

---

# The Optimal Design of International Trade Institutions: Uncertainty and Escape

B. Peter Rosendorff and Helen V. Milner

---

International institutions differ greatly in their forms; the number of states included, the decision-making mechanisms, the range of issues covered, the degree of centralized control, and the extent of flexibility within them all vary substantially from one institution to the next. What accounts for such variation? In this article, as part of the larger Rational Design project on the design of international institutions, we claim that such variation can be accounted for as part of the rational, self-interested behavior of states. We show that at least one important aspect of institutional design can be explained as a rational response of states to their environment.

Almost all international trade agreements include some form of “safeguard” clause, which allows countries to escape the obligations agreed to in the negotiations.<sup>1</sup> On the one hand, such escape clauses are likely to erode both the credibility and the trade liberalizing effect of international trade agreements. On the other hand, they increase the flexibility of the agreement by adding some discretion for national policymakers. The first question we address is the institutional design issue that escape clauses raise: when is such increased flexibility rationally optimal for states making international trade agreements? The answer to this question hinges on the costs of using escape clauses and retaining the overall agreement compared with not using them and abrogating the agreement.

Our second question concerns the effects of different institutional designs. If escape clauses allow states more flexibility in meeting their obligations, what impact does this have on their compliance with the agreement? What are the consequences

We thank the editors of the Rational Design project for their efforts: Barbara Koremenos, Charles Lipson, and Duncan Snidal. We are also grateful to two anonymous reviewers, and to James Morrow, Robert Pahre, Lisa Martin, Chris Canavan, and the editors of *IO*, David Lake and Peter Gourevitch, for many helpful comments and suggestions.

1. Hoekman and Kostecki 1995, 161.

*International Organization* 55, 4, Autumn 2001, pp. 829–857

© 2001 by The IO Foundation and the Massachusetts Institute of Technology

of increased flexibility for institutional performance: is cooperation enhanced, and is it more durable?

An escape clause is any provision of an international agreement that allows a country to suspend the concessions it previously negotiated without violating or abrogating the terms of the agreement. Escape clauses are a prominent feature of many international agreements and are included in most trade agreements. Not all international agreements, however, have such clauses; for instance, some international arms control agreements, such as the Strategic Arms Limitations Treaty (SALT) agreements, do not contain such escape mechanisms. Most trade agreements do contain them, but their nature often differs across agreements and they are usually vigorously contested in negotiations. For example, in both the North American Free Trade Agreement (NAFTA) and the General Agreement on Tariffs and Trade (GATT) Uruguay Round negotiations, antidumping and countervailing duty laws were critical issues that impeded agreement among the countries. Since its inception in the 1940s GATT (and the subsequent World Trade Organization, WTO) has slowly built an arsenal of safeguard mechanisms to protect states from import pressures in the wake of extensive trade liberalizing agreements. These include an escape clause, countervailing duty penalties, antidumping statutes, and a national security exception. For each of these, GATT (now the WTO) specifies the conditions under which a government can grant relief to an industry from import competition, and industries then have the option of choosing which mechanism to file their complaints under. In each of the GATT negotiating rounds, the inclusion and/or modification of these different laws have been the subject of intense debate among the signatories.

Many have noted that these different clauses can be substitutes for one another. Bernard Hoekman and Michael Leidy and Wendy Hansen and Thomas Prusa suggest that countervailing duty and antidumping laws are really “a poor man’s” escape clause.<sup>2</sup> Antidumping and countervailing duty complaints allege that exporting countries are playing unfairly and thus the harmed country avoids the payment of compensation that GATT requires on use of the escape clause. They are thus means for industries to limit import competition on the cheap: they enable a country to abrogate some portion of its treaty obligations under GATT and to pay a lower penalty than were they to use the escape clause. These and other measures, such as the infant industry and balance-of-payments exemptions in GATT, are all designed by governments to reduce domestic pressures to withdraw from the entire agreement when protectionist pressures grow at home. While these laws are generally seen as substitutes, they do differ substantially in the costs they impose on the countries using them. Usually antidumping and countervailing duty clauses are seen as less costly to use than traditional escape clauses. This type of variation is important, as we will explain later.

2. See Hoekman and Leidy 1989; and Hansen and Prusa 1995.

We make three central claims here. First, *escape clauses are an efficient equilibrium under conditions of domestic uncertainty*. When political leaders cannot foresee the extent of future domestic demands for more protection at home (and/or more open markets abroad), such clauses provide the flexibility that allows them to accept an international agreement liberalizing trade. A more general statement is that the greater the uncertainty that political leaders face about their ability to maintain domestic compliance with international agreements in the future, the more likely agreements are to contain escape clauses. In issue-areas where the impact of high uncertainty about domestic pressures to comply is less, governments are less likely to desire such safeguard measures.

We show that the use of an escape clause, a flexibility-enhancing device, in institutional design increases institutional effectiveness whenever there is domestic political uncertainty. We offer support therefore to Rational Design conjecture F1, FLEXIBILITY increases with UNCERTAINTY, as developed in the framing article by Barbara Koremenos, Charles Lipson, and Duncan Snidal.<sup>3</sup> This conjecture suggests that uncertainty about the state of the world rationally leads to the creation of institutional flexibility. Note that flexibility in this context refers to the ability to adapt and respond to unanticipated events within the context of a well-designed institutional system. The system itself is not subject to renewed bargaining. Alternative flexibility-enhancing devices are, of course, available: sunset provisions or anticipated renegotiations are often used. But we think that these mechanisms are even more costly and hence less used than the ones we examine.

Second, *for escape clauses to be useful and efficient they must impose some kind of cost on their use*. That is, countries that invoke an escape clause must pay some cost for doing so, or else they will invoke it all the time, thus vitiating the agreement. Paying this cost signals their intention to comply in the future. But the different costs of alternative escape clause measures will affect the frequency of their use. Less costly measures will be used more often. If governments understand this, they should rationally prefer the set of escape clauses that best matches the extent of protectionist pressure they expect to experience from domestic interests. Thus we anticipate that the architects of international agreements will design such agreements so that the costs of the escape clauses they most desire are balanced by the benefits of future cooperation. Variation in the nature of the escape clause mechanism, primarily its cost, is thus an important feature of different agreements. If states rationally design such agreements, we should expect such variation to be an important element of the bargaining process.

The exact size of the cost will depend on the gains from cooperation relative to the benefits of defection; they are a function of what might be called the “preference configuration.” The costliness of the escape clause is crucial to the effectiveness of the escape clause regime, and the preferences of the domestic players in the negotiating countries will affect the optimal choice of this cost. We claim, therefore,

3. Koremenos, Lipson, and Snidal, this volume.

that domestic preferences and institutions matter in the design of optimal international institutions.

Third, we argue that *including escape clauses makes initial agreements easier to reach*. Their flexibility allows states to be reassured about the division of the long-term gains from the agreement. Indeed we claim that without escape clauses of some sort many trade agreements would never be politically viable for countries. This fits with Rational Design conjecture F2, FLEXIBILITY increases with DISTRIBUTION problems; that is, increased flexibility (necessary to deal with the uncertainty about future states of the world) lessens the problems of bargaining and distribution that may plague an initial agreement.

We use a formal model to examine why countries might desire escape clauses and how this type of institutional design might affect an institution's performance. We examine a two-stage game: an international bargaining game where an agreement over the design of the institution is adopted and then a repeated trade (sub)game where the countries set their trade policies, given the design of the institution.

### *Escape Clauses and Political Uncertainty*

The key factor that renders escape clauses desirable is the presence of uncertainty. In each period the political pressure for protection at home (and/or for more open markets abroad) is subject to a shock. Some unanticipated change in the economy or political system produces a bigger or smaller value for the impact of domestic firms' demands for protection. We model this shock very generally; it is any exogenous and unanticipated change in the state of the world (such as price or supply changes, technological change, political change) that affects domestic firms' demand for, or ability to lobby for, protection of their markets.

