

# International Theory

<http://journals.cambridge.org/INT>

Additional services for *International Theory*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



---

## The dilemma of informal governance with outside option as solution

Daniel Verdier

International Theory / Volume 7 / Issue 01 / March 2015, pp 195 - 229

DOI: 10.1017/S1752971914000335, Published online: 09 February 2015

**Link to this article:** [http://journals.cambridge.org/abstract\\_S1752971914000335](http://journals.cambridge.org/abstract_S1752971914000335)

### How to cite this article:

Daniel Verdier (2015). The dilemma of informal governance with outside option as solution. *International Theory*, 7, pp 195-229 doi:10.1017/S1752971914000335

**Request Permissions :** [Click here](#)

# The dilemma of informal governance with outside option as solution

DANIEL VERDIER

*Department of Political Science, Ohio State University, Columbus, OH, USA*

E-mail: verdier.2@osu.edu

The efficiency-oriented part of the literature on informal governance points to institutional costs as a reason for governments to prefer to cooperate with each other through commitments that are not binding. Left unexplained is what I call the dilemma of informal governance: how informal governance copes with the problem of cheating, to which formal governance has traditionally provided the solution. I show that like-mindedness, the current solution to the dilemma, is convincing but underspecified. Working from a model of governance encompassing the three time-honored dimensions of obligation, precision, and delegation, I analytically explore two other solutions, one that fails, information transmission, another that works, outside option, which I borrow from the power-oriented part of the literature on informal governance. A key finding is that informal governance, despite being neither self-enforceable nor informative, is sustainable for mild Prisoner's Dilemma (PD) types in the presence of outside options. I illustrate the model findings by tracing an historical correlation between power polarization and formalism in the design of security regimes.

**Keywords:** informal governance; security regimes; outside option; obligation; precision; delegation

After decades of exclusive concern for treaties and international organizations, the fields of international law and international institutions have embraced a less formal type of governance, called 'informal governance'. While there are several meanings of informal governance, I borrow mine from Abbott *et al.* (2000, 404): informal governance is a weaker version of legalization, scoring low on obligation, precision, and delegation.

The current popularity of informal governance, according to Slaughter (2004, 8) citing Keohane (2001), reflects the emergence in an ever more globalized world of the so-called 'governance dilemma', by which 'we need more government on a global and regional scale, but we do not want the centralization...' But informal governance, in turn, is pregnant with its own dilemma: How, in the absence of binding commitments, do the parties manage to overcome the risk of cheating – one side opportunistically

suspending cooperation while claiming changed circumstances? Formal governance may have its flaws, but among the reasons for its existence is a well-recognized ability to confront cheating.

There are two approaches to informal governance. A first, pronging efficiency, points to the institutional costs of formal treaty making as reasons for embracing informality. Yet, we are still in need of a general explanation for how it can be made sustainable. Like-mindedness is invoked here and there to eliminate the risk of cheating, but in unspecified dosage and at the peril of eliminating the need for cooperation in the first place. A second line of explanation, emphasizing power and mostly developed by Stone (2011), dismisses the dilemma but it does not provide a theory of informal governance that is generalizable to countries with comparable power.

I use a game between countries of comparable power to analyze the dilemma and look for a solution to it. I model the three substantive components of governance – obligation, precision, and delegation. I first question the common finding that uncertainty of the future provides a rationale for informal governance. I then look for workable solutions to the cheating problem that is endemic to informal governance. Although this problem can be partially solved by restricting membership to harmony types, defined as countries with a dominant strategy of cooperation, this solution is impractical in the context of private information. If a harmony type inadvertently enters into an interaction with a PD type, which has a dominant strategy of non-cooperation, the former risks being exploited for ever by the latter. The risk of exploitation can only be reduced by eliciting information on partners' type – harmony vs. PD. But the need for information transmission is met by an impossibility of supplying it. The standard signaling devices available to formal partners – the spending of *ex-ante* costs or the making of a commitment that will generate costs in the future – are unavailable in informal governance, a format rightly celebrated for short-circuiting the need for a commitment and ridding chancelleries from the *ex-ante* negotiation and ratification costs that are characteristic of treaty making.

