CESifo WORKING PAPERS

6941
2018

March 2018

International Agreements, Economic Sovereignty and Exit

Martin Richardson, Frank Stähler

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6941*
Category 8: Trade Policy

International Agreements, Economic Sovereignty and Exit

Abstract

We develop a model in which it is uncertainty about the future domestic policy environment that both makes international cooperation attractive and induces the possibility of a nation reneging on such an international agreement. We show, in a fairly general setting in which the likelihood of exit is affected by the degree of cooperation, that the possibility of exit reduces the optimal degree of initial cooperation. "Full" cooperation will never be optimal, and the optimal degree of cooperation will never be such as to "squeeze out" any possibility of exit. However, an increase in global uncertainty may imply an increase in cooperation when exit risks are already large to begin with.

JEL-codes: F020, F130, F510, F530.

Keywords: international agreements, international cooperation, exit, sovereignty.

Martin Richardson
Research School of Economics
College of Business and Economics
Australian National University
Australia – Canberra, ACT 0200
martin.richardson@anu.edu.au

Frank Stähler
University of Tübingen
Mohlstr. 36 (V4)
Germany – 72074 Tübingen
frank.staehler@uni-tuebingen.de

February 2017
This paper has been presented at several seminars and conferences, and we wish to thank the participants for very useful comments. Richardson was a visitor at the University of Tübingen when this paper was begun and thanks them for their hospitality.

1 Introduction

In the *Washington Post* of February 25 2015, discussing the proposed Trans-Pacific Partnership (TPP), Elizabeth Warren criticized the Investor-State Dispute Settlement (ISDS) provisions in the proposal and suggested, “This isn’t a partisan issue. Conservatives who believe in U.S. sovereignty should be outraged that ISDS would shift power from American courts, whose authority is derived from our Constitution, to unaccountable international tribunals” (E. Warren, 2015).[1] In June 2016, a referendum of British voters decided, by 51.9% to 48.9%, that the U.K. should exit the European Union: so-called Brexit. Again, “the argument that Britain has lost sovereignty, and even its democracy, by being in the European Union [was] at the heart of the case for Brexit.” (*The Economist*, 2016). Indeed, the primary group proposing exit, the Vote Leave group, had the slogan “Take Control” and the website “voteleavetakecontrol.org” argued that, “[w]e need to be able to hold our lawmakers to account. Over half our laws are made by unelected EU bureaucrats in Brussels for whom we never voted . . . We should take back the power to kick out the people who make our laws.” (Vote Leave, 2016).

Both of these examples illustrate an argument that was very common in the context of both opposition to the TPP and arguments for Brexit: that these international agreements involved a loss of ‘economic sovereignty’. The latter is usually understood to refer to the freedom to set policy of “territorial-jurisdictional entities with independent powers of making and administering” laws and economic policy[2] and two immediate and obvious objections to its use in this context are (1) that any significant international agreement must involve the loss of sovereignty in this respect and (2) that the decision to enter into any such agreement is itself a manifestation of sovereignty and a sovereign decision-maker always has the (sovereign) power to tear up any such agreement. On the first of these points, an agreement to act in a way in which one would choose to act sans the agreement serves little economic purpose.[3] In Dixit’s (1987) terms, in talking of commitments, threats

[1]She was quite right that this is a non-partisan issue. Donald Trump used the TPP extensively in his presidential campaign, arguing that, ”[i]t would give up all of our economic leverage to an international commission that would put the interests of foreign countries above our own” (Trump, 2016) and indicated the US’ withdrawal from it in January 2017 in one of the first executive acts of his presidency.

[2]This is adapted from Rodrik’s (2000) definition of the “nation-state”.

[3]Of course, such agreements can have symbolic value and serve other political ends, but these are not the sorts of agreements that concern us here.

and promises in international trade relations, "[i]f the action we commit ourselves to, or threaten, or promise were the optimal response to the other country's action after the fact, there would be no need for us to do anything in advance" (Dixit, 1987, p.269). Consequently, one wonders why the "loss" of sovereignty is considered a problem in some contexts, but not in others. The answer presumably lies in the perceived extent of the loss of sovereignty, weighed against whatever gain is felt to come, as a *quid pro quo*, from that loss.[4] The important point here is that sovereignty (and its sacrifice) is best thought of – and modelled – as a continuous rather than a discrete choice. The second point – that restricting future "sovereignty" is itself an act of sovereignty and, by the same token, can always be rescinded – is surely brought into high relief by the act of Brexit itself. It is an observation that suggests that the modelling of international agreements should consider the possibility of a party withdrawing from any such agreement.

In this paper, we seek to develop a model, at a fairly abstract level, of the choice made by a sovereign party to an international agreement of how much "sovereignty" to sacrifice in order to join the agreement. In light of the preceding discussion, we model sovereignty as a continuous variable and we also consider the possibility of exit: a party subsequently choosing to rescind its undertakings. In our model, exit occurs precisely because the leaving country perceives that it has sacrificed too much sovereignty. We do not wish to tie the analysis down to a particular kind of agreement – for example, a preferential trading arrangement, a monetary union, an environmental agreement or an agreement on mutual standards recognition – and we also wish to capture the idea of political influences on governments' decisions. So we consider a setup in which countries' policy choices are potentially made "jointly" with others and the extent to which they are coordinated is taken as a measure of the sacrifice of sovereignty. A government seeking to maintain "full" sovereignty will simply choose its policies to maximize its own perceived welfare but one that is willing to cede some sovereignty will take into account its effects

[4] While this argument relates, in a sense, to the depth of sovereignty loss in a particular agreement, the breadth of agreements also provides a dimension in which sovereignty is a continuous variable: "[m]any talk of being sovereign as if it were like being pregnant: one either is or is not. The truth is more complex . . . Britain has signed some 700 international treaties that impinge on sovereignty." (*The Economist*, 2016.) On the 'depth' of sovereignty issue, an opinion piece on ISDS and the TPP in the *Sydney Morning Herald* (Australia) in October 2015 was headlined, 'Trans-Pacific Partnership: we're selling economic sovereignty for little return" (P. Martin, 2015). The essence of the piece was that the TPP's proposed ISDS provisions represented a loss of economic sovereignty for Australia and that estimates of the gains to Australia from the TPP were insufficiently large to make that loss worthwhile.

on others (and will benefit from reciprocal consideration from these same others). The weight it places on the effects on others is our continuous measure of the extent to which it is ceding economic sovereignty in the agreement.

The value of retaining "sovereignty" is that it enables a government to act, in the future, in a way that is not constrained by prior agreements and undertakings. Presumably this is valuable in allowing flexibility in unforeseen future circumstances and suggests that uncertainty about future conditions is essential in modelling this issue. We consider a two-period, two-country model in which the second-period impact of policy on the governments' objectives is uncertain, *ex ante*. The countries can form an agreement regarding the degree to which they cooperate in policy setting but, as sovereign decision makers, each can renege if the realizations of the second-period state of the world make it optimal for them to do so. The agreement is signed in the anticipation of the possibility of exit, and we explore the effects of that anticipation on the initial degree of cooperation. While we set up the problem in an abstract context, we also adopt a particular model for analytical purposes.

There is an extensive literature on the question of sovereignty, particularly with respect to globalization.[5] For our purposes, three analyses are particularly relevant. Rodrik (2000) presented an important "trilemma", analogous to the famous open economy trilemma of international finance, in which he suggested that no more than two of the following three features of a globalized world can be mutually consistent: integrated national economies, the nation state and what he termed mass politics, meaning, "political systems where: a) the franchise is unrestricted; b) there is a high degree of political mobilization; and c) political institutions are responsive to mobilized groups" (Rodrik, 2000, p.180). The implication of this is that, if the nation state is to be operative (and it is here that sovereignty is vested) then either it cannot be responsive to political interest groups (particularly domestic ones) or globalization in the form of international integration is not achievable: free domestic policy choices in a sovereign nation state preclude the "sacrifice" of otherwise sovereign decisions to an international agreement. In our model this trilemma does not apply perfectly but trade-offs across these three desiderata do occur: greater

[5] It is too extensive to canvass here, but three examples from 2001 – Krasner, Quiggin and Philpott – illustrate well the political economy and philosophy discussions in this field. Furthermore, an older literature exists on the stability of international environmental agreements and sovereign debt; see Barrett (1994), Chander and Tulkens (1992), Folmer et al (1993) and Bulow and Rogoff (1989), Eaton and Gersovitz (1981), respectively.

responsiveness to domestic political pressures typically (but not uniformly) maps into a decreased degree of optimal international integration, for example.