Although we model uncertainty as a political shock, we recognize that the strength of the political support for protection (or for liberalization) is determined by many factors, for instance,

- unexpected price or supply shifts that intensify international competition may induce enhanced lobbying efforts by domestic firms;
- changes in production technology that reduce employment in a sector, and hence its political clout;
- changes in a country's political institutions or preferences: tastes might change in favor of enhanced protection, or campaign finance reforms might alter the political pressure that firms can apply;
- changes in domestic political cleavages or alignments that might make a previously pivotal sector less influential in domestic politics, implying that protection is politically less expedient.<sup>4</sup>

4. For analytical tractability, we assume in the model that the shocks in each country are independent. Price shocks—for example, an unexpected rise in the price of an input or the emergence of a

We assume that in the current period leaders in each country know their own domestic political situation but that both sets of leaders are equally uninformed about the degree of political pressure at home and abroad in all future periods. We show that uncertainty about the state of the world creates conditions favorable for the use of escape clauses. That each country has limited knowledge about the domestic politics of the others is central to our argument; furthermore, this uncertainty has a lasting impact because each country faces new shocks in each period that determine the amount of political pressure that domestic groups exert.<sup>5</sup>

The two stages of the model combine the two critical elements of cooperation theory: bargaining and enforcement. The trade game played by the countries is a modified version of a repeated Prisoners' Dilemma. In this second stage enforcement is critical; the temptation to cheat makes cooperation very difficult, especially in international politics where third-party enforcement is absent. Countries must be punished if they protect, but sometimes because of domestic shocks leaders will be forced to protect when they would otherwise want to maintain the agreement—or, forced to undertake “involuntary defection,” as Robert Putnam calls it.<sup>6</sup> Such equilibria to the Prisoners' Dilemma are often supported by the requirement that each player automatically punish the other when cheating is observed, and continue to punish forever or for long periods of time. If their discount value is high enough and punishment is sure and strong enough, then they will resist the temptation to cheat. This set of results has often been used to argue that cooperation in international politics is possible, if not frequent.<sup>7</sup> But such punishment often implies abrogation of the entire agreement.

George Downs and David Rocke show that shorter punishment periods can also support cooperation.<sup>8</sup> They identify domestic political uncertainty as an explanation for “imperfect” treaties, where imperfection is measured relative to the “most cooperative” agreement possible. Using a repeated Prisoners' Dilemma game with trigger strategies, where defections are punished by the other player for a limited number of periods, they argue that domestic political uncertainty leads to agreements with shorter punishment periods and therefore less cooperation. But what if

---

third-country competitor—that affect the lobbying strength of firms at home may simultaneously affect the lobbying strength of firms abroad. Allowing for correlated shocks would not alter our central result; agreements with escape clauses allow countries the option to temporarily exit when political pressure is unexpectedly intense, and when this defection is tolerated by the trading partners in the interests of the system's stability.

5. Uncertainty here concerns the “future state of the world”: the configuration of political pressure in future periods is not known with certainty. Uncertainty regarding the preferences of key domestic players is another possibility, one we consider elsewhere in an investigation of the effect of elections on the design of international agreements (Milner and Rosendorff 1997). Alternatively, the agreement itself is too complex (or time is too valuable) for the domestic policymakers to fully understand the consequences of its passage, and policymakers therefore rely on the information provided by lobbies and other interested third parties. Milner and Rosendorff 1996.

6. Putnam 1988.

7. See Axelrod 1984; and Oye 1986.

8. Downs and Rocke 1995.

countries every now and then face intense pressure to cheat yet do not want to spark retaliation and a breakdown in cooperation? Can an alternative institutional structure be devised to maintain a cooperative agreement, even in these periods of high political pressure to protect? In the presence of exogenous shocks, international institutions may be much better served by allowing countries to make temporary, ad hoc use of escape clauses that permit them to break the rules for a short period and pay a cost to do so. There is no retaliation. The defection is tolerated, exactly because the other side may wish to use the same instrument in the future.<sup>9</sup> Cooperation, as we demonstrate, is deeper and more likely, and international trade institutions are more durable, with escape clauses than without them. In the choice between rules and discretion, therefore, rules with costly discretion may be better than no discretion when the future holds unexpected, unpleasant surprises. Our first key result is that greater domestic uncertainty makes the inclusion of escape clauses more likely in international agreements.

Many trade agreements include such escape clauses; indeed, all GATT agreements have at least one type, if not several types, of such escape clauses. Moreover, these alternative escape mechanisms have different costs for their use. In general a country appealing to an escape clause is allowed, under the institution's rules, to protect the affected industry for the duration of that period, as long as it (in effect) voluntarily and publicly incurs some penalty. This voluntary penalty is consistent with the reciprocity norm of GATT, which requires a country that applies a temporary trade barrier to reciprocate by lowering some other barrier elsewhere so that its trading partners are unaffected by the action or to face equivalent trade barriers by its partners.

But this penalty may take any number of forms. For example, countries using GATT's escape clause must negotiate compensation with the affected exporter or face equivalent retaliation from the exporter. For other safeguard measures, the cost is often smaller and less explicit. Sometimes there is a presumption that a country invoking the escape clause will be forced to devise and implement a plan of structural adjustment for the affected industry; such plans have costs, both economic and political. Moreover, the costs of filing an escape clause, antidumping, or countervailing duty complaint are also part of the cost that the import-competing firms must face. For many of these the technical and legal requirements for producing evidence of injury are sufficiently high to merit consideration. In any case each safeguard mechanism entails some costs when it is used, although these costs do differ in important ways.

After invoking the safeguard, in the next period the country returns to the cooperative regime, having preserved its reputation as a cooperator. Moreover, no supranational enforcement agency must force the country to pay this penalty; the

9. Very little retaliation for treaty violations is actually observed. Under current WTO rules, any punishment can only come after a finding by the dispute settlement procedure at the WTO, and frequently the dispute is "settled" before punishments are applied. The pre-Uruguay Round rules in fact made findings of allowable retaliation quite rare. Rosendorff 1999.

country (and everyone else) realizes that paying the penalty will preserve its credibility in the future. The institution serves as a verification agency, much as the Law Merchants institution does;<sup>10</sup> it monitors whether defection occurs with a penalty.

### *Low Costs, Frequent Escape*

The cost a state must pay for using the escape clause is of great importance. If the penalty is set at an appropriate level, a country may temporarily use the escape clause and then return to the cooperative path. If the cost is too high, countries will abandon the institution and defect when they experience a severe shock. If the cost is too low, there is repeated recourse to the escape clause, and the agreement enforces little actual cooperation over time. Escape clauses will thus be used more often when their costs of use are lower. This implies that policymakers should attempt to design efficient escape clauses; they should act so that the incentive to exercise relief is balanced with the gains from cooperation. Variations in the costs of different escape clause mechanisms will be an important feature of the rational design of international trade agreements.

The first stage in the model focuses on bargaining over the size of the escape clause penalty. When will countries be able to agree to such escape clauses? In particular, when will they be able to agree to impose a cost on themselves for using the escape clause, and when will this be credible? Furthermore, when will they pick a level of costs so that the optimal degree of cooperation is induced? To address this issue, we model a first stage before the trade barrier setting game is played. In this stage the countries bargain with each other over the penalty they are willing to pay for using an escape clause. One can think of this as bargaining over the nature of the trade agreement itself. Thus making an agreement means agreeing on a value for the penalties that all countries will (voluntarily) pay to use an escape clause. We show that the countries negotiate an optimal penalty, one that balances the need for as much cooperation as possible, while allowing some flexibility in times of domestic political pressure. Such a penalty must not be too high or it will eliminate any flexibility and make the system unstable; but it must also not be too low or it will render "cooperation" ineffective. In effect international institutions that are able to adopt an escape clause should do so in ways that generate more durable and stable cooperative regimes.

The escape clause itself is endogenous to the model: choosing a prohibitive cost for invoking the escape clause is equivalent to ruling it out of the institutional structure. Yet in equilibrium we show that the negotiating parties adopt an escape clause with moderate costs. While such bargaining can have distributional consequences, we study only the symmetric case here where the two countries are identical; nevertheless, our model combines both bargaining and enforcement problems.

10. Milgrom, North, and Weingast 1990.



*Agreements Are Easier to Conclude*

Our model also touches on a point made by James Fearon. He uses a model somewhat like ours, which combines a bargaining game in the first stage and a Prisoners' Dilemma game in the second. He points out that "as the shadow of the future lengthens, both states choose tougher and tougher bargaining strategies on average, implying longer and longer delay till cooperation begins."<sup>11</sup> That is, as the possibility of durable cooperation grows in the second-stage Prisoners' Dilemma, the possibility of stalemate in the first-stage bargaining game rises. Hence, making agreements easier to enforce may make them harder to initially conclude, since the distribution of gains set initially will be so important and fixed throughout the future.

Here, the inclusion of escape clauses may reduce this dynamic. That is, if in future periods players can deviate, pay a penalty, and return to cooperation, this escape clause may mean that their initial distributional bargaining is not so important. The pattern of distributive gains agreed upon today may be altered in the future through the use of the escape clause. Therefore, inclusion of an escape clause may have another benefit: it may make agreements easier to conclude initially! We provide some evidence that certain agreements would not have been politically feasible had they not included escape clauses. This is our third result.

