N - g_j \right)$ is the public good spending of country j under the γ-deal.

Similarly, the utility union member i enjoys when country j engages in a γ-deal can be represented by[19]

$$U_i^{U,\gamma}(\alpha_i, \alpha_m) \equiv y - g_m^N + \alpha_i H \left([1 + \beta(N-1)] g_m^N + \gamma \beta g_j^\gamma \right). \tag{20}$$

Equations (19) and (20) give an overview of the welfare consequences for both sides of the agreement. In the current setup, a deal is always beneficial for the union. The following proposition captures the scope for signing a γ-deal.

Proposition 6 *In an international union $\mathcal{U}$, in which the preferences of all the union members satisfy $0 < \alpha_i \leq (1 + \beta N)\alpha_m$, if there exists a country j, which wants to exit the union, then there exists a $\gamma > 0$, such that a post-exit γ-deal as defined by equations (19) and (20) is a Pareto-improvement over a no-deal exit.*

Proof In the appendix.

Proposition 6 shows that both the exiting country and the union can be better off by negotiating a deal that benefits both sides. The γ-deal proposed in this section turns out to be always better than a no-deal relationship. The particular value of the parameter α_j determines the range of parameters γ that are acceptable for country j.

Figure 3 provides further intuition for the utility of the exiting country under a γ-deal. It presents the utility of an exiting union member after signing a γ-deal. As the blue curve in the figure shows, the utility of signing a deal is hump-shaped in γ, with $\bar{\gamma}_j$ being the value at which the country is indifferent between no deal and the $\bar{\gamma}_j$-deal.[20]

Despite the model's simplicity, the γ-deal bears some resemblance to the deals negotiated between the UK and the EU. A proposed post-Brexit relationship popularly known as a soft Brexit is similar to γ-deals with a high value of γ, where the exit does not substantially alter the relationship, except for giving the UK some flexibility in choosing its policies. The hard Brexit scenarios are more in line with a γ-deal with low values of γ - the UK gains much freedom in setting its policy but enjoys very little of the benefits of union membership.

[19] I assume that countries in the union enjoy the spillovers from country j to the degree γ. This symmetry is irrelevant to the results of this section, as they hold for any positive access to the spillovers.

[20] Formally, the proof of Proposition 6 establishes that country j is better off under a deal with γ incrementally larger than zero. It is not difficult to show that the first derivative of U_j^γ is decreasing in γ and positive at zero. Since at $\gamma = 1$ the utility is lower than at U_j^A, there exists a point $\bar{\gamma}_j > 0$ such $U_j^{\bar{\gamma}_j} = U_j^A$, i.e. country j is indifferent between exiting without a deal and accepting the γ-deal. For all values of γ between zero and $\bar{\gamma}_j$, country j is better off signing a γ-deal.

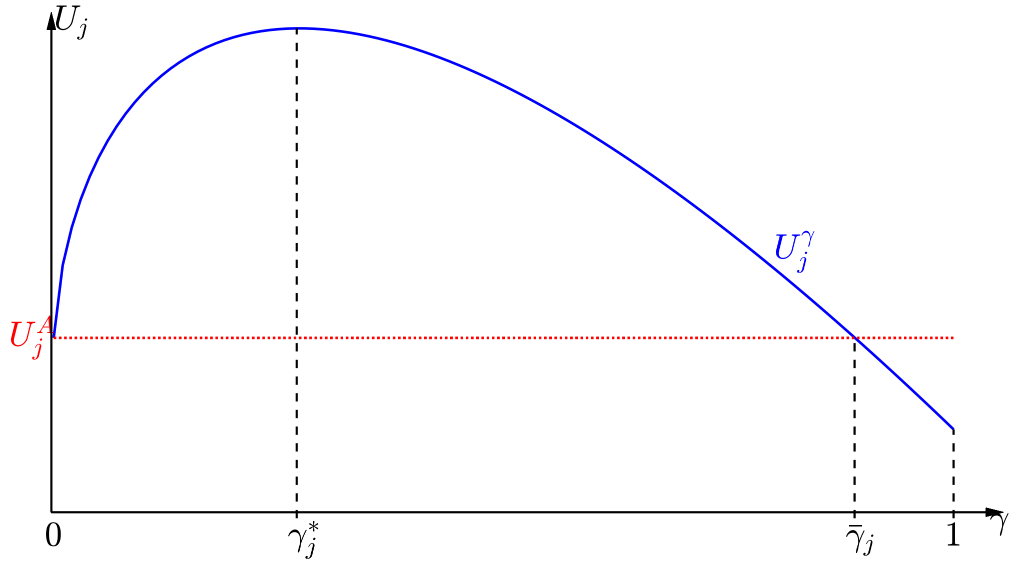


Figure 3: The utility of country j under a γ-deal as a function of γ.

4.1 The power of precedents

From the analysis so far, it seems that the union prefers γ to be as high as possible. There also exists a level of γ_j^* that maximizes the utility of country j. Therefore, any level of γ in the range $[\gamma_j^*; \bar{\gamma}_j]$ is Pareto-efficient. In this subsection, I introduce an extension of the model that renders this Pareto-frontier unachievable.

Consider an institutional setup where any γ-deal signed with country j is immediately available to all other union members. This setup could be realistic if the original deal creates a legal or political precedent. In such a case, the union faces the threat of (immediate) further exits if γ_j is such that

$$\exists_{i \in \mathcal{U}} U_i^{\gamma_j} \left(\alpha_i, \alpha_m \right) > U_i^{U,N} \left(\alpha_i, \alpha_m \right).$$

Any country i satisfying the above condition prefers to exit and sign a γ_j-deal rather than remaining in the union. Let us define $\mathcal{A}^{\gamma_j}$ to be the set of union members, other than j, who prefer the γ_j-deal to full membership. Then, no γ_j will be accepted by the union if

$$\gamma_j g_j^{\gamma_j} \leq \sum_{i \in \mathcal{A}^{\gamma_j}} \left(g_m^N - \gamma_j g_i^{\gamma_j} \right). \tag{21}$$

The union faces the risk that a subset $\mathcal{A}^{\gamma_j}$ of its members reduces their full membership to a γ_j-deal. Therefore, it weights the benefit of having a better relationship with the exiting country j against the losses from further exits.