Bagwell and Staiger (2017) explicitly model the trade-off between economic sovereignty – defined formally in their paper in terms of the "internal affairs" of a country – and international agreements and show that the latter need not impinge on the former, depending on the nature of the interdependencies between countries and the scope of negotiated policies (see their Proposition 3). As their focus is on examining a formal analytical meaning of "economic sovereignty", they do not consider the issue of exit from a pre-agreed arrangement.

A very relevant paper for our purposes is Maggi and Staiger (2015). This paper looks explicitly at trade agreements and, importantly, focuses on a world in which renegotiation is possible after the resolution of uncertainty. Their interests are rather different to ours, focusing on the nature of the optimal contract between cooperating nations and on the consequences for trade liberalization of permitting renegotiation when ex-post compensation between countries is inefficient. They also consider a binary choice for countries – protection or free trade – and they focus on a world in which the benefits of free trade are not *ex post* verifiable, so complete contingent contracts cannot be written. They identify the nature of the optimal contract in this setting (a "property rule" or a "liability rule") and establish that, if renegotiation occurs at all in equilibrium, it will be from protection to free trade: the importer agrees to liberalize trade and the exporter will offer a side payment for this.

While we borrow a number of features from Maggi and Staiger (2015), as described below, we seek a model in which exit can occur in equilibrium – that is, the agreement breaks down – and enquire as to the consequences of that possibility for the nature of the equilibrium *ex ante* agreement. One view of this is that international cooperation requires partner countries to set up a supranational institution that oversees the details of cooperation for at least some time before terms can be renegotiated. While this commitment makes cooperation credible, it will also to be shown to be the reason for the possibility of exit. We will scrutinize how the possibility of exit will affect the design of such an agreement, and we find that an optimal agreement will neither imply full cooperation nor avoid any exit risk completely.

The remainder of the paper is organized as follows. Section 2 presents the general model, and Section 3 shows that neither full cooperation nor compete exit avoidance are equilibrium strategies. Section 4 provides a more specific example of the model, illustrates the results by some numerical simulations and shows that an increase in global uncertainty may also imply more cooperation. Section 5 summarizes and concludes.

2 The general model

We start with a very general two-country setting of two periods in which Home and Foreign, with a common discount factor $\delta \in (0, 1)$, choose some costly policy variables a and a^*, respectively.[6] We will refer to these policy variables as scalars, but our analysis and all our results also hold true if a and a^* are vectors. In the first period, both countries know the potential welfare gains from policy cooperation, but these gains are uncertain for the second period.[7] In particular, welfare will be affected by a parameter θ in the domestic country and a parameter θ^* in the foreign country that measure the "intervention necessity" for the respective country. While the θ terms could represent many aspects of the uncertain policy environment, we interpret them as capturing the extent to which policy actions feed into the relevant government's objective function: the sensitivity of the government's maximand with respect to a particular policy action.[8] The uncertain nature of this captures ambiguity about the political environment in which the government operates and, while it can be interpreted directly as the efficacy of economic policy, our preferred interpretation is in terms of the political support a government receives contingent on policy actions. For example, an open immigration policy may play well (and thus be effective for the policy-maker) when political support flows from employers of an increased workforce, but may be less attractive if the electorate becomes concerned about increased "competition for jobs". Low tariffs may be politically attractive in a full employment environment but less so otherwise; and so on.

As the focus of our interest is not so much on the architecture of the cooperative agree-

[6]All variables and parameters referring to the foreign country are asterisked. We choose the two-country case as the possibility of exit has the strongest implication in this setup, that is, the breakdown of the agreement.

[7]We refer to "welfare" for convenience, but the objective function of the policy maker may capture elements of political support as well as – or instead of – concerns for economic welfare. This is one of a number of dimensions in which our analysis differs from Korinek (2016).

[8]Papers analyzing ISDS follow a similar approach, see for example Horn and Tangerås (2016), Janeba (2016), Kohler and Stähler (2016) and Konrad (2017).

ment, we assume that the θ realizations are public information when they are revealed. If the countries are in a cooperative arrangement, the overlying institution of cooperation observes both policy environments without error in our analysis, so we preclude issues of private information and all that that entails. Nevertheless, as in Maggi and Staiger (2015) these realizations are externally non-verifiable so the governments cannot write complete contingent contracts upon them.

In the first period, these realizations, denoted by θ_1 and θ_1^*, are known. The two governments agree on a degree of cooperation denoted by $\lambda \in [0, 1]$ and explained more fully below, and choose a_1 and a_1^* in light of these known realizations and in anticipation of period 2 realizations. In period 2, θ_2 and θ_2^* (no subscripts will be used henceforth where the meaning is clear from context) are realized and observed and each government can either opt out of the agreement on λ, in which case a_2 (a_2^*) is chosen to maximize W_2 (W_2^*), or continue to abide by the previous agreement. Of course, if one country exits and simply maximizes its own welfare, then the other country will do the same. As for the second period, both countries know only that both the domestic and foreign realizations are distributed between $\underline{\theta}$ and $\overline{\theta}$ and between $\underline{\theta}^*$ and $\overline{\theta}^*$, respectively, according to the joint probability distribution $f(\theta, \theta^*)$ where $0 < \underline{\theta} < \overline{\theta} < \infty$ and $0 < \underline{\theta}^* < \overline{\theta}^* < \infty$. These realizations affect domestic welfare in both periods. Consequently, the objective functions for the domestic and the foreign country are then given respectively by $W(a, a^*, \theta)$ and $W^*(a, a^*, \theta^*)$. Without cooperation, the domestic (foreign) country fully controls its activity levels a (a^*).

We suppose that, in any international agreement, the cooperation parameter λ is chosen to hold for both periods. The reason is that we are interested in fields of international cooperation that require a supranational institution to be set up to follow a certain policy rule for some time. If cooperation could be renegotiated period by period by agreeing on a and a^* and transfers in each period, then no coordination problem would arise in this full information environment. However, the establishment of potentially long-lasting cooperation requires institutional management and commitment, and this is the reason why our model assumes that λ (and transfers) cannot be subject to renegotiation, at least over two periods. In this context, we think of international cooperation as being enforced by a supranational institution following a certain policy rule. This institution serves as a commitment device, as its rules are agreed upon at the beginning of the cooperative agreement and cannot be adjusted after the realization of θ and θ^* in the second period.

Both countries agree that the institution should apply a policy rule that specifies a certain $\lambda \in [0, 1]$ that measures the degree of cooperation. The institution then takes this λ as given and determines both countries' activities to solve the following programs:

$$\max_{a} W(a, a^*, \theta) + \lambda W^*(a, a^*, \theta^*), \max_{a^*} W^*(a, a^*, \theta^*) + \lambda W(a, a^*, \theta). \tag{1}$$

If both countries agree on $\lambda = 1$ we will refer to this case as complete cooperation, if $\lambda < 1$ we refer to this case as incomplete cooperation. If there is no cooperation at all then $\lambda = 0$.

We think of λ as our (inverse) measure of sovereignty and are then interested in what determines its choice. The obvious objective of the parties is to maximize the sum of W and W^*, but it is quite possible that one component – W, for example – is lower when both a and a^* are chosen optimally to maximize $W + W^*$ for $\lambda = 1$ than when a is chosen solely to maximize W (equivalent to $\lambda = 0$) even though the former must represent a potential Pareto improvement on the latter. Consequently, an agreement is reached in period 1 based on expectations of θ and θ^* for period 2, but if realizations of these are sufficiently different to expectations that one country prefers to leave the agreement then the agreement breaks down in period 2. The broad structure of the game is outlined in Table 1.

Table 1: Game structure

Period 1: Both countries learn realizations θ_1 and θ_1^* and agree on λ and a transfer payment if $\lambda \neq 0$. If $\lambda \neq 0$, the policy rule (1) is applied to a_1 and a_1^*.
Period 2: Both countries learn the realizations of θ_2 and θ_2^*. Each country decides on exit or not: effectively $\lambda = 0$ if either or both exit. If no country exits, the policy rule (1) is applied again to a_2 and a_2^*.

We solve this game backwards and demonstrate that λ plays a crucial role in the potential exit decision of a country. For example, if the second period realizations are such that the domestic country realizes a low θ while the foreign country realizes a large θ^*, the domestic country may be better off by exiting the agreement: its own interests are best served by a low policy choice (a) but its partner would prefer a high policy intervention, and if the degree of cooperation is high, then the home country would be obliged, in the agreement, to give considerable weight to its partner's interests; that is, to sacrifice considerable 'sovereignty'. This possibility will be anticipated by both countries when designing the agreement in the first place. Note again that we allow transfers between the two countries when designing the agreement. Hence, both countries always have an interest in maximizing the sum of discounted welfare gains over the two periods, as any asymmetry in individual expected welfare realizations can be managed by transfers.