**The Model**

Consider a world with two countries, home and foreign, that trade a single good. The good is produced by a single firm in each country, and hence there is reciprocal dumping or cross hauling. The profits of the home firm depend therefore on the trade barriers at home,  $t$ , which raise the domestic price and are good for profits, and the trade barriers abroad,  $t^*$ , which reduce exports and induce a fall in the home firm's profits; hence, firm profits are a function of both, that is,  $\Pi(t, t^*)$ .

*Government Objectives*

A government's utility depends on the sum of consumer surplus,  $CS(t)$ , which falls with  $t$ , producer surplus or profits,  $\Pi(t, t^*)$ , which rise with  $t$  and fall with  $t^*$ , and tariff revenues,  $tM(t)$ , which first rise and then fall with the level of the barriers. Let  $\gamma > 0$  denote the weight that a government attaches to the firm's profits. The home government's (one period) utility function then is  $W(t, t^*) = CS(t) + \gamma\Pi(t, t^*) + tM(t)$ . Similarly, for the foreign government,  $W^*(t, t^*) = CS^*(t^*) + g\Pi^*(t^*, t) + t^*M^*(t^*)$ , where  $g > 0$  is the weight of the foreign firm's profits in its government's utility function.

This objective function is "politically realistic" in Richard Baldwin's sense; that is, governments desire to maximize consumer surplus because it helps them recruit

11. Fearon 1998, 282.



votes, but they also value firm profits for the contributions and political support that firms can give them.<sup>12</sup> This utility function is also consistent with the objective function used in Gene M. Grossman and Elhanan Helpman's model of lobbying and campaign contributions.<sup>13</sup> Here, governments are concerned with their reelection and hence have political economy motivations.

In the following sections we use these utility functions to define the payoffs for each outcome that the governments can arrive at in a simple noncooperative trade barrier setting game. These payoffs resemble those of a standard Prisoners' Dilemma: mutual cooperation, which we call the Pareto optimal outcome; mutual defection, or the Nash equilibrium; unilateral defection; and the sucker's payoff. This defines what happens in the second-stage trade game.

### *Political Uncertainty*

Policymakers seeking to maximize their political support choose to adopt trade policies that redistribute revenue among politically salient groups. Here policymakers are balancing the interests of consumers with those of the firms. In each period the political pressure exerted by firms is subject to a shock. Some unanticipated change in the economy or political system allows firms to exert a larger or smaller amount of political pressure. We have deliberately chosen to be vague about the specific nature of this shock—for example, whether it is political or economic. This gives our model greater explanatory breadth. Any exogenous and unexpected change that alters the impact of domestic firms on the demand for protection is relevant. In some periods, firms' political influence will take on a "low" value; in others, however, the pressure applied by the domestic industry is "abnormally" high. The same is true in the other country: its leaders have the same objective function and face the same forms of political pressure. Notice that the firms can be either import competitors or exporters. As defined here, a period of unusually "high" political pressure applied by the firms means a heightened demand by the firms for higher trade barriers at home and lower ones abroad.<sup>14</sup>

In any period,  $\gamma$  and  $g$  are stochastic and are independently and identically distributed with distribution  $\Phi$ : this captures the notion that *ex ante* policymakers are not fully informed about the degree of political pressure to protect local industry that they might experience in any future period. At home, some unanticipated change in the economy or political system creates a larger or smaller value of  $\gamma$ . The same is true in the other country: its leaders have the same objective function and face the same forms of political pressure. For simplicity, we assume that in the

12. Baldwin 1987.

13. Grossman and Helpman 1994.

14. The reader may be tempted to draw a contrast with Milner 1988. There export interests organize in favor of lower domestic tariffs. That is an equilibrium outcome, however, not a statement about preferences. In that model, exporters simply prefer lower tariffs abroad, and adopt, for strategic reasons, political action domestically so that tariff concessions at home can be traded for concessions abroad. A similar dynamic is at work here: firms are willing to trade lower tariffs at home for lower tariffs abroad.

current period each country knows its own state of politics but not the other's, and that both are equally uninformed about the values of  $\gamma$  and  $g$  (at home and abroad) in all future periods. That each country has limited knowledge about the domestic politics of the other is central to our argument; furthermore, this uncertainty has a lasting impact because each country faces new shocks in each period that determine the amount of political pressure domestic groups exert. Uncertainty about the state of the world in the other country creates conditions favorable for the use of escape clauses.

While we model the political uncertainty as exogenous (and hence as uncertainty about the state of the world), national preferences are actually an aggregation of the preferences of the domestic groups. Individual preferences per se do not change, but national ones might as the intensity of firms' demands change. Each player thus is uncertain about how influential various domestic groups are likely to be in the future when policymakers choose their trade policies. In the future each government may be easily capturable by the protectionist lobby, or it may be able to stand firm in the face of protectionist pressure. Neither player knows beforehand which of these types the other is likely to be.

### Without an Escape Clause: Prisoners' Dilemma Game

#### *Under Political Optimum (Cooperation)*

First, we find the pair of trade barriers that maximize the sum of the two governments' utility functions. If  $\gamma$  and  $g$  are known, we can define the cooperative solution:

$$(t^P(\gamma, g), t^{*P}(\gamma, g)) = \arg \max_{(t, t^*)} (W(t, t^*) + W^*(t, t^*)).$$

Denote the utility of each of the governments under the political optimum as

$$P(\gamma, g) = W(t^P, t^{*P}) \quad \text{and} \quad P^*(\gamma, g) = W^*(t^P, t^{*P}).$$

#### *Under Nash Equilibrium*

Under the Nash equilibrium (NE), each player chooses a level of domestic trade barriers as a best response to the behavior of the opponent. In any period in which  $\gamma$  and  $g$  are known we can solve for the Nash equilibrium in trade barriers for that period. Let

$$t(t^*) = \arg \max_t W(t, t^*) \quad \text{and} \quad t^*(t) = \arg \max_{t^*} W^*(t, t^*).$$

Solving these simultaneously leads to the Nash pair of trade barriers  $(t^N, t^{*N})$ . Denote home government's utility under the Nash equilibrium as

$$N(\gamma, g) = W(t^N(\gamma, g), t^{*N}(\gamma, g)).$$

*Defection*

Home’s optimal defection (when foreign cooperates) is

$$t^D = \arg \max_t W(t, t^{*P}),$$

and its utility under the optimal defection is

$$D(\gamma, g) = W(t^D(\gamma, g), t^{*P}(\gamma, g)).$$

If instead foreign defects and home cooperates, home receives the sucker’s payoff:

$$S(\gamma, g) = W(t^P(\gamma, g), t^{*D}(\gamma, g)).$$

*Prisoners’ Dilemma*

So we have  $D(\gamma, g) > P(\gamma, g) > N(\gamma, g) > S(\gamma, g)$ , a Prisoners’ Dilemma game as represented by the standard  $2 \times 2$  normal form matrix:

	$P^*$	$D^*$	
$P$	$P(\gamma, g), P^*(\gamma, g)$	$S(\gamma, g), D^*(\gamma, g)$	(1)
$D$	$D(\gamma, g), S^*(\gamma, g)$	$N(\gamma, g), N^*(\gamma, g)$	

To simplify the notation,  $D(\gamma, g) - P(\gamma, g) \equiv B(\gamma, g)$ . Each player is susceptible to political pressure both to protect against foreign imports and to open export markets; in the future both are equally unsure how much pressure each will experience. Hence, home must make its best guess about the value of raising domestic trade barriers (defecting) in any period by taking expectations over  $g$ ; we denote this best guess by

$$B(\gamma) = \int_g B(\gamma, g) d\Phi.$$

Similarly, both players are completely uninformed about the possible draws of  $\gamma$  and  $g$  in any future period. Hence, the values of  $P(\gamma, g)$  and  $N(\gamma, g)$  are unknown for future periods. Expectations can be formed however; denote

$$P = \int_{\gamma} \int_g P(\gamma, g) d\Phi d\Phi \quad \text{and} \quad N = \int_{\gamma} \int_g N(\gamma, g) d\Phi d\Phi.$$

The Prisoners' Dilemma in Matrix (1) is played in the presence of uncertainty; as in the standard Prisoners' Dilemma; however, a cooperative equilibrium in trigger strategies can be supported by a large enough discount rate.