The solution, I argue, lies in the ability of a country that is engaged in an informal relationship to switch cooperating partner whenever it feels like it, thereby eliminating the risk of long-term exploitation by a non-cooperating partner and making complete information unnecessary. Informal governance cannot exist or, if it does, is not sustainable absent an outside option, defined as the availability, unparalleled under treaty making, of another possible partner to whom to turn in case of failure. I show that, under such circumstances, despite being neither self-enforceable nor informative, informal governance can be sustainable for mild PD types, thereby providing an analytical foundation to the notion of like-mindedness.

The empirical part of the paper considers security cooperation over the last two centuries. Assuming that great powers can avail themselves of

outside options when the system is depolarized, I show that periods of reduced polarization, like the Concert of Europe and the post-Cold-War era, exhibit a freer use of informal legal instruments than periods of rigid alignments.

The paper is thus organized into three parts: review of the literature, modeling of governance, and historical illustrations.

## Literature review

There are two distinctive rationalist approaches to the study of international regimes, one resting on the economic notion of efficiency, the other on the neorealist concept of power, and each one has something distinctive to say about informal governance.

Championed by Keohane (1984), the efficiency approach views regimes as possible solutions to the transaction costs incurred by states in a world of anarchy. Within that tradition, the current emphasis on informalism is a reaction against the earlier enthusiasm for formal institutions (see Keohane and Nye 1974; Dehousse 1997; Raustiala 2002–03; Slaughter 2004; Lipson 2005/2006; Kerwer 2006; Eilstrup-Sangiovanni 2009; Josselin 2009; Mosley 2009; Kahler and Lake 2009; Speyer 2009; Verdier 2009; Helleiner, Pagliari, and Zimmermann 2010). It partly reflects the study of transgovernmental networks, regimes that often are negotiated between lower government levels, typically a department or ministry or agency endowed with regulatory power, sometimes with a sprinkling of experts drawn from the business world. The agreements, dubbed ‘accords’, ‘memoranda of understanding’, ‘communiqués’, or ‘statements’ have no legal force and are often imprecise.

The reasons that have been offered for the informal backlash are what I shall refer to as ‘institutional costs’ so as not to confuse them with the transaction costs identified by the earlier regime literature. They are: high negotiating and ratifying costs, inflexibility in the face of uncertainty, the requisite of publicity, and slowness (see Lipson 1991; Abbott and Snidal 2000; Eilstrup-Sangiovanni 2009; Vabulas and Snidal 2013).

From the list of institutional costs incurred under formalism, typically follows a catalog of explanations for when states would prefer informalism. One claim that is most often put forward is that informalism allows members to deal with uncertainty with greater ease than formalism (see Lipson 1991, 518; Eilstrup-Sangiovanni 2009, 207; Guzman and Meyer 2010, 197–201; Shaffer and Pollack 2010, 789; Stone 2011; Vabulas and Snidal 2013; Kleine 2013). A second rationale is the existence in treaty making of negotiating and ratifying costs, which can be steep in the presence of potential domestic veto players (see Lipson 1991; Abbott and Snidal 2000;

Raustiala 2002–03, 6; Eilstrup-Sangiovanni 2009, 206, 209; Vabulas and Snidal 2013). The need for publicity and/or urgency merely exacerbates the latter problem (see Lipson 1991; Eilstrup-Sangiovanni 2009, 206, 209). Related to the negotiation-cost rationale is the widespread delegation by legislatures and executives of competencies to regulatory agencies in areas of growing international interdependence. In need to cooperate with foreign equivalents, yet without the legal authority to commit their respective governments, these agencies fall back on negotiating non-binding agreements (see Slaughter 2004).