Assuming interior solutions, the policy rule (1) implies, in each period, the first-order conditions

$$W_a(\cdot) + \lambda W_a^*(\cdot) = W_{a^*}^*(\cdot) + \lambda W_{a^*}(\cdot) = 0,$$

where we assume that both $W_a(\cdot)+\lambda W_a^*(\cdot)$ and $W_{a^*}^*(\cdot)+\lambda W_{a^*}(\cdot)$ are strictly concave w.r.t. a and a^* for all $\lambda \in [0,1]$, so that we can determine optimal (and unique) policy activity levels $a(\lambda,\theta,\theta^*)$ and $a^*(\lambda,\theta,\theta^*)$. If, in any period, the countries were to not cooperate then each would simply maximize their own perceived welfare: a program equivalent to that above but with $\lambda = 0$. Consequently, we define the domestic gain from cooperation in a period as

$$\Delta(\theta,\theta^*,\lambda) = W(a(\lambda,\theta,\theta^*), a^*(\lambda,\theta,\theta^*),\theta) - W(a(0,\theta,\theta^*), a^*(0,\theta,\theta^*),\theta).$$

The foreign gain from cooperation $\Delta^*(\theta,\theta^*,\lambda)$ is defined analogously. We make the following assumptions regarding *total* gains from cooperation:

$$\forall\lambda \in [0,1[: \Delta_\lambda(\cdot) + \Delta_\lambda^*(\cdot) > 0; \Delta_\lambda(\theta,\theta^*,1) + \Delta_\lambda^*(\theta,\theta^*,1) = 0. \tag{2}$$

Condition (2) states that *aggregate* gains are always increasing in cooperation and are maximized at $\lambda = 1$; otherwise, complete cooperation is not a socially desirable objective even without an exit option. Condition (2) does not imply, of course, that this is true for each country's welfare individually. We also assume:

$$\forall\lambda \in]0,1] : \Delta_\theta(\cdot) > 0, \Delta_{\theta^*}(\cdot) < 0, \Delta_{\theta^*}^*(\cdot) > 0, \Delta_\theta^*(\cdot) < 0. \tag{3}$$

That is, when cooperating, an increase in a country's own sensitivity of welfare to policy will make that country better off but an increase in the other country's policy sensitivity will make a country worse off. The former property reflects the idea that a country can react optimally to a change in its domestic circumstances – if policy is more effective then that is typically welfare-improving directly but will also induce an increased policy intervention. The latter property, however, reflects the impact of the international agreement: when $\lambda > 0$, the home country must change domestic policy, however slightly, to respond to a change in its partner's policy sensitivity, even though such a policy change is not directly beneficial at home.

Finally, we make the following assumptions concerning the effects of increased cooperation on the marginal welfare impact of changes in policy sensitivity:

$$\forall \lambda \in]0,1] : \Delta_{\theta\lambda}(\cdot) > 0, \Delta_{\theta^*\lambda}(\cdot) < 0, \Delta^*_{\theta^*\lambda}(\cdot) > 0, \Delta^*_{\theta\lambda}(\cdot) < 0. \tag{4}$$

That is, the marginal welfare effect of an increase in a country's policy sensitivity (which is positive, by (3)) is increasing – but that of an increase in the other country's policy sensitivity (which is negative) is decreasing – in the degree of cooperation. Intuitively, the higher is λ, the more a policy-maker must act with a view to its partner's welfare, so an increase in θ, for example, induces both home and foreign countries to act more in home's interest, and this takes the home country closer to maximizing its welfare gains. The foreign country, however, experiences a decline in marginal welfare gains, taking it further away from its maximum welfare gains.

3 The potential exit threat

Consider now the threat of exit (that is, reneging on the agreement) by the home country. This will occur if and only if $\Delta(\cdot) < 0$, and the probability of that depends on the realization not only of θ but of θ^*, too. Let $\theta^{*\prime}(\lambda)$ denote the value of θ^* such that the domestic gain from cooperation is exactly zero at the lowest possible realization of θ.[9] That is,

$$\theta^{*\prime}(\lambda) : \Delta\left(\underline{\theta}, \theta^{*\prime}(\lambda), \lambda\right) = 0.$$

If there exist possible realizations of θ^* in excess of this critical value, at full cooperation ($\lambda = 1$), then a potential exit threat for the home country exists; that is, home exit could

[9]Given the assumed properties from (3), any higher realization of θ must render $\Delta > 0$, meaning no domestic exit will occur.

occur (depending on realizations of θ^*) if and only if $\theta^{*\prime}(1) < \overline{\theta}^*$.

We can also define a critical value for domestic θ realizations such that the home country is indifferent between exiting the agreement or staying in it:

$$\tilde{\theta}(\theta^*, \lambda) : \Delta\left(\tilde{\theta}(\theta^*, \lambda), \theta^*, \lambda\right) = 0. \tag{5}$$

Eq. (5) defines an implicit function that gives all combinations of θ and θ^* for which the domestic welfare gain from the agreement is exactly zero. In order to derive our main results, we first investigate how the critical values $\theta^{*\prime}(\lambda)$ and $\tilde{\theta}(\theta^*, \lambda)$ behave. We find:

Lemma 1. *If*

$$\Delta_\lambda(\tilde{\theta}(\theta^*, \lambda), \theta^*, \lambda) < 0, \tag{6}$$

then

$$\theta^{*\prime}_\lambda(\lambda) = -\frac{\Delta_\lambda(\cdot)}{\Delta_\theta(\cdot)} < 0$$

and $\tilde{\theta}(\theta^*, \lambda)$ *is a monotonically increasing function:*

$$\tilde{\theta}_{\theta^*}(\theta^*, \lambda) = -\frac{\Delta_{\theta^*}(\cdot)}{\Delta_\theta(\cdot)} > 0, \tilde{\theta}_\lambda(\theta^*, \lambda) = -\frac{\Delta_\lambda(\cdot)}{\Delta_\theta(\cdot)} > 0.$$

Proof. Condition (6) together with assumption (4) implies that $\Delta_\lambda(\cdot) < 0$ in the range in which domestic exit will occur.

□

Lemma 1 shows that the behavior of $\theta^{*\prime}(\lambda)$ and $\tilde{\theta}(\theta^*, \lambda)$ depends on the sign of $\Delta_\lambda(\tilde{\theta}(\theta^*, \lambda), \theta^*, \lambda)$. A natural assumption is that this is negative: if the domestic country just 'breaks even' and, consequently, its actions are such that they support the foreign country to a large extent while the foreign country does little in return, we expect that an increase in the degree of cooperation will make the domestic country worse off. In this case, $\theta^{*\prime}(\lambda)$ decreases with the degree of cooperation, making exit more likely, and $\tilde{\theta}(\theta^*, \lambda)$ increases with the degree of cooperation, demanding a larger θ to keep the domestic country indifferent between exit and and no exit. Furthermore, Lemma 1 shows that making the foreign influence on policies stronger – an increase in θ^* – must be matched by an increase in θ (a countervailing effect for the domestic country) for a given λ in order to keep Δ at its zero level.

So home exit will occur only for realizations of θ^* greater than $\theta^{*\prime}(\lambda)$ and realizations of θ less than $\tilde{\theta}(\theta^*, \lambda)$. Consequently, the expected aggregate losses arising from domestic exit, denoted by $\Psi(\lambda)$, are given by

$$\Psi(\lambda) = \int_{\theta^{*\prime}(\lambda)}^{\overline{\theta}^*} \int_{\underline{\theta}}^{\tilde{\theta}(\theta^*)} [\Delta(\theta, \theta^*, \lambda) + \Delta^*(\theta, \theta^*, \lambda)] \, f(\theta, \theta^*) d\theta d\theta^*.$$

If domestic exit occurs, aggregate welfare gains from the agreement fall to zero. These welfare losses due to domestic exit can be avoided by choosing a sufficiently low degree of cooperation, denoted $\tilde{\lambda}$, such that $\tilde{\theta}(\overline{\theta}^*, \tilde{\lambda}) = \underline{\theta}$ and $\theta^{*\prime}(\tilde{\lambda}) = \overline{\theta}^*$, making $\Psi(\lambda) = 0$. Lemma 2 shows that the expected aggregate welfare loss due to domestic exit increases with the degree of cooperation, away from $\tilde{\lambda}$.