LEMMA 1. *A pair of grim trigger strategies (cooperate until a defection is observed, then punish forever) is an equilibrium to the game in Matrix (1) for all*

$$\delta > \frac{\max_{\gamma} B(\gamma)}{P - N + \max_{\gamma} B(\gamma)}.$$

The (expected) incentive to defect in any period with draw  $\gamma$  is  $B(\gamma)$ . The largest value that  $B(\gamma)$  can take is  $\max_{\gamma} B(\gamma)$ . If this maximal incentive to defect is less than the present discounted expected value of future punishments  $[\delta/(1 - \delta)] \times (P - N)$ , cooperation is possible.

### Escape Clause Game

In any period of the Escape Clause game, a player can take the Pareto action, that is, play  $P$  as in the Prisoners' Dilemma above; or it can exercise an escape clause  $EC$  at cost  $k$ ; or it can defect  $D$  as before. The stage game is now 3 by 3:

	$P^*$	$EC^*$	$D^*$
$P$	$P(\gamma, g), P^*(\gamma, g)$	$S(\gamma, g), D^*(\gamma, g) - k$	$S(\gamma, g), D^*(\gamma, g)$
$EC$	$D(\gamma, g) - k, S^*(\gamma, g)$	$N(\gamma, g) - k, N^*(\gamma, g) - k$	$N(\gamma, g) - k, N^*(\gamma, g)$
$D$	$D(\gamma, g), S^*(\gamma, g)$	$N(\gamma, g), N^*(\gamma, g) - k$	$N(\gamma, g), N^*(\gamma, g)$

Define "cooperation" as the play in any period of  $P$  or  $EC$ . Define defection as the play of  $D$  in any period.

DEFINITION 1. *An escape clause strategy (for home) is a strategy in which home plays  $D$  if  $D^*$  has been played in any period in the past, otherwise home plays  $P$  if  $B(\gamma) < k$ , plays  $EC$  if  $k \leq B(\gamma) \leq K$ , and plays  $D$  if  $B(\gamma), k > K$  for some  $K$  to be defined later.*

The extent of the exogenous shock determines the gains to be had from defection in this period; these gains rise with the political pressure that the firms can bring to bear; that is,  $B'(\gamma) > 0$ . If these gains are small ( $B(\gamma) < k$ ), the government sticks to its Pareto optimal strategy, play  $P$ . If the penalty is not too onerous ( $k < K$ ), moderate gains from defection ( $k \leq B(\gamma) \leq K$ ) cause the government to appeal to

the escape clause, *EC*. If the gains from defection are very large and the escape clause penalty is large, that is,  $B(\gamma), k > K$ , the government ceases to cooperate entirely. A useful way to summarize the government's strategy is to say that the government cooperates (by playing *P* or *EC*) when  $\min(B(\gamma), k) \leq K$ , and defects otherwise.

The critical value of  $K$  is determined as the cost that would make any player of this game exactly indifferent between exercising the escape clause and then returning to the cooperative regime, and defecting and exiting the system forever. It is intuitive, therefore, that if the costs of the escape clause and the gains from defection are large, the government will cease to cooperate entirely.

PROPOSITION 1. *A pair of escape clause strategies is a Nash equilibrium. All the proofs are in the appendix. Notice that in the standard Prisoners' Dilemma game, Matrix (1), cooperation is sustained only for discount factors that are large enough; that is,*

$$\delta > \frac{\max_{\gamma} B(\gamma)}{P - N + \max_{\gamma} B(\gamma)}.$$

However, in the escape clause equilibrium here cooperation can be sustained for any value of the discount factor as long as  $k \leq K$ . Recall that at cost  $K$ , any player is indifferent between the escape clause and defection; if  $\delta$  falls, future cooperation is valued less, and the critical  $K$  falls. Hence, the cost of exercising the escape clause must fall as well. So a low discount factor can still produce cooperation. Cooperation now is more flexible in that temporary defection is now possible—unlike in the standard Prisoners' Dilemma, where no defection of any kind was permissible.

One particularly appealing aspect of this equilibrium in the context of institutional design is that the penalty associated with the escape clause is self-enforcing. Any country that wishes to exercise the escape clause in an agreement must visibly penalize itself; no external enforcement agency is required. For a defector to avoid being punished, it must pay the penalty  $k$  in a visible way. The international institution is an information provider rather than an enforcer here: it is entrusted as an agent of the contracting states to check that each country that adopts an escape clause pays a penalty and to inform the others of this. Only when penalties are not paid do the states need to punish each other.

COROLLARY 1. *There exists an agreement with an escape clause that Pareto dominates one without it in the presence of political uncertainty.*

In any period in which the escape clause is exercised, there is no "true" cooperation: the escaping player is defecting, and the defection is being tolerated. Hence, the value of the game under an escape clause equilibrium will decrease as the use of the escape clause increases. If the escape clause is used infrequently or not at all, there is more "true" cooperation; however, domestic political uncertainty is likely to lead at some point to a complete breakdown of the regime, and then the

punishment phase will be applied forever. This corollary establishes that either there is an escape clause with a level of cost that induces enough cooperation and no breakdown such that the value of the game in an escape clause equilibrium is larger than that of the same game without an escape clause, or the cost of escape is too high and the escape clause equilibrium is the same as the grim-trigger equilibrium of the standard Prisoners' Dilemma. Hence, there is an escape clause cost such that the escape clause equilibrium Pareto dominates (perhaps weakly) the grim-trigger equilibrium of the game without an escape clause.

Notice that the more salient the domestic political uncertainty, or the greater its likely impact on electoral returns, the more likely are political leaders to view an escape clause as an essential element of any agreement.

### **Uncertainty and Escape Clauses: Implications and Some Evidence**

As already noted, most international trade agreements include at least one form of escape clause, and many include several. Our claim is that this prevalence of escape clauses is due to the high levels of domestic uncertainty that surround trade politics. We predict that domestic uncertainty affects the use of escape clauses. Greater domestic uncertainty, or situations where political leaders are more sensitive to unanticipated changes in political pressures, should be associated with more reliance on escape mechanisms. An interesting test of our model would be to identify those political institutions that magnify the effect of unanticipated shocks and see whether countries with these types of institutions are more likely to devise and use escape clauses in their trade relations. Another test would be to deduce which issue-areas are more subject to unanticipated domestic shocks and see if they are more likely to have escape clauses associated with them. Such an exercise, unfortunately, is beyond the scope of this article. However, we can suggest two facts about escape clauses that accord with our theory: certain countries that arguably are more sensitive to domestic pressures are the main proponents and users of escape mechanisms, and certain issue-areas seem more likely to have escape clauses than others due to their greater levels of uncertainty.

Escape clauses in trade policy exist both at the national and the international level. Interestingly, international usage has often copied domestic laws. It is notable that several countries dominate the international use of all forms of escape clauses and that all of these countries have tended to use escape clauses domestically first. The main countries using GATT (now WTO) antidumping, countervailing duty, and safeguard clauses are the same ones that earlier developed a battery of domestic laws to use these trade remedies. By and large, the United States, Canada, the European Union, and Australia are the main users of these clauses.<sup>15</sup> These are the same countries that initially built domestic trade laws around such escape mecha-

15. Trebilcock and Howse 1995.

nisms. The first instance of an antidumping law was Canada's 1904 dumping regime.<sup>16</sup> In 1947 the United States instituted the world's first safeguard clause.<sup>17</sup> And the United States and Canada were both the early designers of countervailing duty laws. This suggests that the need for escape clauses may be associated with democracies. It may well be that unanticipated shocks are far more damaging for political leaders in democracies than in nondemocracies. These shocks may be more likely to get them ejected from office as the negatively affected groups mobilize against the incumbents in election periods. If so, this would account for why these types of countries are more likely to have such national escape clause provisions and why they are also more likely to be proponents of these provisions at the international level.

In the realm of safeguard clauses, for example, it is the United States that has the oldest domestic laws and has been the most vocal proponent of them in international trade negotiations. U.S. trade law puts the escape clause into practice through Section 201 of the Trade Act of 1974. Following a petition—from the industry or from government (the president, the U.S. Trade Representative, or Congress, among others)—the U.S. International Trade Commission (ITC) conducts an investigation to evaluate whether imports have threatened to injure or been a substantial cause of injury to the domestic industry. After an affirmative finding by the ITC, the president may grant protection for up to five years, with the possibility of extending it for another three.<sup>18</sup> This practice has been followed closely in GATT, largely at the United States' insistence. Article XIX of GATT permits a member to escape from its obligations not to raise trade barriers when one of its industries is suffering an economic downturn and is experiencing "serious injury."