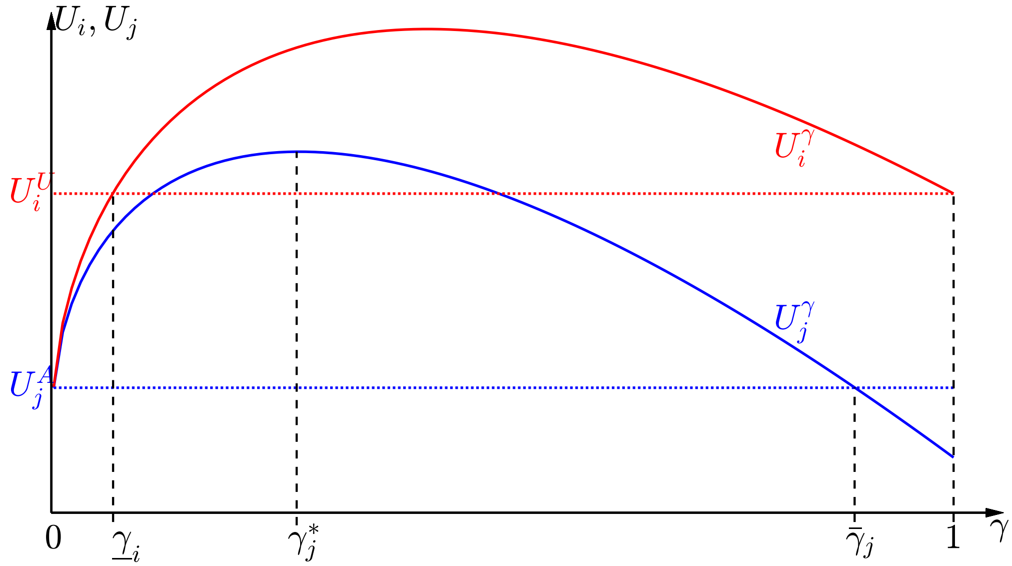


Figure 4: The utility of country j and country i under a γ-deal as a function of γ. The set of available deals might be now limited by the risk of country i exiting. Country i prefers the γ_j-deal over full membership whenever γ_j exceeds $\underline{\gamma}_i$.

Figure 4 presents a case where the preferences of country i substantially limit the range of available γ-deals in an institutional setup that allows precedent. The union only accepts deals with $\gamma_j < \underline{\gamma}_i$. In the presented case, this is lower than the union's preferred deal under no-precedent ($\bar{\gamma}_j$) and even lower than country j's preferred deal γ_j^*.

This extension sheds light on the Brexit negotiations. A soft Brexit would benefit both sides of the agreement, limiting the economic damage from the UK's departure. Nevertheless, the EU and the UK ended up signing a limited trade agreement. Part of the reason might be the strict requirements dictated by the EU for a more favorable deal. The lack of concessions on the union's side might be motivated by the threat of future exits if Brexit turns out to be successful.

5 Concluding remarks

The analysis in this paper shows that in an international union, individually-made exit decisions can be inefficient. Fiscal transfers between union members can eliminate this inefficiency but might be politically difficult. Policy adjustments catering to the preferences of the exiting country can also prevent an exit, but they are not necessarily Pareto-efficient. Similarly, the ex-ante introduction of exit costs may reduce the inefficiency of exits but is unlikely to incentivize first-best exit decisions. When failing to prevent an exit, the union can negotiate a post-exit relationship with the exiting country that benefits both sides while being in between membership and a no-deal exit. The scope for such an agreement is substantially reduced if the union has to consider the risk of other countries using the deal as a precedent.

Throughout the paper, I relate those results to real-world examples of interactions between international organizations and their members. The relationship between the UK and the EU, which had its most recent chapter in the British exit from the EU, is the single most inspiring example. Over the nearly four decades of EU membership, the UK was able to obtain fiscal transfers in the form of the UK rebate, heavily affect EU policymaking, and finally exited the union in 2020. The various economic analyses of potential post-Brexit EU-UK relationships clearly show the scope for mutual benefits of close cooperation. Nevertheless, the difficulty of the Brexit negotiations and the shortcomings of the final agreement might imply that the EU is aware of the risk of Brexit serving as a precedent for other members.

I keep the model simple for the tractability of the analysis and to make it more general. It can be extended to better represent a particular type of international union or to include more interesting economic dynamics. In particular, the framework could be adjusted to represent environmental treaties, which are a crucial element of the international political landscape and feature more intertemporal trade-offs.

A Appendix: Proofs

A.1 Proof of Proposition 2

The proof consists of two parts:

i. showing that if an exit is socially inefficient, then there exists a set of transfers satisfying the conditions in the proposition,

ii. showing that if such a set of transfers exists, then the exit by country j is socially inefficient.

Let us start with part i., for which I construct such a set. Consider a division of the union $\mathcal{U}$ in the following way:[21]

$$\begin{aligned}\mathcal{U}^- &= \left\{ i : i \in \mathcal{U}, i \neq j, U_i^{U,N+1}(\alpha_i, \alpha_m) - U_i^{U,N}(\alpha_i, \alpha_m) < 0 \right\}, \\ \mathcal{U}^+ &= \left\{ i : i \in \mathcal{U}, i \neq j, U_i^{U,N+1}(\alpha_i, \alpha_m) - U_i^{U,N}(\alpha_i, \alpha_m) \geq 0 \right\}.\end{aligned}$$

I can now construct a set of transfers

$$\tau_{i,j} = \begin{cases} U_i^{U,N+1}(\alpha_i, \alpha_m) - U_i^{U,N}(\alpha_i, \alpha_m) & \text{if } i \in \mathcal{U}^- \\ \chi \left[U_i^{U,N+1}(\alpha_i, \alpha_m) - U_i^{U,N}(\alpha_i, \alpha_m) \right] & \text{if } i \in \mathcal{U}^+ \end{cases}, \tag{22}$$

where χ is the ratio of net gains from the j-exit of those who gain from it, and the net gains from preventing the exit of those who lose from an exit.

$$\chi \equiv \frac{U_j^A(\tilde{\alpha}_j) - U_j^{U,N+1}(\tilde{\alpha}_j, \alpha_m) + \sum_{i \neq j, i \in \mathcal{U}^-} \left[U_i^{U,N}(\alpha_i, \alpha_m) - U_i^{U,N+1}(\alpha_i, \alpha_m) \right]}{\sum_{i \neq j, i \in \mathcal{U}^+} \left[U_i^{U,N+1}(\alpha_i, \alpha_m) - U_i^{U,N}(\alpha_i, \alpha_m) \right]}. \tag{23}$$

The inefficiency of the exit implies that condition (7) is violated, i.e.

$$U_j^{U,N+1}(\tilde{\alpha}_j, \alpha_m) - U_j^A(\tilde{\alpha}_j) + \sum_{i \neq j, i \in \mathcal{U}} \left[U_i^{U,N+1}(\alpha_i, \alpha_m) - U_i^{U,N}(\alpha_i, \alpha_m) \right] \geq 0,$$

which in turn guarantees that $\chi \leq 1$. The construction of the $\mathcal{U}^-$ and $\mathcal{U}^+$ sets guarantees that both the numerator and the denominator in equation (23) are positive.