Lemma 2. *If* $\theta^{*\prime}(\lambda) < \overline{\theta}^*$ *and condition* (6) *holds, then* $\Psi_\lambda(\lambda) > 0$. $\Psi_\lambda(\tilde{\lambda}) = 0$.

Proof. See Appendix A.1. □

Definitionally, while a very ambitious agreement may lead to substantial welfare gains if exit does not occur, it also increases the expected aggregate welfare losses due to potential exit. Similar analysis and expressions define the expected aggregate losses associated with the possibility of foreign exit, $\Psi^*(\cdot)$. The sum of discounted aggregate welfare gains from cooperation, then, can be written as follows:

$$\begin{aligned} \widehat{\Omega}(\lambda) &= \Delta(\theta_1, \theta_1^*, \lambda) + \Delta^*(\theta_1, \theta_1^*, \lambda) \\ &+ \delta \left(\int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}^*}^{\overline{\theta}^*} (\Delta(\theta, \theta^*, \lambda) + \Delta^*(\theta, \theta^*, \lambda)) \, f(\theta, \theta^*) d\theta d\theta^* - \Psi(\lambda) - \Psi^*(\lambda) \right). \end{aligned}$$

The two terms on the first line here give the welfare gains as they are realized in the first period when both θ_1 and θ_1^* are known and the supranational institution is set up to run cooperative policies using policy rule (1). The second line gives the discounted welfare gain from this policy, adjusted for the expected aggregate welfare losses that can arise from both domestic and foreign exit. When designing the agreement – that is, specifying λ – both governments anticipate the risk of exit and maximize $\widehat{\Omega}(\lambda)$ w.r.t. λ. We assume that $\widehat{\Omega}(\lambda)$ is continuous and that its derivative exists and is continuous. A direct implication is that (i) by the Extreme Value Theorem a maximum exists, and (ii) we can use first derivatives to state our two central results. We first scrutinize whether there is a chance of complete cooperation.

Proposition 1. *If condition* (6) *holds and* $\theta^{*\prime}(1) < \overline{\theta}^*$ *and/or* $\theta'(1) < \overline{\theta}$*, then cooperation will never be complete.*

Proof. From (2),

$$\widehat{\Omega}_\lambda(\lambda = 1) = -\Psi_\lambda(1) - \Psi^*_\lambda(1) < 0$$

if $\theta^{*\prime}(1) < \overline{\theta}^*$ and/or $\theta'(1) < \overline{\theta}$. □

Proposition 1 shows – under very general conditions – that cooperation will never be complete in the face of a serious exit threat because the sum of discounted aggregate welfare gains can be increased by reducing λ from $\lambda = 1$. If the exit threat is real for $\lambda = 1$, both governments are willing to sacrifice part of the welfare gains as to reduce the exit risk. The second central result scrutinizes whether they will want to go so far as to completely avoid exit.

Proposition 2. *In any agreement, the possibility of exit will never be completely avoided.*

Proof. Without loss of generality, suppose exit by the home country is the binding constraint such that $\tilde{\lambda} < \tilde{\lambda}^*$, implying $\Psi_\lambda(\tilde{\lambda}) = 0$ (see Lemma 2) and $\Psi^*_\lambda(\tilde{\lambda}) = 0$ because the foreign exit constraint is not binding. As $\tilde{\lambda} < 1$, $\widehat{\Omega}_\lambda(\tilde{\lambda}) > 0$. □

Proposition 2 shows that the sum of discounted aggregate welfare gains can be increased by introducing an exit risk. Consequently, any agreement that is subject to an exit risk will neither feature complete cooperation nor will it want to eliminate the exit risk completely. Thus, potential exit is an equilibrium result of the optimal design of the agreement, and not necessarily due to some oversight on the part of the governments.[10] This is a noteworthy result as it depends neither on the specification of welfare gains nor on any details of the underlying joint distribution of θ and θ^*. In any agreement where full cooperation would lead to an exit risk that is exacerbated by an increase in λ, cooperation will neither be complete nor such that this risk will completely disappear.

[10]Note carefully that our results do not rely on $\widehat{\Omega}(\lambda)$ to be strictly quasiconcave. Both Proposition 1 and 2 show that neither full cooperation nor the complete elimination of any exit possibility maximize $\widehat{\Omega}(\lambda)$.

4 A specific model

As an illustration of this general approach and in order to explore the effect of an increase in uncertainty on cooperation, we now consider a more closely specified set-up. We suppose henceforth that there is an increasing quadratic cost to policy actions, that a government's objective function depends on a simple weighted sum of their own policy and the policy choice of the other government and that the variable θ interacts with this policy sum multiplicatively. Specifically:

$$\begin{aligned} W(a, a^*, \theta) &= (a + \mu a^*)\theta - \frac{a^2}{2\gamma} \\ W^*(a^*, a, \theta) &= (a^* + \mu^* a)\theta^* - \frac{a^{*2}}{2\gamma}, \end{aligned} \tag{7}$$

where $\mu(\mu^*)$ captures the sensitivity of the home (foreign) government's objectives to foreign (home) actions and may be either positive or negative,[11] and γ is an inverse measure of the costliness of policy. From (7), we have $W_a = \theta - (a/\gamma), W_{aa} = -1/\gamma < 0, W_{aa^*} = 0, W_{a^*} = \mu\theta$ and $W_{a^*a^*} = 0$. The supranational institution, given some λ, observing some $\Theta = (\theta, \theta^*)$ and for a cooperative Home, solves $\max_a(W + \lambda W^*)$ and does a similar exercise for a cooperative Foreign, leading to

$$a_c = \gamma(\theta + \lambda\mu^*\theta^*), a_c^* = \gamma(\theta^* + \lambda\mu\theta), \tag{8}$$

where the subscript c denotes cooperation. Plugging back into W yields

$$V_c(\lambda, \theta, \theta^*) \equiv W(a_c, a_c^*, \theta) = \frac{1}{2}\gamma\left[(\theta + 2\mu\theta^*)\theta + \lambda(2(\mu\theta)^2) - \lambda(\mu^*\theta^{*2})\right].$$

Suppose, in period 2, the countries have previously agreed to cooperate and chosen some non-zero level of λ. New realizations of Θ are then observed and continued cooperation means the decision problem just described applies, yielding welfare in period 2, if the agreement holds, of $V_c(.)$ evaluated at the realizations of the random variables. Reneging on the agreement (exiting) means choosing a to maximize just W, which is the same as above but with $\lambda = 0$ leading to activity levels $a_n = \gamma\theta$ and $a_n^* = \gamma\theta^*$ and thus $V_n(\theta, \theta^*) =$

[11] We assume, however, that $\text{sign}(\mu) = \text{sign}(\mu^*)$ throughout this section.

$\frac{1}{2}\gamma\theta(\theta+2\mu\theta^*)$, where the subscript n denotes non-cooperation. But $V_n(\theta,\theta^*) = V_c(0,\theta,\theta^*)$ is just the first term in $V_c(\lambda,\theta,\theta^*)$, and thus is less or greater than $V_c(.)$, depending on the second term. So we can write $V_c(.)$ as:

$$V_c = V_n + \Delta \text{ where } \Delta = \frac{1}{2}\gamma\lambda\left[\left(2(\mu\theta)^2 - \lambda(\mu^*\theta^*)^2\right)\right].$$

Note that Δ here corresponds exactly to Δ in our earlier notation, and it is easily verified that the assumptions in (3), (4) and (6) all hold here. For non-cooperation to be strictly preferred to cooperation in period 2 we need $\Delta < 0$, that is, $2(\mu\theta)^2 < \lambda(\mu^*\theta^*)^2$. Consequently, Home will renege on the deal if realizations are such that $\theta < \kappa\theta^*$ where

$$\kappa = \kappa(\lambda,\mu,\mu^*) = M\sqrt{\lambda/2} \text{ for } M = \mu^*/\mu$$

Similarly, Foreign will renege if $\theta^* < \kappa^*\theta$ where $\kappa^* = \kappa^*(\lambda,\mu,\mu^*) = 1/M\sqrt{\lambda/2} = \kappa/M^2$. A useful way to visualize this is to think of it in the entire $\mathbb{R}^2_+$ space that contains the supports of θ and θ^*. We can show the ranges of Θ realizations that will lead one country or the other to renege and can then consider different distribution supports as rectangles in this space, indicating – for given λ and thus κ – the kinds of outcomes we might see. Figure 1 exemplifies this, and shows that the foreign country will renege if Θ is such that it is below the respective $\kappa(\cdot)$–line in the rectangle. In the same setting we can illustrate the effects of changing λ where an increase in λ increases κ and so makes the two boundary lines scissor together, as in Figure 1.[12]

We assume henceforth that, in period 1, the observed realizations of the θs when the agreement is negotiated are θ_1 and θ_1^*, and that these are the expected values of θ and θ^* across the distributions in period 2.[13] Further mathematical details are relegated to Appendix A.2, and in what follows, we want to explore the effect of changes in the dispersion of Θ on welfare, the degree of cooperation and the exit probability. To do so, we need to consider specific distributions, and for our analysis of uncertainty and cooperation we assume that θ and $\theta^*$are uniformly and independently distributed on supports $[\underline{\theta},\overline{\theta}]$ and

[12]Note that the slope of the "H reneges" boundary, $1/\kappa$, always exceeds that of the "F reneges" boundary, κ/M^2, for all possible values of M and λ: $1/\kappa = \kappa/\kappa^2 = \kappa/((\lambda M^2)^2) = (2/\lambda)(\kappa/M^2) > \kappa/M^2$. So we can never observe a realization of Θ such that both parties would wish to renege on any agreement.