Most authors working in the efficiency tradition agree that the main drawback of informal against formal governance is the lack of enforcement mechanism in the sense that there is no legal basis for retaliating against a country for cheating if this country did not legally commit itself not to cheat in the first place (see Verdier 2009, 115; Eilstrup-Sangiovanni 2009, 200; Shaffer and Pollack 2011, 1164). Compliance with binding agreements is generally higher than with non-binding agreements, although non-binding agreements may be deeper and more ambitious than binding ones (see Downs, Rocke, and Barsoom 1996; von Stein 2008; Victor 2011, 227–29; Johns 2012).

The efficiency literature thus leaves us with a trade-off or, what I call, the dilemma of informal governance: How can an arrangement that tolerates cheating be a solution to the high costs and inefficiencies of formalism, given that the reason why formalism has high costs and inefficiencies in the first place is to enable participants to deter the risk of cheating?

From the point of view of the earlier regime literature, the institutional costs incurred in treaty making were not mere inconvenience that governments had to cope with in order to build a state of law. They also delivered the credible commitments that made the law self-enforcing in the first place. The time that a government took to thoroughly consult with interested bureaucratic agencies, plus the resources that were spent extracting the consent of the legislature, had the effect of signaling to other signatories the extent of the government's will to honor the commitment (see Koremenos, Lipson, and Snidal 1999; Martin 2005). An informal agreement will never be as credible as a formal agreement.

Similarly, uncertainty of the future does make flexibility a desirable feature. Such is why after all many treaties contain an escape clause providing for momentary suspension or permanent renegotiation (see Downs and Rocke 1995: Ch. 4; Rosendorf and Milner 2001; Koremenos 2005). Escape clauses have the side effect of weakening the strength of each member government's commitment, but they may still deter cheating by inviting the other members' scrutiny (see Johns 2014). Informal governance, in contrast, frees a 'defecting' government from any formal monitoring; cooperation being voluntary to

begin with, it can be withdrawn at no cost and without notice. Informalism accommodates uncertainty at the cost of emptying the agreement from its *raison d'être*.

One way to solve the dilemma is to look for a facilitating condition that would mitigate cheating and thus lessen the need for high costs *ex ante* and rigidity *ex post*. The most commonly invoked fix along such lines is 'like-mindedness', by which a government cooperates only with governments' holding policy goals convergent with its own.<sup>1</sup> Like-mindedness may result from like convictions, shared norms of behavior, technical expertise, or any other form of 'social capital'.

Although intuitively attractive, the like-mindedness hypothesis needs further specification. Like-mindedness does not work as a dummy variable – the proposition that governments *i* and *j* are or are not like-minded – for like-mindedness then means harmony, which, by definition, needs no cooperation.<sup>2</sup> Like-mindedness may work as a continuous variable that stretches between the two extremities of harmony and, say, PD with extreme prejudice, but the cut point has not been located yet.

Furthermore, the argument that views like-mindedness as a solution to cheating under informal governance runs into a contradiction of its own making. Like-mindedness cannot be ascertained in a context of incomplete information, where types are private information, if signaling is not possible. Signaling requires an actually incurred cost, a feature that is more characteristic of treaty making than of informal governance.

The power approach is an alternative to the efficiency approach. It builds on Krasner's (1991) proposition that international institutions are instruments of state power. A useful operationalization of the power variable is the notion of outside option, according to which the country that has the most capacity, in Gruber's (2000, 7) words, to 'go it alone' is the country with the most bargaining power. In his study of the UNSC, Voeten (2001, 845) argues that 'the ability to act outside helps the superpower to reach agreements that would be vetoed in the absence of outside options'. Finally, and this is where the power literature intersects with the topic of informal governance, in his study of the IMF and the WTO, Stone (2011) argues that the United States' outside option allows it, whenever its core interests are at stake, to override the organization's formal procedures and revert to self-serving informalism with the weaker members' blessing (see also Steinberg 2002). In sum, informalism works for he who has an option

<sup>1</sup> See Raustiala (2002, 24), Whytock (2005, 20, 30), Eilstrup-Sangiovanni (2009, 205), Kerwer (2006), and Kahler and Lake (2009, 273). For another line of argument, one emphasizing an informal delegation to a third party, the European Council Presidency, see Kleine (2013).