Summing up, the total transfers make country j indifferent between union-membership and autonomy, the negative transfers from the $\mathcal{U}^-$ members compensate them for the potential gains they would have experienced after j-exit, and countries in the subset $\mathcal{U}^+$ are at least as good off under the transfer scheme as they would after exit. The latter are better off whenever $\chi < 1$.

The proof of part ii. is simpler. Let us denote the set of such transfers as $\{\tilde{\tau}_{i,j}\}_{i \neq j, i \in \mathcal{U}}$. Then,

$$U_j^{U,N+1}(\tilde{\alpha}_j, \alpha_m) + \sum_{i \neq j, i \in \mathcal{U}} \tilde{\tau}_{i,j} \geq U_j^A(\tilde{\alpha}_j) \tag{24}$$

$$\forall_{i \neq j, i \in \mathcal{U}} U_i^{U,N+1}(\alpha_i, \alpha_m) - \tilde{\tau}_{i,j} \geq U_i^{U,N}(\alpha_i, \alpha_m). \tag{25}$$

[21] In fact, it is easy to show that the set $\mathcal{U}^-$ is empty under the dense core assumption. I keep the general form of the proof to show that Proposition 2 is robust to relaxing the assumption.

Summing the second inequality over all i's and bringing it together with the first inequality yields

$$U_j^{U,N+1}(\alpha_i,\alpha_m)+\sum_{i\neq j,i\in\mathcal{U}}U_i^{U,N+1}(\alpha_i,\alpha_m)\geq U_j^A(\tilde{\alpha}_j)+\sum_{i\neq j,i\in\mathcal{U}}U_i^{U,N}(\alpha_i,\alpha_m)\,, \tag{26}$$

which violates condition (7), implying that the exit of country j is socially inefficient.

A.2 Lemma 7 and proof of Proposition 3

To facilitate the proof, let me first introduce Lemma 7, which establishes the pivotal member in the union.

Lemma 7 *For any adjusted public spending level $\tilde{g}_{m,j}$, that prevents country j from exiting the union, it holds that if the adjustment is preferred over j-exit by the median country, then it is also preferred by the majority of union members.*

Proof The median member prefers the adjusted policy level over an exit of country j, hence

$$\begin{aligned}
U_m(\tilde{g}_{m,j},N+1;\alpha_m)-U_m^{U,N}(\alpha_m,\alpha_m) &\geq 0,\\
[y-\tilde{g}_{m,j}+\alpha_m H((1+\beta N)\tilde{g}_{m,j})]-\left[y-g_m^N+\alpha_m H\left((1+\beta(N-1))g_m^N\right)\right] &\geq 0,\\
-\tilde{g}_{m,j}+g_m^N+\alpha_m\left[H((1+\beta N)\tilde{g}_{m,j})-H\left((1+\beta(N-1))g_m^N\right)\right] &\geq 0.
\end{aligned}$$

Let us consider two cases

1. Case 1: $\tilde{g}_{m,j}\geq\frac{1+\beta(N-1)}{1+\beta N}g_m^N$.

   Then, for any $\alpha_i\geq\alpha_m$

   $$\begin{aligned}
   \alpha_i\left[H((1+\beta N)\tilde{g}_{m,j})-H\left((1+\beta(N-1))g_m^N\right)\right] &\geq \alpha_m\left[H((1+\beta N)\tilde{g}_{m,j})+\right.\\
   &\qquad \left.-H\left((1+\beta(N-1))g_m^N\right)\right],
   \end{aligned}$$

   as all α_i's are positive. Hence,

   $$\begin{aligned}
   - &\quad \tilde{g}_{m,j}+g_m^N+\alpha_i\left[H((1+\beta N)\tilde{g}_{m,j})-H\left((1+\beta(N-1))g_m^N\right)\right]\geq\\
   \geq &\quad -\tilde{g}_{m,j}+g_m^N+\alpha_m\left[H((1+\beta N)\tilde{g}_{m,j})-H\left((1+\beta(N-1))g_m^N\right)\right].
   \end{aligned}$$

The above inequality states that the gain from preventing the exit by adjusting policy is larger for i than it is for m, or, formally

$$U_i\left(\tilde{g}_{m,j}, N+1; \alpha_i\right) - U_i^{U,N}\left(\alpha_i, \alpha_m\right) \geq U_m\left(\tilde{g}_{m,j}, N+1; \alpha_m\right) - U_m^{U,N}\left(\alpha_m, \alpha_m\right) \geq 0. \quad (27)$$

This means that all countries i, for whom $\alpha_i \geq \alpha_m$ support the policy adjustment, which (together with the median country), constitutes a majority of union members.