[13]Nothing significant hinges on this assumption – essentially it means that, looking forward, the governments expect the same conditions that hold at the time of negotiation to also prevail in future – and it can be relaxed at the cost of some messier algebra.

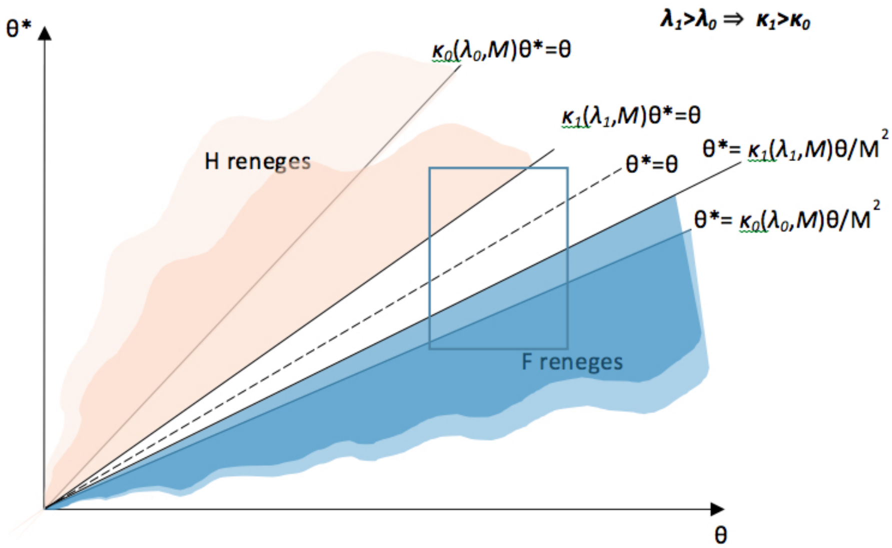

Figure 1: Distribution support and changes in λ

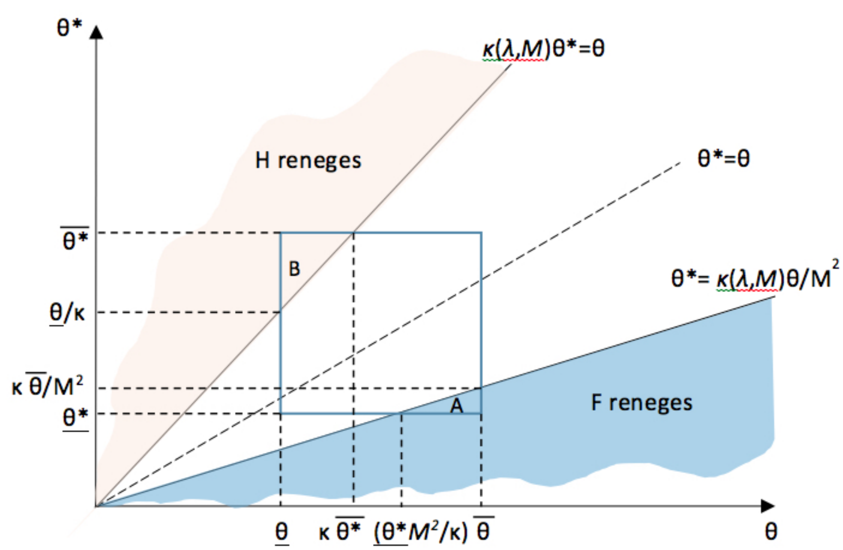

Figure 2: Outcomes for a given λ

$[\underline{\theta}^*, \overline{\theta}^*]$, respectively.[14] More specifically, we assume that the width of each support is 2Γ and $2\Gamma^*$, respectively, so, with means θ_1 and θ_1^*, we have $\overline{\theta}^* = \theta_1^* + \Gamma^*$, $\underline{\theta}^* = \theta_1^* - \Gamma^*$, $\overline{\theta} = \theta_1 + \Gamma$ and $\underline{\theta} = \theta_1 - \Gamma$. Figure 2 is a generic representation of our analysis. In that Figure, triangle A represents the realizations of Θ such that the foreign country will renege on the agreement and B represents realizations of Θ such that Home will renege.

The uniform distribution allows us to investigate an increase in uncertainty by a mean-preserving spread (MPS), that is, an increase in Γ and Γ^*. We start with a case in which both countries are fully symmetric with identical policy sensitivity (so $\mu = \mu^*$ and $M = 1$) and identical supports for the θ's and in which we investigate an MPS in both distributions simultaneously. That is, in what follows we have $\Gamma = \Gamma^*$ throughout.[15] Exit by one country or the other becomes now more likely as cooperation increases (λ rises) and as the dispersion of the distributions increases.[16] Absent any possibility of exit, it is easily shown that aggregate welfare is strictly increasing in λ (so the optimal degree of cooperation is always complete cooperation at $\lambda = 1$) and increasing (and convex) in the Γ terms. Allowing the possibility of exit, however, when the distributions are fairly compact (Γ and Γ^* are low) then nearly full cooperation is optimal but, when the distributions spread and exit becomes more likely, the equilibrium gains from cooperation fall.

What does this increase in dispersion do to welfare? Let the welfare potential of an agreement be defined as the ratio of the sum of discounted maximized welfare gains when exit can occur to the equivalent sum absent the possibility of exit. Figure 3 shows the maximized welfare potential, the optimal degree of cooperation (λ) and the aggregate probability of some exit plotted against the dispersion in the distribution of the θ's (Γ and Γ^*). As to be expected, the maximized welfare potential is falling in Γ. However, the optimal degree of cooperation is non-monotonic in this dispersion: a small MPS in the θ's induces lower optimal cooperation initially but, as the θ's become more dispersed, a further MPS induces a greater λ. It is also clear from Figure 3 that exit and welfare potential are less sensitive to λ the larger is Γ. This is what drives the eventual increase in the optimal

[14] A more comprehensive discussion of these and other simulation results is presented in a Simulations Appendix, available on request from the authors.

[15] In what follows we have set $\theta_1 = \theta_1^* = 1, \mu = \mu^* = 1$ and $\delta = 0.99$. We then vary Γ (and Γ^*) from 0.01 to 0.99 and λ from 0.05 to 1. Furthermore, all references to welfare henceforth should be understood as references to expected welfare.

[16] In terms of Figure 2, given the uniform distributions, this is simply the ratio of the areas $A + B$ to the entire area $(\overline{\theta} - \underline{\theta})(\overline{\theta}^* - \underline{\theta}^*)$.