In the realm of antidumping and countervailing duties the same association is apparent. U.S. and Canadian laws have preceded international ones and set the pattern for them. Article VI of GATT, and the Second Antidumping Code of the Tokyo Round, which define practice in antidumping and countervailing duty law, allows member states to apply duties when imports are sold at "less than fair value," following U.S. practice. Ronald A. Cass and his colleagues describe the U.S. antidumping laws (and those of other countries) as "miniature escape clauses," in that the antidumping code extends protection to smaller cases on which agreement would be impossible *ex ante*.<sup>19</sup> Similarly, the U.S. countervailing duty code (which is consistent with GATT's Art. VI) allows member states to apply a countervailing duty when a subsidy is being provided to the foreign industry.<sup>20</sup> Other forms of the

16. *Ibid.*, 172.

17. *Ibid.*, 227.

18. Between 1975 and 1990, ninety-two cases under sec. 201 were initiated, of which thirteen industries received relief and seven more received trade adjustment assistance. High profile cases included color televisions in 1982, which received protection on \$1,543 million of imports that year, and nonrubber footwear, \$2,480 million in 1981. Hufbauer and Rosen 1986.

19. Cass et al. 1997, 24.

20. Between 1994 and 1999 alone, 77 antidumping petitions were filed in the United States. Stern 1997. Worldwide, the antidumping clause has been invoked over two thousand times since 1970.



escape clause appear throughout GATT. Balance-of-payments exceptions (Art. XVIII and XII), infant-industry protection (Art. XVIII), and tariff renegotiation (Art. XXVII) allow temporary escape from a member's obligations under the agreement.

Trade is, of course, an area where governments are likely to face strong domestic pressures for import protection from time to time. When imports surge or when economic conditions facing an industry turn downward, pressures for protection may suddenly appear. Unfortunately, governments may not be able to anticipate perfectly the magnitude of such pressures or their origin. Cass and his colleagues claim that these safeguard mechanisms allow "protectionist sentiment to hold sway" when political pressures are large.<sup>21</sup> Democratic leaders may be especially vulnerable to such unexpected changes, and hence may seek escape clause protection more than leaders in other systems. The greater impact of uncertainty in democratic systems may make their leaders particularly desirous of escape clause mechanisms in trade.

The need for escape clauses may also vary by issue-area. It is widely believed that trade is an area where governments face domestic uncertainty that has significant costs; such international economic exchanges are susceptible to swift changes due to price or supply shocks, technological change, and/or foreign government policy changes. The same is true in the macroeconomic area. Fixed exchange-rate systems may be especially vulnerable to unanticipated domestic pressures to devalue. High uncertainty over the timing and magnitude of these domestic pressures seems likely. Thus we see escape clause measures in many fixed exchange-rate agreements. In the Bretton Woods regime, for example, the simple rule was the requirement to maintain fixed exchange rates. But a country could devalue in the event of "fundamental disequilibrium," a vague phrase allowing escape from the simple rule since even economists were unable to agree on what balance-of-payments equilibrium meant. The regime did not dictate in advance the size of the devaluation. Instead, it required a member state to seek approval from the International Monetary Fund (at least for an exchange-rate realignment of more than 10 percent).

The European Payments Union, the postwar multilateral trade-deficit clearing system, gave signatories the right to suspend liberalization measures in the event of serious economic disturbance or if liberalization was too disruptive.<sup>22</sup> Similarly, Europe's Exchange Rate Mechanism required member states to maintain bilateral exchange rates within clearly demarcated target zones, but did allow for realignments of the parity. While the architects of the mechanism recognized the need for occasional parity realignments, they did not specify exactly when such realignments should take place. Instead, realignments were required to be negotiated among all

21. Cass et al. 1997, 24.

22. Oatley, this volume.

members.<sup>23</sup> In all three cases, escape clause mechanisms were included in the design of these institutions to deal with situations where policymakers face high levels of domestic uncertainty over the pressures that will arise for them to abrogate any international agreement they sign.

Notice that under all three regimes (Bretton Woods, European Payments Union, and Exchange Rate Mechanism), devaluation (the use of an escape clause) was not without cost. Devaluation was permitted only in concert with other measures designed to bring core macroeconomic aggregates back to within “acceptable” levels. Devaluation was therefore frequently associated with fiscal and monetary contraction and policy liberalization and reform, all of which come at a domestic political price.

In some noneconomic issues, uncertainty may be consequential enough so that temporary noncooperation may arise as an equilibrium in isolated cases. James D. Morrow, for instance, argues that prisoners of war treaties are often robust in the face of frequent battlefield violations of the rules of war in an environment where monitoring and acquiring accurate information are very costly.<sup>24</sup> Moreover, similar to our model, violations must be policed by the violators themselves, and punishment (in the case of gross violation) must be publicly implemented for cooperation to be sustained. But in other noneconomic areas, it seems that domestic uncertainty is less pervasive and consequential. In an area like arms control, the public and interest groups tend to be less organized and involved. The most important constituent of these agreements is often the military, which may take part in the negotiations and hence shape them directly. The impact of unexpected changes in this area may be less for political leaders than in areas like trade. Notably, arms control agreements have frequently not included escape clauses. The Antiballistic Missile Treaty, most of the SALT agreements, and the Intermediate-range Nuclear Forces treaties do not contain escape mechanisms; some of these allow countries to withdraw with certain notification provisions, and some have definite time limits, but none seem to contain clauses that allow temporary abrogation of the agreements. This suggests, if our claims are correct, that arms control is an area where domestic uncertainty is less important for leaders. Unexpected shocks that greatly increase pressures for leaders to cheat on the agreement (or pay substantial domestic costs), are less common in this area. Hence, one would not expect states to be as concerned about including escape clauses in these agreements as they are in trade and the monetary area. Where domestic uncertainty is less consequential for leaders, escape clauses will be less important and hence less used. We return to this question later.

### **The Optimal Penalty: Institutional Design**

If the cost of exercising the escape clause is too high, the gains from temporary defection and preserving one’s cooperative reputation are more than likely out-

23. Canavan and Rosendorff 1997.

24. Morrow, this volume.

weighed by the penalty associated with the use of the escape clause. In such circumstances, the cost of exercising the escape clause is too high, that is,  $k > K$ . Then, in any period where a large shock is experienced, the escape clause option is too expensive, and the system breaks down entirely. As a corollary to Proposition 1 above, the same equilibrium strategies in an environment where  $k > K$  lead to an equilibrium path in which  $P$  is played until  $B(\gamma) > k$ , in which case home plays  $D$  and the system breaks down. Over time, if the escape clause is too costly, the system breaks down with probability 1 (as long as the discount rate is not too high).

But this raises a question of implementation: when will countries be able to agree to escape clauses that do not lead to the breakdown of all cooperation? In particular, when will they be able to agree to impose a cost on themselves for using the escape clause, and when will this be credible? Furthermore, will they pick the optimal level of costs so that the optimal degree of cooperation is induced? To address this issue we model a first stage before the trade barrier setting game is played. In this stage the countries bargain with each other over the penalty they are willing to pay for invoking an escape clause. One can think of this as bargaining over the nature of the trade agreement itself. Much of the bargaining in trade talks concerns escape clauses and exceptions to the agreement rather than the general amount of liberalization. Thus making an agreement means agreeing on a value for the penalties that all countries will (voluntarily) pay to use an escape clause.

Therefore, we add a pregame negotiation phase over the size of  $k$ . We consider the symmetric case where both countries are identical. Each wants to choose a penalty that maximizes the value of playing the game. But the value of the game is the same for both players (they are identical), so they agree merely to the level of  $k$  that maximizes the value of the game.

**PROPOSITION 2.** *Let  $V_C$  and  $V_C^*$  be the present discounted expected value of the escape clause equilibrium for home and foreign, respectively. Then both countries agree on  $k^* = \arg \max_k (V_C + V_C^*)$  when  $k^* \leq K$ ; they agree on  $K$  otherwise.*

Larger distributional questions arise when the assumption of symmetry is relaxed. If one country has a greater capacity to absorb exogenous shocks, or alternatively is immune to capture by political interests, this country would prefer a larger value of  $k$ ; a country that is easily captured by special interests will instead prefer a smaller  $k$ . The outcome of this bargaining among asymmetric countries will have important consequences for the international institutions, but it is a subject that we leave for future consideration.