2. Case 2: $\tilde{g}_{m,j} < \frac{1+\beta(N-1)}{1+\beta N} g_m^N$.

   Then, for any $\alpha_i \leq \alpha_m$

$$\begin{aligned} \alpha_i\left[H\left((1+\beta N)\tilde{g}_{m,j}\right) - H\left((1+\beta(N-1))g_m^N\right)\right] \;\; \geq \;\; & \alpha_m\left[H\left((1+\beta N)\tilde{g}_{m,j}\right) + \right. \\ & \left. -H\left((1+\beta(N-1))g_m^N\right)\right], \end{aligned}$$

   as all α_i's are positive. Hence,

$$\begin{aligned} - \;\; & \tilde{g}_{m,j} + g_m^N + \alpha_i\left[H\left((1+\beta N)\tilde{g}_{m,j}\right) - H\left((1+\beta(N-1))g_m^N\right)\right] \geq \\ \geq \;\; & -\tilde{g}_{m,j} + g_m^N + \alpha_m\left[H\left((1+\beta N)\tilde{g}_{m,j}\right) - H\left((1+\beta(N-1))g_m^N\right)\right]. \end{aligned}$$

   The above inequality states that the gain from preventing the exit by adjusting policy is larger for i than it is for m, or, formally

$$U_i\left(\tilde{g}_{m,j}, N+1; \alpha_i\right) - U_i^{U,N}\left(\alpha_i, \alpha_m\right) \geq U_m\left(\tilde{g}_{m,j}, N+1; \alpha_m\right) - U_m^{U,N}\left(\alpha_m, \alpha_m\right) \geq 0. \quad (28)$$

   This means that all countries i, for whom $\alpha_i \leq \alpha_m$ support the policy adjustment, which (together with the median country), constitutes a majority of union members.

Depending on the type of policy adjustment - whether it is increasing total public spending or decreasing it - the deal is supported by a majority consisting of the median voter and either members with a high preference for the public good, or members with a low preference for the public good.[22] ∎

Lemma 7 allows us to limit attention to the median preferences in the union. The result does not directly follow from the single-peakedness of utility in g. The decision over the

[22] Depending on the size of the spillovers and the distribution of preferences, the policy adjustment could be unanimously preferred over the exit of country j.

policy adjustment involves the level of public spending and the number of countries in the union.

Proof of Proposition 3

The policy change described in the proposition needs to:

- be supported by a majority of the members,
- prevent the exit of country j,
- be acceptable for all members, i.e. do not trigger other exits.

Condition i. guarantees that the policy adjustment is sufficient to keep country j in the union. By Lemma 7, condition ii. ensures that most members support the policy adjustment (if it prevents country j from exiting). Finally, condition iii. deals with the unintended consequences of a policy adjustment. For the former two conditions to be correct, we need there to be $N+1$ members, i.e., none of the other countries decides to exit after a change in the policy g. Condition iii. guarantees that none of the members prefers to leave the union rather than staying in it under $\tilde{g}_{m,j}$. $\blacksquare$

A.3 Proof of Lemma 4

Proof Equation (16) gives the difference in expected utility conditional on $C = C_j$ and $C = 0$ as

$$\begin{aligned}
&\mathbb{E}\left[U_j|C=C_j\right] - \mathbb{E}\left[U_j|C=0\right] = \\
&\quad = p_j^{\varepsilon}\left[\int_{-\infty}^{\infty} \max\left\{U_j^A(\tilde{\alpha}_j) - C_j, U_j^{U,N+1}\left(\tilde{\alpha}_j, \alpha_m\right)\right\} f_j\left(\tilde{\alpha}_j\right) d\tilde{\alpha}_j\right] \\
&\quad - p_j^{\varepsilon}\left[\int_{-\infty}^{\infty} \max\left\{U_j^A(\tilde{\alpha}_j), U_j^{U,N+1}\left(\tilde{\alpha}_j, \alpha_m\right)\right\} f_j\left(\tilde{\alpha}_j\right) d\tilde{\alpha}_j\right] \\
&\quad + \sum_{k\neq j, k\in\mathcal{U}} p_k^{\varepsilon}\left[F_k\left(\underline{\alpha}\right) - F_k\left(\underline{\alpha}^{C_j}\right)\right]\left[U_j^{U,N+1}\left(\alpha_j, \alpha_m\right) - U_j^{U,N}\left(\alpha_j, \alpha_m\right)\right].
\end{aligned}$$

The two integrals have the same utility functions in them, so what they really differ with is the exit costs, which shifts the point at which the maximization problem flips in favor of union membership. In particular, in the range $\left[\underline{\alpha}^{C_j}, \underline{\alpha}\right]$ country j chooses to exit under no costs, and continues membership under the exit costs C_j. This allows me to re-write the full expression as

$$
\begin{aligned}
&\mathbb{E}\left[U_j|C=C_j\right]-\mathbb{E}\left[U_j|C=0\right]= \\
&\quad =p_j^{\varepsilon}\left[\int_{\underline{\alpha}^{C_j}}^{\underline{\alpha}}\left[U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right)-U_j^{A}\left(\tilde{\alpha}_j\right)\right]f_j\left(\tilde{\alpha}_j\right)d\tilde{\alpha}_j\right]-p_j^{\varepsilon}F_j\left(\underline{\alpha}^{C_j}\right)C_j \\
&\quad +\left[U_j^{U,N+1}\left(\alpha_j,\alpha_m\right)-U_j^{U,N}\left(\alpha_j,\alpha_m\right)\right]\sum_{k\neq j,k\in\mathcal{U}}p_k^{\varepsilon}\left[F_k\left(\underline{\alpha}\right)-F_k\left(\underline{\alpha}^{C_j}\right)\right]. \qquad (29)
\end{aligned}
$$

The preference shock ε_j is drawn from a uniform distribution $U\left[-\bar{\varepsilon},\bar{\varepsilon}\right]$, which means that $\alpha_i+\varepsilon_i$ follows also a uniform distribution $U\left[\alpha_i-\bar{\varepsilon},\alpha_i+\bar{\varepsilon}\right]$. Hence

$$
\begin{aligned}
f_i\left(\alpha\right) &= \begin{cases} 0 & \text{if } \alpha<\alpha_i-\bar{\varepsilon} \\ \frac{1}{2\bar{\varepsilon}} & \text{if } \alpha_i-\bar{\varepsilon}\leq\alpha\leq\alpha_i+\bar{\varepsilon} \\ 0 & \text{if } \alpha>\alpha_i+\bar{\varepsilon}\end{cases}, \\
F_i\left(\alpha\right) &= \begin{cases} 0 & \text{if } \alpha<\alpha_i-\bar{\varepsilon} \\ \frac{\alpha-\alpha_i+\bar{\varepsilon}}{2\bar{\varepsilon}} & \text{if } \alpha_i-\bar{\varepsilon}\leq\alpha\leq\alpha_i+\bar{\varepsilon} \\ 1 & \text{if } \alpha>\alpha_i+\bar{\varepsilon}\end{cases}.
\end{aligned}
$$