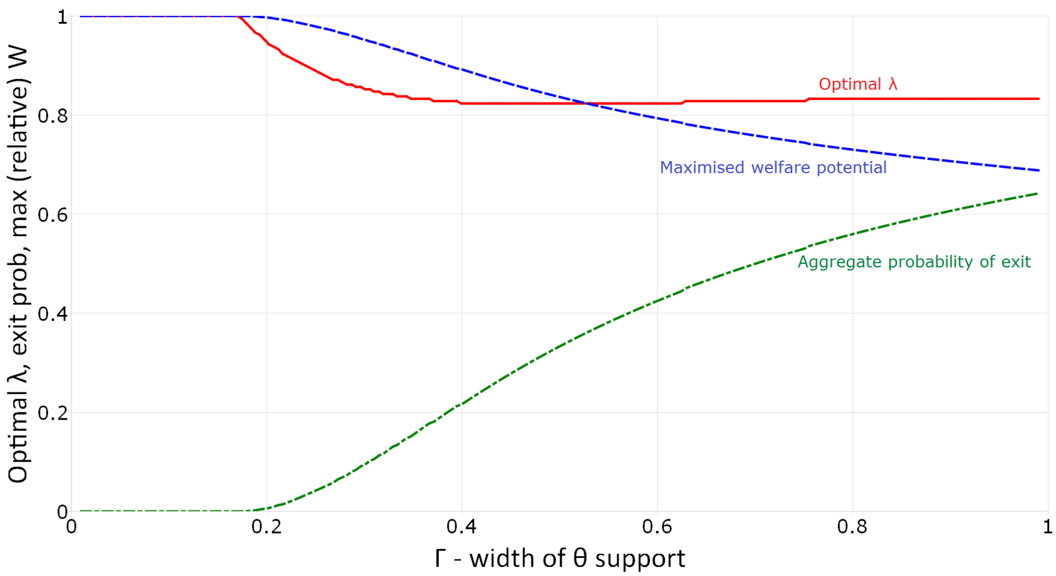

Figure 3: Optimal λ, welfare potential, probability of exit: the symmetric case

degree of cooperation as uncertainty rises: there are still gains from cooperation but the losses generated by the increased likelihood of exit are less severe as uncertainty becomes sufficiently high.

The exercise here is an MPS in both distributions: a way of modeling an increase in global uncertainty. The simulations suggest that, while this makes full cooperation less desirable, the relationship is not monotonic, and thus, an increase in global uncertainty may actually increase cooperation when global uncertainty is already on a level that implies a substantial exit risk. At high levels of uncertainty, further increases might induce more cooperation rather than less. Increasing uncertainty is always reducing the welfare potential of an agreement. It highlights the critical role of exit: the reason that the gains from cooperation fall when uncertainty increases is entirely due to the fact that the probability of the agreement breaking down goes up.

We now turn to a case of asymmetry in which the sensitivity of welfare to foreign policies – the μ parameter – varies between the two countries; we think of this as an approximation of a small and large country where the latter is less sensitive to the policies of the former than *vice versa*. The exercise here fixes $\mu > \mu^*$ (so we think of Home as 'small' relative to Foreign) and looks at the impact of mean-preserving spreads in the θ and θ^* distributions.

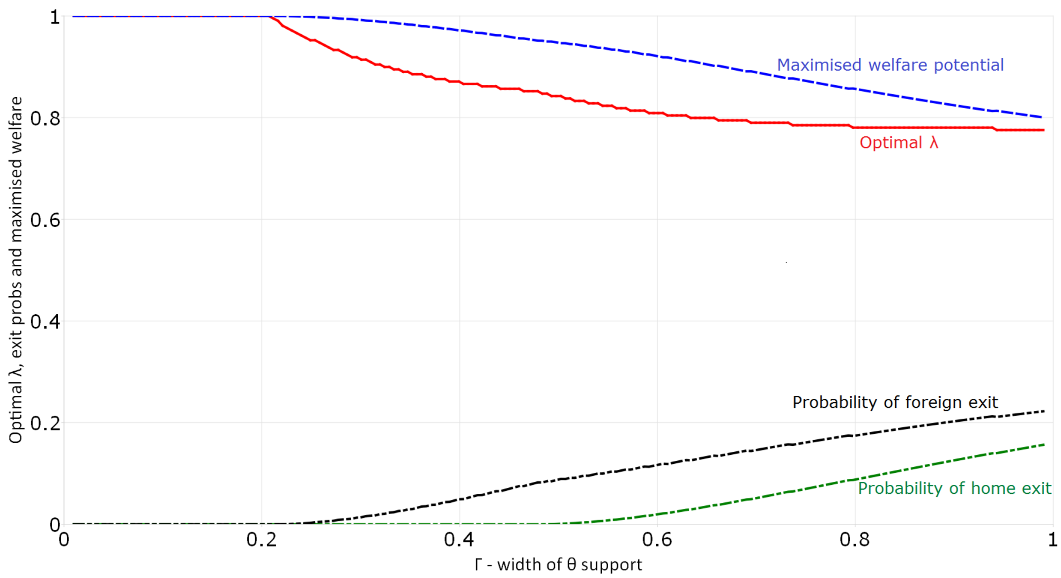

Figure 4: Optimal λ,welfare potential, probability of exit: the asymmetric case

This exercise is directly comparable with that shown in Figure 3 above, but with $\mu = 1.9$ and $\mu^* = 1$ (whereas $\mu = \mu^* = 1$ in Figure 3). As in Figure 3, we find here that a MPS in the θ and θ^* distributions reduces the welfare potential but here the consequences for exit are much less extreme. Interestingly, it is foreign exit that is always more likely than domestic exit – the country that is less sensitive to foreign policies is more likely to exit as the uncertainty increases. But the optimal degree of cooperation now is monotonically decreasing in the MPS, as shown in Figure 4. In contrast to the symmetric case of Figure 3, the negative welfare consequences of an MPS are lessened here, even as the optimal degree of cooperation and the probability of exit are lower.

The analysis underlying Figures 3 and 4 has relied upon the assumption of independent realizations of θ and θ^*. We can also consider the special cases in which the two random variables θ and θ^* are perfectly positively or negatively correlated. Suppose that $\theta^* = \zeta + \eta\theta$ for some constants $\zeta > 0$ and η. As shown in Figure 2 we will observe cooperation only if realizations of θ and θ^* lie between the two exit boundaries; i.e. if $\theta > \kappa(\lambda, M)\,\theta^*$ and $\theta^* > \kappa(.)\,\theta/M^2$. Consider first the case of positive correlation so $\eta > 0$. Full cooperation ($\lambda = 1$) yields $\kappa = M/\sqrt{2}$ so the slope of the lower boundary (that for foreign exit) is $1/M\sqrt{2}$ while that of the upper boundary is $\sqrt{2}/M$. Given the symmetric supports centered on the initial (first period) values of the thetas, full cooperation will then occur if and only if

$$1/M\sqrt{2} < \eta < \sqrt{2}/M \implies 1 < \eta M\sqrt{2} < 2$$

If the countries have sufficiently similar sensitivities to each others' policies, so M is close to one, then η sufficiently close to one will always yield full cooperation. In terms of Figure 2 this result simply states that realizations lie on a straight line through the means that always falls between the two exit thresholds. Intuitively, if, say, country H gets a high realization then it wishes to be active in policy, but F will also get a high realization so policy activity by H also serves the interests of F (and *vice versa.*) With perfect negative correlation, on the other hand, we have $\eta < 0$. Now realizations lie on a downward-sloping line through the means and clearly exit could occur with more extreme realizations but a sufficiently compact distribution will lead to zero exit and full cooperation. However, a more disperse distribution may imply a non-zero probability of exit for $\lambda = 1$, and consequently the optimal degree of cooperation will fall. The intuition for this is isomorphic to that for the case of perfect positive correlation.

5 Concluding remarks

We mentioned in the paper's Introduction two recent episodes that have, at least in part, motivated this analysis. The first case was that of "Brexit": the decision taken by Great Britain in June 2016 to withdraw from the Economic Union. In our analysis, exit can occur as an equilibrium phenomenon, but we can also consider "shocks" that induce exit. An equilibrium interpretation of Brexit would suggest that policy-makers in Great Britain experienced an unusually low payoff from policy at the same time as the rest of the EU experienced high payoffs. The former means that British policy-makers would be less inclined to activist (and costly) policies but the latter means that, in a highly cooperative agreement British policy-makers would be expected to be policy active (as this benefits their partners in such a case). One of the critical factors in Brexit was immigration policy and one could certainly characterize British and European differences over this policy in this sort of fashion. Furthermore, to the extent that Britain under a Conservative government could be perceived as being less policy active than their European counterparts, this corresponds in our model to a lower mean policy sensitivity. This is, again, at least consistent with Brexit. Finally, to the extent that Britain was the smaller partner – thinking of the rest of the EU as a single partner – one would anticipate a greater sensitivity of Britain to EU policies than vice versa; this, too, is consistent with Brexit in our model.

One might also think of Brexit as a phenomenon driven by some external shock, rather than by particular realizations of variables where those realizations were known to be possible, if not likely. So, for example, if Britain were to experience a sudden, negative shock to its policy sensitivity (or the EU a positive shock of theirs) then, in our analysis this makes Brexit more likely. An equilibrium response to such shocks would be a lower degree of cooperation, and our analysis indicates that this would lead to a reduced overall probability of exit but, nevertheless, before such an equilibrium response occurs, realizations could be such that Brexit is induced.