### *On the Design of Escape Clauses*

We have established that escape clause equilibria exist, and that for the escape clause to be exercised in equilibrium, it cannot be too expensive to adopt. This also points to an important trade-off in the design of international institutions between rigidity and stability. As the system becomes too rigid—or as  $k$  rises—it becomes

TABLE 1. *The trade-off between rigidity and stability*

<i>Size of penalty</i>	<i>Regime stability</i>	<i>Regime rigidity</i>
$k \leq K$	High	Low
$k > K$	Low	High

increasingly unstable (Table 1). At low values of  $k$ , the system is stable. For any value of the shocks, either pure cooperation or the escape clause is exercised; there is never any exit from the system and hence the regime is very stable. But this comes at a cost: At low values of  $k$ , the escape clause is cheap to adopt, leading to many periods in which defection is being tolerated in exchange for the benefits of long-term stability.

Instead, if the cost of exercising the escape clause is too high, it is never used, and as soon as the shocks become severe, the system breaks down and exit occurs. The regime is now too rigid and becomes unstable. It becomes clear then that the traditional Prisoners' Dilemma game without an escape clause is equivalent to this game with a large  $k$ : cooperation will break down at some point.

*COROLLARY 2. As the costs of using an escape clause rise, it will be used less frequently.*

### **Costs and Use of Escape Clauses: Some Empirical Implications and Evidence**

If we are right that governments rationally design escape clause mechanisms, we should see that variations in their cost lead to variations in their usage. Low-cost escape mechanisms should have much appeal; those with high costs should not. A good deal of evidence seems to suggest that this argument is valid. For instance, in U.S. trade law, the escape clause (Sec. 201) has been used far less often than have various other safeguard mechanisms. Wendy Hansen and Thomas Prusa show that the average number of escape clause cases filed has never gone above eleven a year, whereas for antidumping and countervailing duty cases the average reached a peak of ninety-two a year in the early 1980s.<sup>25</sup> Moreover, escape clause complaints have been decreasing steadily, with less than one a year filed in the early 1990s. In contrast, antidumping and countervailing duty cases have been growing over time. What accounts for this difference in usage?

We argue that it is the greater cost of invoking escape clauses that makes firms less likely to do so. Hansen and Prusa claim that the lower probability of success

25. Hansen and Prusa 1995, 299, tab. 1.

encourages firms to file antidumping and countervailing duty complaints instead. But our claim is that the lower probability of success results from the fact that escape clause actions when implemented cost the importing country more and thus make policymakers less likely to accept petitions for them. Thus firms see the mechanism as less successful and choose other means. The main reason they cost more is that exporters have a right to demand compensation for escape clause relief and, if it is not forthcoming, to retaliate. Compensation and retaliation create large domestic costs for governments, and thus they try to avoid such measures.

GATT also provides evidence that greater costs mean less use. Under GATT rules, exporters were entitled to compensation or retaliatory action if Article XIX, which involved the escape clause, was invoked. Moreover, the standards of proof for “serious injury” caused by imports needed to invoke the escape clause have been the highest of all. Among all the various safeguard means in GATT, Article XIX was among the least used. It was invoked only 150 times from 1950 to 1994. And its use has declined over time: 3.9 times a year from 1950 to 1984, and 3.2 times a year from 1985 to 1994. In contrast, the antidumping clause is much more frequently invoked: over two thousand times since 1970 alone.<sup>26</sup>

Moreover, scholars have noted that the costliness of escape clause actions has led to the proliferation of so-called voluntary export restraints. As Jeffrey Schott states,

Most major trading countries, however, have been deterred from invoking Article XIX less by its requirements than by the availability of less onerous and more flexible channels of protection. These have included coercing trading partners to accept VERs [voluntary export restraints] and other so-called gray area measures, as well as frequent recourse to unilateral relief actions under Article VI (i.e., antidumping and countervailing duties).<sup>27</sup>

Voluntary export restraints are less costly to use than escape clauses because they do not assume compensation or allow retaliation from the affected exporter. But an importing country using them may incur costs. Unlike a tariff or quota, which provides rents for the importing country, a voluntary export restraint transfers those rents to the exporter. As Bernard M. Hoekman and Michael Kostecki maintain,

Affected exporters tended to accept VERs because they were better than the alternative—often an AD [antidumping] duty—as they allowed them to capture part of the rent that was created. Instead of being confronted with a tariff, the revenue of which is captured by the levying government, a VER involves voluntary cut-backs by exporters in their supplies to a market. This reduction in supply will raise prices—assuming that others do not take up the slack. Exporters therefore get more per unit sold than they would under an equivalent tariff. . . . The key point to remember about VERs is that they imply some direct compensation of affected exporters and selectively target exporters. Thus

26. Hoekman and Kostecki 1995.

27. Schott 1994, 94.

they practically meet GATT-1947's compensation requirement, while allowing for circumvention of its nondiscrimination requirement.<sup>28</sup>

Hence, voluntary export restraints were preferred to escape clause actions because they were less expensive to employ, but even they imposed costs on the importing country.<sup>29</sup>

Interestingly, GATT recognized that the costliness of using the escape clause was hurting the system and pushing states to develop other means, such as voluntary export restraints, to deal with domestic pressures. Many GATT officials found other safeguard remedies—such as antidumping, voluntary export restraints, and countervailing duties—very undesirable. They preferred that countries use the escape clause mechanism. But they also realized that this process was too costly and thus underused.

In the Uruguay Round, GATT officials made several changes to reduce the costs of the escape clause relative to other safeguards. First, they banned the use of voluntary export restraints in the agreement on safeguards.<sup>30</sup> This in effect raised the costs of such measures. Second, they decided that it was necessary to reduce the costs of the escape clause option. So they proposed, and countries agreed, that one way to do this was to eliminate the right of retaliation. In the WTO, countries that use the escape clause no longer have to pay compensation and the injured exporters can no longer legally retaliate for the first three years of its use.<sup>31</sup> As Hoekman and Kostecki note, “by the time of the Uruguay round the major objective of ‘target’ countries was to constrain the use of AD and VERs and assert the dominance of Article XIX in safeguard cases . . . Two options were available: either to tighten the discipline on the use of AD, or to reduce the disincentives to use Article XIX. Both approaches were pursued.”<sup>32</sup> Lowering the costs of using the escape clause was therefore seen as a key way to shift countries away from using alternative safeguards like antidumping and countervailing duties, and toward using more escape clause actions. This seems to provide some evidence that leaders do indeed rationally design international agreements.

In the international monetary arena, the costs of exercising relief have varied both across institutions and within institutions over time. Again, one could argue that these variations are the rational responses of political leaders to the problems associated in part with domestic uncertainty. The Bretton Woods system's vagueness about the conditions under which a devaluation could occur meant that it was frequently appealed to, and effective cooperation was limited. The European Payments Union and the Exchange Rate Mechanism both were more specific about

28. Hoekman and Kostecki 1995, 168–69.

29. Similarly, Rosendorff establishes that voluntary export restraints are preferred by policymakers to antidumping duties because they generate higher electoral returns at lower costs when policymakers experience political pressures for protection. Rosendorff 1996.

30. Schott 1994, 94.

31. See Preeg 1995, 100–101; and Schott 1994, 94–97.

32. Hoekman and Kostecki 1995, 169.

the terms of realignments; moreover, the Exchange Rate Mechanism became increasingly more restrictive about the conditions under which escape was possible as the system moved toward monetary union, and accordingly less tolerant of realignments. Consequently, the system became somewhat more rigid and less flexible, leading to more periods of instability and exit, as happened in Britain and Italy in 1992.<sup>33</sup>

### Fearon's Dynamic

The escape clause adds flexibility to an agreement that might be difficult to sustain in the presence of uncertainty. Hence, bargainers are not stuck in a commitment to a distributional outcome for the infinite horizon, thereby making initial bargains easier to strike. This result lies in contrast to Fearon's concern that infinite horizon models with large discount factors make agreements difficult to strike.

*COROLLARY 3. Agreements should be easier to achieve when escape clauses are included than otherwise.*

As many analysts have noted about GATT, signing would have been impossible for many countries had it not included various safeguards. John Gerard Ruggie, for example, has argued that all of the international economic agreements, or regimes, negotiated after World War II had to embody the norms of "embedded liberalism," by which he meant that they had to combine multilateralism with the requirements of domestic stability.<sup>34</sup> Domestic safeguards that allowed countries to protect their economies were thus essential parts of this norm in both the trade and monetary areas. Without such safeguards, countries would never have signed the trade and monetary agreements.