From the above properties of the distribution it follows that

$$\forall_{i,j\in\mathcal{U}}\ \alpha_i>\alpha_j\Rightarrow F_i(\alpha)\leq F_j(\alpha), \qquad (30)$$

and

$$\forall_{i,j\in\mathcal{U}}\ \alpha_i>\alpha_j\Rightarrow\forall_{\alpha\leq\alpha_j+\bar{\varepsilon}}\ f_i(\alpha)\leq f_j(\alpha). \qquad (31)$$

I also need to show that

$$\forall_{i,j\in\mathcal{U}}\alpha_i>\alpha_j\Rightarrow U_i^{U,N+1}\left(\alpha_i,\alpha_m\right)-U_i^{U,N}\left(\alpha_i,\alpha_m\right)>U_j^{U,N+1}\left(\alpha_j,\alpha_m\right)-U_j^{U,N}\left(\alpha_j,\alpha_m\right). \qquad (32)$$

The above inequality boils down to

$$-\left(g_m^{N+1}-g_m^{N}\right)+\alpha_i\left[H\left(G_m^{N+1}\right)-H\left(G_m^{N}\right)\right]>-\left(g_m^{N+1}-g_m^{N}\right)+\alpha_j\left[H\left(G_m^{N+1}\right)-H\left(G_m^{N}\right)\right],$$

where $G_m^N\equiv\left(1+\beta(N-1)\right)g_m^N$ and $G_m^{N+1}\equiv\left(1+\beta N\right)g_m^{N+1}$. This inequality requires $G_m^{N+1}>G_m^N$, which can be seen from comparing the condition for the public spending level selected by the median country (5) for a union with N and $N+1$ countries

$$H_g\left(G_m^N\right) = \frac{\alpha_m}{1+\beta(N-1)} > \frac{\alpha_m}{1+\beta N} = H_g\left(G_m^{N+1}\right),$$

and the fact that $H(\cdot)$ is increasing and concave, which imply that $H_g(\cdot)$ is decreasing.

All union members willingly participate in the union in period 1, which implies that $\forall_{i\in\mathcal{U}}\alpha_i \geq \underline{\alpha}^0$. Given that, I can use property (31) to conclude that for $\alpha_i > \alpha_j$:

$$\begin{aligned}&\int_{\underline{\alpha}^{C_j}}^{\underline{\alpha}} \left[U_i^{U,N+1}\left(\tilde{\alpha}_i,\alpha_m\right) - U_i^A\left(\tilde{\alpha}_i\right)\right] f_i\left(\tilde{\alpha}_i\right) d\tilde{\alpha}_i \geq \\ &\geq \int_{\underline{\alpha}^{C_j}}^{\underline{\alpha}} \left[U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right) - U_j^A\left(\tilde{\alpha}_j\right)\right] f_j\left(\tilde{\alpha}_j\right) d\tilde{\alpha}_j, \end{aligned} \tag{33}$$

where it is key to keep in mind that in the range $\left[\underline{\alpha}^{C_j}, \underline{\alpha}\right]$ the utility of staying in the union is smaller than the utility of being autonomous, so the utility difference inside the integral is negative.

Similarly, from equation (30) it follows that

$$-F_i\left(\underline{\alpha}^{C_j}\right) C_j \geq -F_j\left(\underline{\alpha}^{C_j}\right) C_j. \tag{34}$$

The final part of the proof requires a comparison of the *value of committing others* between different union members. A simple application of property (32) yields[23]

$$\begin{aligned}&\sum_{k\neq i,j;\ k\in\mathcal{U}} p_k^{\varepsilon}\left[F_k\left(\underline{\alpha}\right) - F_k\left(\underline{\alpha}^{C_j}\right)\right]\left[U_i^{U,N+1}\left(\alpha_i,\alpha_m\right) - U_i^{U,N}\left(\alpha_i,\alpha_m\right)\right] \geq \\ &\geq \sum_{k\neq i,j;\ k\in\mathcal{U}} p_k^{\varepsilon}\left[F_k\left(\underline{\alpha}\right) - F_k\left(\underline{\alpha}^{C_j}\right)\right]\left[U_j^{U,N+1}\left(\alpha_j,\alpha_m\right) - U_j^{U,N}\left(\alpha_j,\alpha_m\right)\right]. \end{aligned} \tag{35}$$

This leaves the j and i terms of the sums on each side, respectively

$$\begin{aligned}&p_j^{\varepsilon}\left[F_j\left(\underline{\alpha}\right) - F_j\left(\underline{\alpha}^{C_j}\right)\right]\left[U_i^{U,N+1}\left(\alpha_i,\alpha_m\right) - U_i^{U,N}\left(\alpha_i,\alpha_m\right)\right] \geq \\ &\geq p_i^{\varepsilon}\left[F_i\left(\underline{\alpha}\right) - F_i\left(\underline{\alpha}^{C_j}\right)\right]\left[U_j^{U,N+1}\left(\alpha_i,\alpha_m\right) - U_j^{U,N}\left(\alpha_i,\alpha_m\right)\right], \end{aligned}$$

[23]The expression in equation (35) becomes an equality when $\sum_{k\in\mathcal{U}} F_k(\underline{\alpha}) = 0$, i.e., when no country is at risk of exit. In such a union, exit costs have no impact on the welfare of the members. In all other cases, the inequality is strict and it transforms the inequality in Lemma 4 into a strict one.

or equivalently

$$p_j^{\varepsilon}\int_{\underline{\alpha}^{C_j}}^{\alpha} f_j\left(\tilde{\alpha}_j\right)d\tilde{\alpha}_j\left[U_i^{U,N+1}\left(\alpha_i,\alpha_m\right)-U_i^{U,N}\left(\alpha_i,\alpha_m\right)\right]\geq$$
$$\geq p_i^{\varepsilon}\int_{\underline{\alpha}^{C_j}}^{\alpha} f_i\left(\tilde{\alpha}_i\right)d\tilde{\alpha}_i\left[U_j^{U,N+1}\left(\alpha_j,\alpha_m\right)-U_j^{U,N}\left(\alpha_j,\alpha_m\right)\right], \tag{36}$$

which follows from (32) and (31).