<sup>2</sup> See Keohane's (1984, 51) on harmony.

outside the formal setting. And informalism is defined as the possibility of suspending formal and more egalitarian procedures.

There is no dilemma of informal governance from the power perspective. If there is any dilemma at all, as Stone quips, it is why weaker powers are willing to join the organization in the first place, given that they are being used to share the burden and confer legitimacy. His answer, Voeten's too, is that the superpower is willing to 'lose' on issues that are secondary to its national interest.

Stone offers a parsimonious model that nicely captures the United States' current ambivalence toward international treaties and organizations. Yet, the argument is specific to a particular power structure. The argument equates informalism with opportunism and does not generalize to non-superpowers. It is also an argument about the optimal mix of formal and informal strategies within a single organization, not one about the choice between the two opposite forms of governance, which is how most of the literature approaches the question.

In the following, I pick up the loose threads on the efficiency side of the debate, while borrowing the notion of outside option from the power side. I show that in a world of countries of comparable power, mild PD types can be made to cooperate with harmony types in the absence of any legal framework provided that all protagonists have equal access to outside options. This new finding is derived from a model of formalism that builds on the legalization triad offered by Abbott *et al.* (2000): obligation, precision, and delegation.

### The model

Abbott *et al.* (2000) tell us that *obligation* means that the cooperating states are pledged to norms that are considered binding under the rules of international law. Any breach of obligation is susceptible to trigger an enforcement mechanism known as 'self-enforcement'. Unlike domestic agreements, international agreements are not enforceable by police or judge, but through the direct threat of mutual retribution between directly interested parties. Game theory shows that cooperation can be supported in indefinitely iterated mixed-sum games through the use of realistic punishment strategies when players have a sufficiently long time horizon – a fair characterization of trading nations. Equally important, are the strict conditions that empirical reality must satisfy in order to submit to the theoretical predictions. Notions of what constitutes cooperation and defection must be defined with *precision*, lest cooperation break down and mutual punishment ensue, while notions of what constitutes fair punishment must also be defined with precision, lest the players get caught into an endless

defection loop. Obligation and precision thus complement each other in the sense that obligation requires precision. Precision, in turn, often requires *delegation* to a third party. I begin with precision.

### *Precision*

There are good reasons why negotiators would let imprecision creep into their work. One form of imprecision, also known as '*constructive ambiguity*',<sup>3</sup> is deliberate. It might be employed in an agreement in order to bypass a contentious matter and make progress on other matters. It is nothing more than a temporary agreement to disagree and postpone hard choices to the future. Constructive ambiguity may also take the form of deliberate omission of relevant, yet disputed issues. The problem with papering over intractable issues, of course, is that the issue is likely to reappear at one point or another and force a breakdown of cooperation.

There is a second form of imprecision, known in economics as *incomplete contracting*. Due to the costs to write clauses and individuals' cognitive limitations, it is not efficient for parties to try to pre-specify all future contingencies in a comprehensive contract, but seems advisable, instead, to allow for some 'flexibility' (see Milgrom and Roberts 1992, 129–40). Extended to treaty making, this line of reasoning translates into the proposition that informal governance should perform better under uncertainty than rigid treaty making.

If precision is realized, self-enforcement is viable and the notion of obligation realistic. I offer a formal definition of imprecision on which I later build a model of cooperation.

Consider the following game of coordination between two countries,  $i$  and  $j$  (Table 1).

The optimal way for the countries to play this game is to play it repeatedly and somehow alternate between the strategy pairs  $\{x, y\}$  and  $\{y, x\}$ . To that effect they define an agreement that envisions  $n$  different possible contingencies, only one of which will actually happen at each round, randomly chosen by Nature.