Moreover, Hoekman and Kostecki claim that "political realities often dictate that there be a mechanism allowing for the temporary reimposition of protection in instances where competition from imports proves to be too fierce to allow the restructuring process to be socially sustainable. Indeed, a safeguard mechanism is likely to be a pre-condition for far-reaching liberalization to be politically feasible."<sup>35</sup> Or as Alan Sykes has shown, "when self-interested political officials must decide whether to make trade concessions under conditions of uncertainty about their political consequences, the knowledge that those concessions are in fact 'escapable' facilitates initial trade concessions."<sup>36</sup> Following Kenneth Dam,<sup>37</sup> Sykes maintains that

unanticipated changes in economic conditions may create circumstances in which the political rewards to an increase in protection (or the political costs of

33. Canavan and Rosendorff 1997.

34. Ruggie 1982.

35. Hoekman and Kostecki 1995, 191.

36. Sykes 1991, 259.

37. Dam 1970, 99.



an irrevocable commitment to reduce protection) are great. Consequently, in the absence of an escape clause, trade negotiators may decline to make certain reciprocal concessions for fear of adverse political consequences in the future. But, with an escape clause in place the negotiators will agree on a greater number of reciprocal concessions, knowing that those concessions can be avoided later if political conditions so dictate.<sup>38</sup>

Our point is that the inclusion of escape clauses should make reaching an initial agreement easier.

This argument shares much with the theory of efficient breach used in legal theory. This theory advances the idea that “there are circumstances where breach of contract is more efficient than performance and that the law ought to facilitate breach in such circumstances.”<sup>39</sup> In order to do so, there must be mechanisms that can determine and compel payment of the appropriate levels of damages for such breach. Jeffrey L. Dunoff and Joel P. Trachtman also note that “entry into contract may be facilitated by the understanding of parties that breach may be permitted under certain circumstances.”<sup>40</sup> They point out that the WTO’s safeguard system and its notion of compensation or retaliation provides just such a mechanism for efficient breach.

An alternative flexibility-enhancing device is to build into any agreement the opportunity for regular renegotiation, as in GATT, or the International Coffee Agreement.<sup>41</sup> John E. Richards notes that the International Air Transport Association, an airfare-setting cartel, allowed suspension of current agreements for the one-year period in which renegotiation occurred.<sup>42</sup> In the same way that an escape clause adds the necessary flexibility and does not fix the distributional impact immutably, Barbara Koremenos suggests that allowing for renegotiation and finite duration reduces the distributional impact of the agreement, making bargaining over an initial agreement easier, without reducing the effect of the “shadow of the future” in enforcing the agreement.<sup>43</sup> The escape clause, like the opportunity for renegotiation, reduces the effects of Fearon’s dynamic. We do think, however, as does Sykes, that renegotiation of an entire agreement is likely to be the most costly means by far and to have a lower probability of success than will the mere inclusion of escape clauses in the original agreement.<sup>44</sup>

There is a second reason escape clauses may diminish Fearon’s dynamic. In our model the countries are in a position similar to John Rawls’s “initial position,” where one is behind the veil of ignorance and cannot tell exactly how one will benefit (or lose) in the future from agreements made now.<sup>45</sup> Because shocks occur in each

38. Sykes 1991, 279.

39. Dunoff and Trachtman 1999, 24.

40. *Ibid.*, 26.

41. Koremenos, Lipson, and Snidal, this volume.

42. Richards, this volume.

43. Koremenos 1998.

44. Sykes 1991, 280.

45. Rawls 1971.

*future period that cannot be predicted beforehand, the players do not know the future distribution of gains and losses from the initial agreement with certainty. Hence this is likely to mitigate how hard they bargain in the first place. For these two reasons in our model, Fearon's argument may not hold: the length of the shadow of the future may play no role in affecting the bitterness of bargaining over the initial agreement. Moreover, including escape clauses may make both enforcement and distributive bargaining easier!*

## Conclusion

International institutions vary substantially. Their design reflects the rational calculations of, as well as the strategic interaction among, countries creating them. These different designs also have implications for the functioning of these institutions. International institutions matter, but so do their forms.

We have shown that *international institutions that include an escape clause can generate more durable and stable cooperative regimes. The escape clause itself is endogenous to the model: choosing a prohibitive cost for using the escape clause is equivalent to ruling it out of the institutional structure. Yet we have shown that in equilibrium the negotiating parties will adopt an escape clause with moderate costs when faced with domestic political uncertainty. Indeed, this particular institutional feature—the escape clause—is determined endogenously as an equilibrium outcome to the strategic game between the countries. Thus our model not only derives the rational form of an institution but also shows the impact of that institution once in place. We think future research should explore this result when more than two players are involved and/or when the countries are assumed to be different, such as Giovanni Maggi does.*<sup>46</sup>

We make three claims here. One is that escape clauses are an efficient equilibrium under conditions of domestic political uncertainty. When political leaders cannot foresee the extent of future domestic demands for protection, such clauses provide the flexibility that allows them to accept an international agreement liberalizing trade. One testable proposition is that the greater the domestic uncertainty that political leaders face about their ability to maintain domestic compliance with international agreements, the more likely leaders are to negotiate agreements that contain escape clauses. In issue-areas where governments face less uncertainty about future domestic pressures to comply, they are less likely to design such safeguard measures. This may help account for the differences between international trade agreements, where escape clauses are prevalent, and arms control agreements, where they appear to be less salient. Another testable proposition would involve examining whether certain domestic political institutions that reduce domestic uncertainty reduce the incentives for leaders in these countries to pursue escape clauses. Our model's results thus support Rational Design conjecture F1, FLEXIBILITY

46. Maggi 1999.

increases with UNCERTAINTY ABOUT THE STATE OF THE WORLD. Future research to examine the empirical hypotheses we have outlined would lend credence to this conjecture.

Our second claim is that escape clauses are useful and efficient only when they impose some kind of cost for their use; that is, importing countries must pay for invoking them or else they will be invoked all the time, thus vitiating the agreement. Paying the cost signals an intention to comply with the agreement in the future. Hence, another testable proposition is that the different costs of different escape clause measures should affect their use. Less costly measures for the importer should be used more often. We assume that governments understand this dynamic. And we anticipate that the architects of international agreements will rationally design such agreements so that the types of escape clauses they most desire will be neither too cheap (encouraging frequent use) nor too expensive (discouraging their use altogether). Furthermore, since paying the penalty is self-enforcing, we expect that the institution's role will be less that of an enforcer making countries pay this penalty and more that of an information provider telling others that the penalty has been paid. Thus we expect that countries will pay penalties, while looking to international institutions for information on whether others have done the same. The role of international institutions here is to provide a particular kind of information about other states' behavior. Again, this is a testable proposition that might warrant future attention.

Our third claim is that escape clauses make initial agreements easier to reach. Fearon's dynamic breaks down; the flexibility provided by escape clauses ensures that the division of the long-term gains from the agreement is not immutable. This result of our model provides theoretical support for Rational Design conjecture F2, FLEXIBILITY increases with DISTRIBUTION problems. Our argument also shares much with the legal theory of efficient breach, where the inclusion of measures allowing parties to later breach a contract may make initial agreement on a contract more likely. Indeed, we claim that without escape clauses of some sort many international agreements would never be politically viable for political leaders to sign in the first place. And this explains why rational political leaders design flexibility into their international commitments when they are uncertain about the future.

Here we have investigated whether the inclusion of escape clauses in international agreements could be a rational response of political leaders to their domestic problems, especially to unanticipated domestic political pressures. These escape mechanisms help political leaders to maintain international cooperation without sacrificing their domestic political positions; they thus reduce the costly, contradictory pressures that can emanate from domestic and international politics, helping to make international cooperation more compatible with domestic political success. As we have argued elsewhere,<sup>47</sup> such solutions to the two-level game faced by political leaders are essential for successful international cooperation. Rationally designing

47. See Milner and Rosendorff 1996; and Milner 1997.

flexibility into international agreements thus is important for political leaders when faced with domestic uncertainty and international distributional problems. The likelihood of and the probability of success of international institutions thus depends on their internal design, as well as other factors.

**Appendix**

DEFINITION 2. Let  $N(\gamma, g) - S(\gamma, g) \equiv A(\gamma, g)$ .

DEFINITION 3. Denote

$$I(\gamma) = \int_g I(\gamma, g)d\Phi \quad \text{and} \quad I = \int_\gamma \int_g I(\gamma, g)d\Phi d\Phi$$

for any function  $I = A, B, P, D, N, S$ .

DEFINITION 4. Let  $p = Pr(P|cooperation)$ .