Taking together inequalities (33)-(36) yields

$$\begin{aligned}
&\mathbb{E}\left[U_i|C=C_j\right]-\mathbb{E}\left[U_i|C=0\right]\\
&\quad=p_i^{\varepsilon}\left[\int_{\underline{\alpha}^{C_j}}^{\alpha}\left[U_i^{U,N+1}\left(\tilde{\alpha}_i,\alpha_m\right)-U_i^{A}\left(\tilde{\alpha}_i\right)\right]f_i\left(\tilde{\alpha}_i\right)d\tilde{\alpha}_j-F_i\left(\underline{\alpha}^{C_j}\right)C_j\right]\\
&\quad+\sum_{k\neq i,\ k\in\mathcal{U}}p_k^{\varepsilon}\left[F_k\left(\underline{\alpha}^{0}\right)-F_k\left(\underline{\alpha}^{C_j}\right)\right]\left[U_i^{U,N+1}\left(\alpha_i,\alpha_m\right)-U_i^{U,N}\left(\alpha_i,\alpha_m\right)\right]\geq\\
&\quad\geq p_j^{\varepsilon}\left[\int_{\underline{\alpha}^{C_j}}^{\alpha}\left[U_j^{U,N+1}\left(\tilde{\alpha}_j,\alpha_m\right)-U_i^{A}\left(\tilde{\alpha}_j\right)\right]f_j\left(\tilde{\alpha}_j\right)d\tilde{\alpha}_j-F_j\left(\underline{\alpha}^{C_j}\right)C_j\right]\\
&\quad+\sum_{k\neq j,\ k\in\mathcal{U}}p_k^{\varepsilon}\left[F_k\left(\underline{\alpha}\right)-F_k\left(\underline{\alpha}^{C_j}\right)\right]\left[U_j^{U,N+1}\left(\alpha_j,\alpha_m\right)-U_j^{U,N}\left(\alpha_j,\alpha_m\right)\right]=\\
&\quad=\mathbb{E}\left[U_j|C=C_j\right]-\mathbb{E}\left[U_j|C=0\right]. \quad \blacksquare
\end{aligned}$$

A.4 Lemma 8 and proof of Proposition 5

To facilitate the proof of Proposition 5, let me first introduce Lemma 8.

Lemma 8

$$\forall_{i\in\mathcal{U}}\forall_{C>0}\mathbb{E}\left[U_i|C\right]>U_i^{A}\left(\alpha_i\right).$$

Proof The expected utility conditional on C can be rewritten as

$$\begin{aligned}
\mathbb{E}\left[U_i|C\right] &= p_i^{\varepsilon}\left[\int_{-\infty}^{\infty}\max\left\{U_i^{A}\left(\tilde{\alpha}_i\right)-C,U_i^{U,N+1}\left(\tilde{\alpha}_i,\alpha_m\right)\right\}f_i\left(\tilde{\alpha}_i\right)d\tilde{\alpha}_i\right]\\
&+\sum_{k\neq i,k\in\mathcal{U}}p_k^{\varepsilon}\left\{F_k\left(\underline{\alpha}^{C}\right)\max\left\{U_i^{U,N}\left(\alpha_i,\alpha_m\right),U_i^{A}\left(\alpha_i\right)\right\}+\left[1-F_k\left(\underline{\alpha}^{C}\right)\right]U_i^{U,N+1}\left(\alpha_i,\alpha_m\right)\right\}\\
&+\left(1-\sum_{k\in\mathcal{U}}p_k^{\varepsilon}\right)U_i^{U,N+1}\left(\alpha_i,\alpha_m\right)\geq\\
&\geq p_i^{\varepsilon}\left[\int_{-\infty}^{\infty}U_i^{U,N+1}\left(\tilde{\alpha}_i,\alpha_m\right)f_i\left(\tilde{\alpha}_i\right)d\tilde{\alpha}_i\right]
\end{aligned}$$

$$
\begin{aligned}
& + \sum_{k\neq i, k\in\mathcal{U}} p_k^{\varepsilon} \left\{ F_k\left(\underline{\alpha}^C\right) U_i^A\left(\alpha_i\right) + \left[1 - F_k\left(\underline{\alpha}^C\right)\right] U_i^{U,N+1}\left(\alpha_i, \alpha_m\right) \right\} \\
& + \left(1 - \sum_{k\in\mathcal{U}} p_k^{\varepsilon}\right) U_i^{U,N+1}\left(\alpha_i, \alpha_m\right) = \\
= \ & \left[1 - \sum_{k\neq i, k\in\mathcal{U}} p_k^{\varepsilon} F_k\left(\underline{\alpha}^C\right)\right] U_i^{U,N+1}\left(\alpha_i, \alpha_m\right) + \sum_{k\neq i, k\in\mathcal{U}} p_k^{\varepsilon} F_k\left(\underline{\alpha}^C\right) U_i^A\left(\alpha_i\right) > \\
> \ & U_i^A\left(\alpha_i\right) \quad \blacksquare
\end{aligned}
$$

Proof of Proposition 5

Lemma 4 establishes that any level of exit costs C preferred over no exit costs by the median country is also preferred by all countries with a preference parameter $\alpha \geq \alpha_m$. This means that any costs level chosen by the median country will have majority support in the union over no exit costs. Moreover, Lemma 8 shows that no potential union member will decide against joining the union because of the introduction of exit costs, no matter the level. The two results brought together imply that we can concentrate solely on the expected utility of the median country under alternative exit costs levels for discovering what level will be chosen by the union.

To simplify notation, let us define the gain of country i from participating in a union of $N+1$ members as $\mathcal{M}_i\left(\alpha\right)$,

$$
\begin{aligned}
\mathcal{M}_i\left(\alpha\right) & \equiv U_i^{U,N+1}\left(\alpha, \alpha_m\right) - U_i^A\left(\alpha\right) \\
& = -g_m^{N+1} + g_i^A + \alpha\left[H\left((1+\beta N) g_m^{N+1}\right) - H\left(g_i^A\right)\right]. \qquad (37)
\end{aligned}
$$

Then,

$$
\begin{aligned}
\frac{\partial \mathcal{M}_i\left(\alpha\right)}{\partial \alpha} & = \frac{\partial g_i^A}{\partial \alpha}\left[1 - \alpha H_g\left(g_i^A\right)\right] + H\left((1+\beta N) g_m^{N+1}\right) - H\left(g_i^A\right) \\
& = H\left((1+\beta N) g_m^{N+1}\right) - H\left(g_i^A\right) > 0 \qquad (38)
\end{aligned}
$$

The transition between the first and second line follows from equation (3). The inequality is based on the properties of the function $H(\cdot)$ and a comparison between equations (3) and (5). It holds for all $\alpha_i < (1+\beta N)\alpha_m$.