Assume that the countries divide the  $n$  contingencies into three subsets (with the following agreed distribution denoted  $\Delta$ ). The first two subsets are precise in the sense that if Nature draws any component from the first subset, country  $i$  does  $x$  while country  $j$  does  $y$ , whereas if the component is drawn from the second subset, country  $i$  plays  $y$  while country  $j$  plays  $x$ . The third subset, in contrast, is imprecise in the sense that its components either are assigned to both countries (constructive ambiguity) or to neither

<sup>3</sup> A phrase attributed to Henry Kissinger.



Table 1. A game of coordination

Country <i>i</i>	Country <i>j</i>	
	<i>x</i>	<i>y</i>
<i>x</i>	0,0	<i>f</i> , $-b_j$
<i>y</i>	$-b_i$ , <i>f</i>	0,0

With  $f > b_k = b_{ij} > 0$ .

(incomplete contracting). Consequently, whenever Nature draws a contingency from this third subset, each country claims it as its own, with the result that they both play *x*.

As an illustration, consider the case of the Proliferation Security Initiative (PSI), a ‘Statement’ urging participating states to interdict the trade or shipment of weapons of mass destruction (WMD) and their delivery systems. As undersecretary of state John Bolton appositely put it, ‘there are essentially an infinite number of potential circumstances and variations and permutations where interdictions could take place’ (see Boese 2003). The text along with the shared interpretation reached in posterior meetings by the member parties contemplated a series of contingencies. Although the negotiations were kept secret, one can say with the advantage of hindsight that some of these contingencies were precisely defined: for instance, both Russia and the United States seemed to agree that Iran, North Korea, and Poland should be denied access to WMD, though with a different sense of urgency – higher for Washington than Moscow with respect to Iran and North Korea; vice versa with respect to Poland. Other contingencies were not defined with any precision: for instance, the cases of shipments to Pakistan or Assad’s Syria, or any type of move necessitating a potentially illegal interdiction, such as the boarding of a ship flying a national flag on the high seas.

To simplify notation, I assume that there is a continuum of possible contingencies distributed over the [0,1] interval. Given a contingency, it has an  $\alpha$  probability of being precise and a residual  $1 - \alpha$  probability of being imprecise; in turn, assuming it is precise, it has a  $\beta$  probability of having been assigned to country *i* and a residual  $1 - \beta$  probability of having been assigned to country *j*.

A fully precise agreement can be defined as a mathematical (perfect) partition of rights and obligations, that is, one with  $\alpha = 1$ . In contrast, maximum imprecision manifests itself in the form of  $\alpha = 0$ . Both  $\alpha$  and  $\beta$  are objectively known.<sup>4</sup>

<sup>4</sup> I assume that the probability distribution of the precision/imprecision variable,  $\alpha$  is known in advance. While this notion of uncertainty may seem a bit faint-hearted – it does not embrace

A precise distribution is efficient in the sense that the aggregate expected payoff for  $i$  and  $j$  is the highest achievable. The aggregate payoff for one round of cooperation, indeed, is equal to

$$\sum_{b=i,j} U_b(\cdot) = (\beta\alpha f + (1-\beta)\alpha(-b_i)) + (\beta\alpha(-b_j) + (1-\beta)\alpha f). \quad (1)$$

This expression simplifies to  $\alpha(f - (1-\beta)b_i - \beta b_j)$ , which is increasing in  $\alpha$ , measuring the degree of precision.

### *Delegation*

Where precision was not possible at the contracting time, because the negotiators differed or could not discern future circumstances, it can still be had *ex post* by delegating to a third party the responsibility of coordinating differing interpretations. Delegation clears ambiguity and provides completeness *ex post*.<sup>5</sup>

Delegation may not be possible. The contracting parties may not be able to find a mutually agreeable third party. For instance, the United States is not a member of the International Criminal Court not because there is no domestic support for punishing genocide and crimes against humanity among the American public, but out of fear that the court be turned against US military personnel out of pure mischief on the part of its overwhelming Third-World membership (see Galbraith 2003). Delegation may not be possible also because the parties are risk averse and balk at the possibility of an adverse ruling in the future. For instance, Washington fears that even an impartial court might turn against the United States following an unpredictable, yet hard to guard against, misuse of American power such as the bombing of civilians or the torture of prisoners. Such a condemnation would be unacceptable to a government committed to justice for all.