That is,  $p$  is the probability of playing  $P$  given that  $P$  or  $EC$  is to be played. Consider the current period in which nature has drawn  $(\gamma, g)$ . Home knows  $\gamma$  but is unsure of  $g$ , and hence is unsure of the behavior of the foreign country. Since the countries are symmetric, we know that foreign plays  $P^*$  with probability  $p$  and plays  $EC^*$  with probability  $1 - p$ . If home plays  $P$ , then home earns in that period  $pP(\gamma) + (1 - p)S(\gamma)$ ; whereas if home plays  $EC$ , home earns  $p(D(\gamma) - k) + (1 - p)(N(\gamma) - k)$ . Then  $P$  is played if  $pP(\gamma) + (1 - p)S(\gamma) > pD(\gamma) + (1 - p)N(\gamma) - k$ , that is, if  $k > pB(\gamma) + (1 - p)A(\gamma)$ . Hence,  $p = Pr(k > pB(\gamma) + (1 - p)A(\gamma))$ .

LEMMA 2. For any  $k$ , the function  $\Lambda(p; k) = Pr(k > pB(\gamma) + (1 - p)A(\gamma))$  has a fixed point,  $p = \Lambda(p; k)$ .

*Proof.* For any  $k$ ,  $\Lambda$  is a continuous function of  $p$  mapping from  $[0, 1]$  into  $[0, 1]$ . Now  $[0, 1]$  is a compact, convex set. Therefore, a fixed point exists by Brouwer's Fixed Point theorem.

Lemma 2 implies that there exists a distribution function  $\Gamma$  such that  $p = \Gamma(k)$ .

LEMMA 3.  $\Gamma(0) = 0$  and  $\lim_{k \rightarrow \infty} \Gamma(k) = 1$ .

*Proof.*  $\Gamma(0) = Pr(0 > A(\gamma)) = 0$ , since  $A(\gamma) > 0$  for all  $\gamma$ ;  $\lim_{k \rightarrow \infty} \Gamma(k) = \lim_{k \rightarrow \infty} Pr(k > pB(\gamma) + (1 - p)A(\gamma)) \rightarrow 1$ , since  $B(\gamma), A(\gamma)$  are finite for all  $\gamma$  and  $p \in [0, 1] \forall k$ , since  $p$  is a distribution function.

PROOF OF PROPOSITION 1. A pair of escape clause strategies is a Nash equilibrium.

The expected current period return from defection at home is  $D(\gamma)$ , and hence the gains from defection are  $D(\gamma) - \max(P(\gamma), D(\gamma) - k) = \min(B(\gamma), k)$ . Consider the event that a deviation has been observed in some period. From then on, the one-shot Nash strategies are

played, yielding the Nash payoff (in expectation, since the draws in the future periods are unknown, forever). That is, the aggregate Nash payoff (starting in the next period) is

$$V_D = \frac{\delta}{1 - \delta} N.$$

What is the forgone cooperative aggregate payoff? If cooperation occurred in the last period, in the next each player has the option of cooperating again or defecting. The value of the game in a cooperative phase is the earnings from the play in that period plus the continuation value,  $V = p[p(P + \delta V) + (1 - p)(S + \delta V)] + (1 - p)p(D - k + \delta V) + (1 - p) \times (N - k + \delta V)$ .

Solving we have

$$\begin{aligned} V &= \frac{1}{1 - \delta} (p^2 P + p(1 - p)(S + D) + (1 - p)^2 N - k(1 - p)) \\ &= \frac{1}{1 - \delta} (p^2(A - B) + p(-A + D - N) + N - k(1 - p)). \end{aligned}$$

Hence,

$$V - V_D = \frac{1}{1 - \delta} (p^2(A - B) + p(D - N - A + k) + N(1 - \delta) - k).$$

Recall that  $p = \Gamma(k)$ . The no defect condition in any period is therefore

$$\min(B(\gamma), k) < \frac{1}{1 - \delta} ((\Gamma(k))^2(A - B) + \Gamma(k)(D - N - A + k) + N(1 - \delta) - k).$$

Let

$$Z(k) \equiv \frac{1}{1 - \delta} ((\Gamma(k))^2(A - B) + \Gamma(k)(D - N - A + k) + N(1 - \delta) - k),$$

and define  $K$  to be a fixed point of  $Z(k)$ , that is,  $Z(K) = K$ . Setting  $z(k) = Z(k) - k$ , we have

$$z(k) \equiv \frac{1}{1 - \delta} ((\Gamma(k))^2(A - B) + \Gamma(k)(D - N - A + k) + N(1 - \delta) - k(2 - \delta)).$$

Now  $z(0) = N > 0$  and as  $k \rightarrow \infty$ ,  $\Gamma(k) \rightarrow 1$ , and  $z(k) \rightarrow -\infty < 0$  from Lemma 2. Then we have a nondegenerate fixed point by the intermediate value theorem. Then  $K$  is the upper bound on any penalty in order to invoke *EC*, and home plays *P* if  $B(\gamma) < k < K$ ; plays *EC* if  $k \leq B(\gamma) \leq K$ ; and plays *D* if both  $B(\gamma), k > K$ . Hence, a pair of escape clause strategies is an equilibrium.

PROOF OF PROPOSITION 2. Let  $k^*$  satisfy

$$k^* = \frac{1 - \Gamma(k^*)}{\Gamma'(k^*)} - 2\Gamma(k^*)(A - B) - (D - N - A),$$

then both countries agree on  $k^*$  when  $k^* \leq K$  and agree on  $K$  otherwise.

The value of the game to either player in which an escape clause equilibrium is played is

$$V(k) = \frac{1}{1 - \delta} ((\Gamma(k))^2(N - S - D + P) + \Gamma(k)(S + D - 2N + k) + N - k)$$

when  $k < K$ . What value of  $k$  maximizes this value? We solve  $k^* = \arg \max_k V(k)$ . The first order condition  $V'(k^*) = 0$  yields

$$k^* = \frac{1 - \Gamma(k^*)}{\Gamma'(k^*)} - 2\Gamma(k^*)(N - S - D + P) - (S + D - 2N).$$

Checking the second order condition, note that  $V''(k^*) < 0$  iff

$$\Gamma''(k^*) < (1 + (A - B)\Gamma'(k^*)) \frac{-2(\Gamma'(k^*))^2}{1 - \Gamma(k^*)}.$$

A sufficient condition for this to hold is that  $\Gamma(\cdot)$  has an increasing hazard rate, and  $A - B > 0$ . Moreover, we know that at  $k = K$ , each player is indifferent between exercising the escape clause and defecting permanently. If  $k^* > K$ , then  $V(k^*) < V(K)$ , implying the optimal choice of penalty is  $K$ .

PROOF OF COROLLARY 1. An agreement with an escape clause Pareto dominates one without in the presence of political uncertainty.

This follows from the previous proposition. Any escape clause game with  $k > K$  is equivalent to a game without an escape clause. This is because if  $k > K$ , the escape clause is never exercised, and at some point defection occurs (unless the discount rates are very high). However, countries optimally choose  $k \leq K$ ; hence, an agreement with an escape clause dominates one without.

PROOF OF COROLLARY 2. As the costs of using an escape clause rise, it will be used less frequently.

In any escape equilibrium, the probability that the escape clause is used is  $1 - p = 1 - \Gamma(k)$ . As  $k$  rises,  $1 - \Gamma(k)$  falls, reducing the frequency with which the escape clause is exercised.

PROOF OF COROLLARY 3. Agreements should be easier to achieve when escape clauses are included than otherwise.

With escape clauses, true cooperation occurs as long as  $k \leq K$ ; there is no restriction on the discount factor  $\delta$ . That is, given any discount factor  $\delta$ , there exists a penalty  $k \leq K$  such that an escape clause equilibrium exists. In the standard Prisoners' Dilemma in the face of uncertainty, cooperation occurs whenever

$$\delta > \frac{\max_{\gamma} B(\gamma)}{P - N + \max_{\gamma} B(\gamma)}.$$

Hence, the set of discount factors under which the standard Prisoners' Dilemma under uncertainty can support a cooperative equilibrium is

$$\left[ \frac{\max_{\gamma} B(\gamma)}{P - N + \max_{\gamma} B(\gamma)}, 1 \right] \subset (0, 1],$$

the set of discount factors under which an escape clause equilibrium exists. Hence, if we were to draw a discount factor at random, we are more likely to be able to support an escape clause equilibrium than a cooperative equilibrium in a game without an escape clause.