The other case we mentioned in the introduction was TPP. While this is not an agreement that has actually been entered into, our interest was in the fact that so many critics of the proposal cite a loss of economic sovereignty as a reason to reject the proposed agreement. In the context of our analysis this amounts to a suggestion that the degree of cooperation required by TPP is too high given the projected gains to a party to the agreement. So this rather turns our analysis on its head: in our model the optimal degree of cooperation is chosen in order to maximize the expected gains from the agreement, but the implication of this criticism of TPP is that the proposed degree of cooperation is, essentially, exogenous and, in the eyes of its critics, so high that exit would be preferable. Again this is at least consistent with our analysis in that, *ceteris paribus*, increased cooperation in our model raises the probability of exit.

The bottom line of our analysis is that complete cooperation is unlikely to be achievable if it involves an exit risk due to the sovereignty constraint of countries. At the same time, the sobering implication is that any agreement will never try to avoid exit completely. These agreements therefore will be inherently unstable to some extent, and they will be so by design, not by oversight. However, we could also show that cooperation may increase with an increase in global uncertainty once the exit risk is already sufficiently large. Hence, exit risks will never completely avoided, but that does not necessarily imply that an increase in global uncertainty will reduce cooperation.

Our analysis also has stressed that it is the substantial difference in policy sensitivity realizations that makes exit likely. In the terminology of shocks, as explored above, it means that cooperation is much more likely to be stable if these shocks are positively correlated. Thus, our results are consistent with earlier results on the necessity of countries to follow similar business cycle patterns to form a successful currency union (see Mundell, 1961, for the seminal paper). Our results show that cooperation is less likely to be at risk when policy preferences are aligned, but at the same time, the exit threat will never completely disappear.

Appendix

A.1 Proof of Lemma 2

Define

$$H(\theta,\theta^*,\lambda)=\int_{\underline{\theta}}^{\tilde{\theta}(\theta^*)}\left[\Delta(\theta,\theta^*,\lambda)+\Delta^*(\theta,\theta^*,\lambda)\right]f(\theta,\theta^*)d\theta,$$

so

$$\Psi(\lambda)=\int_{\theta^{*\prime}(\lambda)}^{\overline{\theta}^*}H(\theta,\theta^*,\lambda)d\theta^*.$$

and

$$\Psi_\lambda(\lambda)=-\theta^{*\prime}_\lambda(\lambda)H(\theta,\theta^{*\prime}(\lambda),\lambda)+\int_{\theta^{*\prime}(\lambda)}^{\overline{\theta}^*}H_\lambda(\theta,\theta^*,\lambda)>0.$$

The sign here follows from $\theta^{*\prime}_\lambda(\lambda)<0$ and the fact that

$$\begin{aligned}H_\lambda(\theta,\theta^*,\lambda) &= \tilde{\theta}_\lambda(\theta^*,\lambda)\left[\underbrace{\Delta(\tilde{\theta}(\theta^*,\lambda),\theta^*,\lambda)}_{=0}+\Delta^*(\tilde{\theta}(\theta^*,\lambda),\theta^*,\lambda)\right]f(\tilde{\theta}(\theta^*,\lambda),\theta^*)\\ &+\int_{\underline{\theta}}^{\tilde{\theta}(\theta^*)}\left[\Delta_\lambda(\theta,\theta^*,\lambda)+\Delta^*_\lambda(\theta,\theta^*,\lambda)\right]f(\theta,\theta^*)d\theta>0.\end{aligned}$$

So, $H_\lambda(\theta,\theta^*,\lambda)>0$ because $\tilde{\theta}_\lambda(\theta^*,\lambda)>0$, $\Delta(\cdot)+\Delta^*(\cdot)>0$ for all $\lambda>0$, and $\Delta_\lambda(\cdot)+\Delta^*_\lambda(\cdot)>(\geq)0$ for all $\lambda<(\leq)1$. For $\lambda=\tilde{\lambda}$, we have $\tilde{\theta}(\overline{\theta}^*,\tilde{\lambda})=\underline{\theta}$ and $\theta^{*\prime}(\tilde{\lambda})=\overline{\theta}^*$, implying that

$$H(\theta,\theta^*,\tilde{\lambda})=\int_{\theta^{*\prime}(\lambda)}^{\overline{\theta}^*}H_\lambda(\theta,\theta^*,\tilde{\lambda})=0,$$

and, consequently, that $\Psi_\lambda(\tilde{\lambda})=0$.

A.2 The specific model

In period 1, the realizations of the θs are known and, analogously to what was shown before, the (known) aggregate gains from cooperation are

$$\begin{aligned}\Omega_1 &\equiv \frac{\gamma\lambda}{2}\left(2(\mu\theta_1)^2 - \lambda(\mu^*\theta_1^*)^2 + 2(\mu^*\theta_1^*)^2 - \lambda(\mu\theta_1)^2\right)\\ &= \frac{1}{2}\gamma\lambda(2-\lambda)\left((\mu\theta_1)^2 + (\mu^*\theta_1^*)^2\right)\end{aligned}$$

In period 2, the aggregate potential gains from cooperation are defined analogously for any particular realizations of Θ as $(\gamma\lambda(2-\lambda)((\mu\theta)^2 + (\mu^*\theta^*)^2))/2$. But, as Θ is random, so the expected aggregate potential gains from cooperation in period 2 are $\gamma\lambda\widetilde{W}/2$ where

$$\widetilde{W} = \int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}^*}^{\overline{\theta}^*}(2-\lambda)\left(\mu\theta)^2 + (\mu^*\theta^*)^2\right) f(\theta,\theta^*)d\theta^* d\theta.$$

However, this ignores the possibility of exit. If the realization of Θ were such that either country were to choose to exit, then both countries would be back in the non-cooperative setting. Consequently, the expression above for the expected aggregate potential gains from cooperation in period 2 needs to be adjusted for the probability of exit. Consider first the likelihood of domestic exit. As discussed, this occurs with positive probability iff $\kappa\theta^* > \theta$, and this can only be the case, given the supports of θ and θ^*, if $\underline{\theta}/\kappa < \overline{\theta}^*$. Under this assumption, we have four permutations of the losses from domestic exit, in terms of Figure 2: the "Home reneges" line either crosses the support of θ^* at θ in its interior or lies below $\underline{\theta}$, and either crosses the support of θ at $\overline{\theta}^*$ in its interior or lies everywhere below $\overline{\theta}^*$.

If we define $\theta^*_{max} = \max\{\underline{\theta}^*, \underline{\theta}/\kappa\}$, then the losses – again, actually foregone gains – associated with domestic exit are equal to $\gamma\lambda\Psi(\kappa)/2$ where

$$\Psi(\kappa) = \int_{\theta^*_{max}}^{\overline{\theta}^*}\int_{\underline{\theta}}^{\kappa\theta^*}(2-\lambda)\left(\mu\theta)^2 + (\mu^*\theta^*)^2\right) f(\theta,\theta^*)d\theta d\theta^*.$$

Similarly, the losses associated with foreign exit are $\gamma\lambda\Psi^*(\kappa^*)/2$ with a θ_{max} defined analogously, *mutatis mutandis*, and

$$\Psi^*(\kappa^*) = \int_{\theta_{max}}^{\overline{\theta}}\int_{\underline{\theta}^*}^{\kappa^*\theta}(2-\lambda)\left(\mu\theta)^2 + (\mu^*\theta^*)^2\right) f(\theta,\theta^*)d\theta^* d\theta.$$

where $\kappa^*\theta = \kappa\theta/M^2$. All up, then, the expected aggregate potential gains from cooperation in period 2, adjusted for the possibility of exit, are given by

$$\begin{aligned}\Omega_2 &\equiv \frac{\gamma\lambda}{2}\left(\widetilde{W}-\Psi-\Psi^*\right)\\ &= \frac{\gamma\lambda}{2}\left(\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}^*}^{\overline{\theta}^*}(2-\lambda)\left((\mu\theta)^2+(\mu^*\theta^*)^2\right)f(\theta,\theta^*)d\theta^*d\theta-\Psi-\Psi^*\right).\end{aligned}$$

With the common discount factor of δ, the overall expected aggregate gain from international cooperation is then equal to $\widehat{\Omega}=\Omega_1+\delta\Omega_2$. We then have the optimal λ defined implicitly by the first-order condition for λ of $d\widehat{\Omega}/d\lambda=0$ or, after simplifications

$$\frac{2(1-\lambda)}{\lambda(2-\lambda)}(\Omega_1+\delta\Omega_2)-\delta\lambda(\theta_1^2+\theta_1^{*2})-\frac{\delta\gamma\lambda(2-\lambda)}{2}\left(\frac{d\Psi}{d\kappa}+\frac{d\Psi^*}{d\kappa^*}\frac{\partial\kappa^*}{\partial\kappa}\right)\frac{\partial\kappa}{\partial\lambda}=0. \quad \text{(A.1)}$$

We observe from (A.1) immediately that $\lambda=1$ can never fulfill the first-order condition, as the first term would be equal to zero while the second term is negative for $\lambda=1$.