$\underline{\alpha}^C$ is a continuous function of C for non-prohibitive values of C.[24] Neither $F_i(\alpha)$, nor $f_i(\alpha)$ is continuous, as they are derived from a uniform distribution on a limited range.

[24] By non-prohibitive I mean values of C for which $\underline{\alpha}^C > 0$.

However, assuming that all union members are better off participating in the union, i.e. $\forall_{i\in\mathcal{U}}\alpha_i > \underline{\alpha}$, the two functions are locally continuous around $C = 0$. I can, therefore, take a first order derivative of the expected utility of country i with respect to C, as $C \to 0^+$.

Then, starting from equation (29),

$$\frac{\partial}{\partial C_0}\left[\mathbb{E}\left[U_i|C=C_0\right]-\mathbb{E}\left[U_i|C=0\right]\right] = -p_i^{\varepsilon}\left[\mathcal{M}_i\left(\underline{\alpha}^{C_0}\right)f_i\left(\underline{\alpha}^{C_0}\right)\frac{\partial\underline{\alpha}^{C_0}}{\partial C_0}\right] - p_i^{\varepsilon}F_i\left(\underline{\alpha}^{C_0}\right)$$
$$- p_i^{\varepsilon}f_i\left(\underline{\alpha}^{C_0}\right)C_0\frac{\partial\underline{\alpha}^{C_0}}{\partial C_0} - \frac{\partial\underline{\alpha}^{C_0}}{\partial C_0}\sum_{k\neq i,k\in\mathcal{U}} p_k^{\varepsilon}f_k\left(\underline{\alpha}^{C_0}\right)\left[U_i^{U,N+1}\left(\alpha_i,\alpha_m\right)-U_i^{U,N}\left(\alpha_i,\alpha_m\right)\right].$$

I take the derivative at $C_0 \to 0$, hence, $\underline{\alpha}^{C_0} \to \underline{\alpha}$. Moreover, the gain from being in a monetary union for a country with $\underline{\alpha}^{C_0}$ is negligible, i.e., $\mathcal{M}\left(\underline{\alpha}^{C_0}\right) = 0$. This simplifies the above equation to

$$\frac{\partial}{\partial C_0}\left[\mathbb{E}\left[U_i|C=C_0\right]-\mathbb{E}\left[U_i|C=0\right]\right] = -p_i^{\varepsilon}F_i\left(\underline{\alpha}\right)$$
$$- \frac{\partial\underline{\alpha}^{C_0}}{\partial C_0}\sum_{k\neq i,k\in\mathcal{U}} p_k^{\varepsilon}f_k\left(\underline{\alpha}\right)\left[U_i^{U,N+1}\left(\alpha_i,\alpha_m\right)-U_i^{U,N}\left(\alpha_i,\alpha_m\right)\right]. \tag{39}$$

To analyze equation (39) further, I can obtain $\frac{\partial\underline{\alpha}^{C_0}}{\partial C_0}$ using the implicit function theorem. Let me first define an auxiliary function S,

$$S \equiv \mathcal{M}\left(\underline{\alpha}^{C}\right) + C = 0. \tag{40}$$

The two relevant partial derivatives are

$$\begin{aligned}\frac{\partial S}{\partial\underline{\alpha}^{C}} &= \mathcal{M}'\left(\underline{\alpha}^{C}\right),\\ \frac{\partial S}{\partial C} &= 1,\end{aligned}$$

where the first derivative is also positive, as can be seen in equation (38). This brings us to

$$\frac{\partial\underline{\alpha}^{C}}{\partial C} = -\frac{1}{\mathcal{M}'\left(\underline{\alpha}^{C}\right)} < 0. \tag{41}$$

I can further derive $\mathcal{M}'(\underline{\alpha})$ as

$$\mathcal{M}'(\underline{\alpha}) = H\left((1+\beta N)g_m^{N+1}\right) - H\left(g\left(\underline{\alpha}\right)\right) = \frac{g_m^{N+1} - g\left(\underline{\alpha}\right)}{\underline{\alpha}}, \tag{42}$$

where $g(\underline{\alpha})$ is the spending level chosen by an autonomous country with public good preferences $\underline{\alpha}$. The second equality follows from the fact that

$$0 = \mathcal{M}(\underline{\alpha}) = -g_m^{N+1} + g(\underline{\alpha}) + \underline{\alpha}\left[H\left((1+\beta N)g_m^{N+1}\right) - H\left(g\left(\underline{\alpha}\right)\right)\right].$$

Combining the above findings with equation (39) allows me to state that $\frac{\partial}{\partial C_0}\left[\mathbb{E}\left[U_i|C=C_0\right] - \mathbb{E}\left[U_i|C=0\right]\right] \geq 0$ if and only if

$$\frac{\underline{\alpha}\left[U_i^{U,N+1}\left(\alpha_i,\alpha_m\right) - U_i^{U,N}\left(\alpha_i,\alpha_m\right)\right]}{g_m^{N+1} - g\left(\underline{\alpha}\right)} \sum_{k\neq i, k\in\mathcal{U}} p_k^{\varepsilon} f_k\left(\underline{\alpha}\right) \geq p_i^{\varepsilon} F_i\left(\underline{\alpha}\right).$$

Inserting the uniform distribution into the above condition yields

$$\frac{\underline{\alpha}\left[U_i^{U,N+1}\left(\alpha_i,\alpha_m\right) - U_i^{U,N}\left(\alpha_i,\alpha_m\right)\right]}{g_m^{N+1} - g\left(\underline{\alpha}\right)} \sum_{\substack{k\neq i, k\in\mathcal{U},\\ \alpha_k\leq\underline{\alpha}+\bar{\varepsilon}}} p_k^{\varepsilon} \geq p_i^{\varepsilon}\max\left\{\underline{\alpha} - \alpha_i + \bar{\varepsilon}, 0\right\}. \tag{43}$$

For positive deadweight exit costs to be chosen by the union, we need the median country to support positive costs over no costs. This is guaranteed by condition (43) applied to the median country, which is equation (17).