We now consider specific distributions and we consider the case in which θ and $\theta^*$are uniformly and independently distributed on supports $[\underline{\theta},\overline{\theta}]$ and $[\underline{\theta}^*,\overline{\theta}^*]$, respectively.[17] More specifically, we assume that the width of each support is 2Γ and $2\Gamma^*$, respectively, so, with means θ_1 and θ_1^*, we have $\overline{\theta}^*=\theta_1^*+\Gamma^*,\underline{\theta}^*=\theta_1^*-\Gamma^*,\overline{\theta}=\theta_1+\Gamma$ and $\underline{\theta}=\theta_1-\Gamma$.

For given λ (and hence κ), we calculate the *ex ante* probability of exit by each country. Denoting by $\text{PrEx}(i)$ the probability of exit by $i=H,F$, we have,

$$\text{PrEx}(F)=\frac{cM^2}{2\kappa}\left(\frac{\kappa\overline{\theta}}{M^2}-\underline{\theta}^*\right)^2,\ \text{where } c\equiv\frac{1}{(\overline{\theta}-\underline{\theta})(\overline{\theta}^*-\underline{\theta}^*)}=\frac{1}{4\Gamma\Gamma^*}.$$

Similarly,

$$\text{PrEx}(H)=\frac{c}{2\kappa}\left(\kappa\overline{\theta}^*-\underline{\theta}\right)^2.$$

Suppose there were no possibility of exit at all. Then the expected second period aggregate welfare gains from cooperation would be:

$$\frac{1}{2}\gamma\lambda\tilde{W}=\frac{1}{2}\gamma\lambda\left(\frac{1}{3}c(2-\lambda)\left[(\overline{\theta}^*-\underline{\theta}^*)\mu^2(\overline{\theta}^3-\underline{\theta}^3)+(\overline{\theta}-\underline{\theta})\mu^{*2}((\overline{\theta}^*)^3-(\underline{\theta}^*)^3)\right]\right).$$

Given the distribution of Θ, we can plug in for these various limits to obtain the following:

[17] A more comprehensive discussion of these and other simulation results is presented in a Simulations Appendix, available on request from the authors.

$$\tilde{W} = (2-\lambda)\mu^2\theta_1^2 + (2-\lambda)\mu^{*2}\theta_1^{*2} + \frac{1}{3}(2-\lambda)\left(\mu^2\Gamma^2 + \mu^{*2}\Gamma^{*2}\right).$$

From this, if $\mu = \mu^*$, so $M = 1$,

$$\tilde{W}(\mu = \mu^*) = (2-\lambda)\mu^2\left(\theta_1^2 + \theta_1^{*2} + \frac{1}{3}\left(\Gamma^2 + \Gamma^{*2}\right)\right).$$

Note that this is increasing and convex in Γ and Γ^*, so a mean-preserving spread (MPS) in either distribution increases expected welfare, absent the possibility of exit.[18]

Turning to the losses from exit, Ψ andΨ^*, and putting together the expressions for $\tilde{W}, \Psi, \Psi^*, \underline{\theta}, \overline{\theta}, \underline{\theta}^*, \overline{\theta}^*, \kappa, M$ and c, we can derive a messy but simulatable expression for $\widetilde{\Omega}(\lambda, \theta_0, \theta_0^*, \Gamma, \Gamma^*)$.[19]

References

[1] Bagwell, K., R. Staiger (2017), National sovereignty in an interdependent world. Mimeo. Available at http://www.dartmouth.edu/ rstaiger/sovereignty_11052017.pdf. Accessed 13/01/2018.

[2] Barrett, S. (1994), Self-Enforcing International Environmental Agreements, *Oxford Economic Papers*, 44: 878-894.

[3] Bulow, J., K. Rogoff (1989), A constant recontracting model of sovereign debt, *Journal of Political Economy*, 97: 155-178.

[4] Chander, P., H. Tulkens (1992), Theoretical foundations of negotiations and cost sharing in transfrontier pollution problems, *European Economic Review*, 36: 388-398.

[5] Dixit, A. (1987), How should the United States respond to other countries' trade policies? In R. Stern (ed.) U.S. trade policies in a changing world economy. MIT Press: Cambridge MA.

[6] Eaton, J., M. Gersovitz (1981), Debt with potential repudiation: Theoretical and empirical analysis, *Review of Economic Studies*, 48: 289-309.

[18] Given the uniform distributions, we have $\mathrm{E}(\theta) = \theta_1$ and $\mathrm{Var}(\theta) = \Gamma^2/3$ (and similarly for θ^*), so we can write $\tilde{W} = (2-\lambda)(\mu^2\mathrm{E}(\theta)^2 + \mu^{*2}\mathrm{E}(\theta^*)^2 + \mu^2\mathrm{Var}(\theta) + \mu^{*2}\mathrm{Var}(\theta^*))$.

[19] This is an illustration only. As noted before, the actual losses from exit will depend on how the exit threshold lines "cut" the support boxes in our figures and the simulations allow for all the various permutations that can arise.

[7] Folmer, H., P. van Mouche, S. Ragland (1993), Interconnected games and international environmental problems, *Environmental and Resource Economics*, 3:. 313-353.

[8] Horn, H., Tangerås, T. (2016), Economics and Politics of International Investment Agreements, IFN Working Paper No. 1140, Research Institute of Industrial Economics.

[9] Janeba, E. (2016), Regulatory Chill and the Effect of Investor State Dispute Settlements, mimeo.

[10] Kohler, W., Stähler, F. (2016), The Economics of Investor Protection: ISDS versus National Treatment, CESifo Working Paper No. 5766.

[11] Konrad, K. (2017), Large investors, regulatory taking and investor-state dispute settlement, *European Economic Review*, forthcoming.

[12] Korinek, A. (2016), Currency wars or efficient spillovers? A general theory of international policy cooperation. NBER Working Paper No. 23004.

[13] Krasner, S (2001), Rethinking the sovereign state model, *Review of International Studies*, 27: 17-42.

[14] Maggi, G. and R. Staiger (2015), Optimal design of trade agreements in the presence of renegotiation, *American Economic Journal: Microeconomics*, 7: 109-143.

[15] Martin, P. (2015), "Trans-Pacific Partnership: we're selling economic sovereignty for little return." Sydney Morning Herald October 13 2015. Available at http://www.smh.com.au/comment/transpacific-partnership-were-selling-our-sovereignty-for-little-return-20151011-gk6fhw.html. Accessed 30/9/2016.

[16] Mundell, R.A. (1961), A Theory of Optimum Currency Areas, *American Economic Review*, 51: 657-665.

[17] Philpott, D. (2001), Usurping the sovereignty of sovereignty? *World Politics*, 53: 297-324.

[18] Quiggin, J. (2001), Globalization and economic sovereignty, *Journal of Political Philosophy*, 9: 56-80.

[19] Rodrik, D. (2000), How far will international economic integration go? *Journal of Economic Perspectives*, 14: 177-186.

[20] The Economist (2016), "Brexit brief: Dreaming of sovereignty." March 19 2016. Available at http://www.economist.com/news/britain/21695056-talk-taking-back-power-may-be-delusional-more-democracy-not-dreaming-sovereignty. Accessed 30/9/2016.

[21] Trump, D. (2016), "Declaring American Economic Independence." Available at https://assets.donaldjtrump.com/DJT_DeclaringAmericanEconomicIndependence.pdf. Accessed 30/9/2016.

[22] Vote Leave (2016), http://voteleavetakecontrol.org/why_vote_leave.html. Accessed 30/9/2016.

[23] Warren, E. (2015), "The Trans-Pacific Partnership clause everyone should oppose." The Washington Post February 25 2015. Available at https://www.washingtonpost.com/opinions/kill-the-dispute-settlement-language-in-the-trans-pacific-partnership/2015/02/25/ec7705a2-bd1e-11e4-b274-e5209a3bc9a9_story.html?utm_term=.192c1abcc1b7. Accessed 30/9/2016.