For all countries to be better off under some level of positive exit costs, it is sufficient to show that the country with the lowest parameter α in the union is better off and apply Lemma 8 to extend the argument to all other union members. A sufficient condition for the lowest α country to be better off under some positive level of exit costs C is condition (43) applied to that country, which becomes then equation (18). ∎

A.5 Proof of Proposition 6

The proof proceeds in two steps:

1. I show that for all $\gamma > 0$ all remaining union members are better off with a γ-deal as compared to a no-deal exit of country j,

2. I show that there exists a $\gamma > 0$ such that country j is better off with a γ-deal as compared to no relationship with the union.

The first part follows naturally from the fact that the benefits of the public good are non-competitive, i.e. the union members do not incur any cost of allowing country j to enjoy a fraction γ of the spillovers. This means that for any $\gamma > 0$ their utility under a γ-deal is strictly higher than under no-deal, as they gain partial access to the spillovers from country j without paying anything for it, i.e.

$$\forall_{i \in \mathcal{U}, i \neq j} \forall_{\gamma > 0} U_i^{U,\gamma} (\alpha_i, \alpha_m) > U_i^{U,N} (\alpha_i, \alpha_m) .$$

For the second part I first need to highlight that U_j^γ is continuous in γ and differentiable. Then,

$$\frac{\partial U_j^\gamma}{\partial \gamma} = - \left(g_m^N - g_j \right) + \alpha_j \left[\left(g_m^N - g_j \right) + \beta N g_m^N \right] H_g \left(g_j^\gamma + \gamma \beta N g_m^N \right) . \tag{44}$$

In the special case of $\gamma = 0$, the γ-deal is equivalent to no deal, and $U_j^\gamma|_{\gamma=0} = U_j^A$. Considering the first derivative at this special case

$$\frac{\partial U_j^\gamma}{\partial \gamma}|_{\gamma=0} = \beta N g_m^N > 0,$$

where the additional terms from equation (44) cancel out when substituting $\alpha_j H_g(g_j) = 1$ from condition (3).

The above result implies that country j is better off by a marginally positive γ as compared to $\gamma = 0$. ∎

B Appendix: Parameter values for numerical examples

For all numerical examples, I assume that the public good utility function $H(\cdot)$ is a CRRA function,

$$H(g) = \frac{g^{1-\theta}}{1-\theta},$$

where $\theta < 1$.

Table 1: Parameter values used in the numerical examples presented in the figures

Parameter	Description	Fig. 1a	Fig. 1b	Fig. 2	Fig. 3	Fig. 4
β	International spillovers	0.15	0.15	0.9	0.9	0.9
θ	Utility parameter	0.5	0.5	0.3	0.3	0.3
$N+1$	Number of countries in the union	7	7	4	11	11
α_j	Lowest preference in the union	0.25	0.25	0.18	0.19	0.19
α_i	Second lowest pref. in the union	-	-	0.25	-	0.23
α_m	Median country's preference	0.5	0.5	0.25	0.30	0.30
α_h	Highest pref. in the union	0.75	1.25	0.28	-	-
y	Income	5	5	5	5	5
p^{ε}	Prob. of preference shock	-	-	0.20	-	-
$\bar{\varepsilon}$	Max. size of the preference shock	-	-	0.15	-	-
C	Exit costs	-	-	0.10	-	-

References

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

Athanassiou, P. (2009). Withdrawal and expulsion from the EU and EMU. ECB Legal Working Paper Series 10, European Cental Bank.

Battaglini, M. and Harstad, B. (2016). Participation and duration of environmental agreements. *Journal of Political Economy*, 124(1):160–204.

Battaglini, M. and Harstad, B. (2020). The political economy of weak treaties. *Journal of Political Economy*, 128(2):544–590.

Benitez, J. (2019). U.s. nato policy in the age of trump: Controversy and consistency. *The Fletcher Forum of World Affairs*, 43(1):179–200.

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

Born, B., Müller, G. J., Schularick, M., and Sedláček, P. (2019). The Costs of Economic Nationalism: Evidence from the Brexit Experiment*. *The Economic Journal*, 129(623):2722–2744.

Broadbent, B., Pace, F. D., Drechsel, T., Harrison, R., and Tenreyro, S. (2019). The Brexit

Vote, Productivity Growth and Macroeconomic Adjustments in the United Kingdom. Discussion Papers 1916, Centre for Macroeconomics (CFM).

Eijffinger, S. C., Kobielarz, M. L., and Uras, B. R. (2018). Sovereign Default, Exit and Contagion in a Monetary Union. *Journal of International Economics*, 113:1 – 19.

Fuchs, W. and Lippi, F. (2006). Monetary union with voluntary participation. *The Review of Economic Studies*, 73(2):437–457.

Gancia, G., Ponzetto, G. A., and Ventura, J. (2020). A theory of economic unions. *Journal of Monetary Economics*, 109:107–127.

Gancia, G., Ponzetto, G. A., and Ventura, J. (2021). Globalization and Political Structure. NBER Working Papers 22046, National Bureau of Economic Research, Inc.

Gourinchas, P. O., Martin, P., and Messer, T. (2020). The Economics of Sovereign Debt, Bailouts and the Eurozone Crisis. NBER Working Papers 27403, National Bureau of Economic Research, Inc.

Grossman, G. M., McCalman, P., and Staiger, R. W. (2021). The "new" economics of trade agreements: From trade liberalization to regulatory convergence? *Econometrica*, 89(1):215–249.

Harstad, B. (2006). Flexible integration? mandatory and minimum participation rules. *The Scandinavian Journal of Economics*, 108(4):683–702.

Kobielarz, M. (2021). Exits and Bailouts in a Monetary Union. Discussion Paper Series DPS21.07, KU Leuven – Faculty of Economics and Business.

Maggi, G. and Ossa, R. (2021). The political economy of deep integration. *Annual Review of Economics*, 13(1):19–38.

Maggi, G. and Staiger, R. W. (2015). Optimal design of trade agreements in the presence of renegotiation. *American Economic Journal: Microeconomics*, 7(1):109–43.

McGrattan, E. R. and Waddle, A. (2020). The impact of brexit on foreign investment and production. *American Economic Journal: Macroeconomics*, 12(1):76–103.

Oates, W. (1972). *Fiscal federalism*. Harcourt Brace Jovanovich, New York, NY.