Technically speaking, introducing delegation in the agreement has two purposes. First, to reduce imprecision. If delegation to a third party was made, then it is as if the agreement had provided for a partition *ex ante*, except that, *ex ante*, the countries only had a probabilistic knowledge of how the third party might assign the imprecise cases. For instance, I assume that in the case where the countries fail to coordinate on  $\{x, y\}$  or  $\{y, x\}$ , but instead pool on  $x$ , they incur one round of zero payoffs and then expect, in the following round, the third party to rule in favor of  $i$  with  $\theta$  probability

Donald Rumsfeld's notion of 'unknown unknowns' – it corresponds to the stochastic notion of risk, which is the standard approach to uncertainty in the field; see Koremenos (2005).

<sup>5</sup> For evidence of a correlation between imprecision and delegation, see Koremenos (2008, 175). For a study of the trade dispute settlement mechanism as an exemplar of *post-facto* coordination, see Reinhardt (2001) and Johns (2012).

and in favor of  $j$  with  $1 - \theta$  probability. Hence, for country  $i$ , the delegation payoff is equal to

$$U_i(\text{delegation}) = \delta\tau(\theta f - (1 - \theta)b_i)$$

This delegation payoff yields an average payoff comparable to that of cooperation, yet discounted twice, first by a factor  $\delta$  to reflect the one-period delay caused by the *ex-ante* imprecision, second by a factor  $\tau$  (with  $0 \leq \tau < 1$ ) to reflect the politics of delegation – the fact that the winner of the draw earns more (and the loser loses less) from spontaneous coordination than from coordination mediated by a third party.<sup>6</sup>

The second purpose of delegation is to redress defection. Hence, if a country plays  $x$  when it was expected to play  $y$ , the case is submitted to third-party arbitration, which, I assume, will automatically rule against the defector. The cooperation payoffs are also delayed by one period and discounted by  $\tau$ .

### Obligation

I model obligation as an indefinitely repeated game in which each country plays a well-defined punishment strategy. This section, which may be skipped by the non-technical reader, merely restates the well-known, benchmark finding that mutual cooperation is possible in a formal setting.

To cooperate in this game is to apply the above-defined agreed distribution  $\Delta$ . If both countries do so repeatedly, they can expect to receive, for country  $i$ ,  $U_i(\Delta|\Delta) = \alpha\beta(f + \delta C_i) + \alpha(1 - \beta)(-b_i + \delta C_i) + (1 - \alpha)\delta(\theta(\tau f + \delta C_i) + (1 - \theta)(-\tau b_i + \delta C_i))$ , and for country  $j$ ,  $U_j(\Delta|\Delta) = \alpha(1 - \beta)(f + \delta C_j) + \alpha\beta(-b_j + \delta C_j) + (1 - \alpha)\delta((1 - \theta)(\tau f + \delta C_j) + \theta(-\tau b_j + \delta C_j))$ .

In both expressions, the first term is the payoff in case the country wins the draw, the second term is the payoff if the country loses the draw, and the last term is the delegation payoff in the event of imprecision. The delegation payoff merely is another draw, discounted twice, a first time by  $\delta$  because it is put off by one period, and a second time by  $\tau$  for the political reasons already mentioned.  $C_{k=i,j}$  stands for the continuation value of the repeated game.

Defection (call it  $\Xi$ ) may take several forms, but the most obvious one consists in playing  $x$  under any circumstances – this way, the country cashes in  $f$  part of the time if the other country cooperates without ever disbursing  $b_{k=i,j}$ . I also assume that defection implies refusing to enforce a third party's negative ruling. Defection triggers punishment, which I assume to take the simple and realistic form of a one round of mutual retaliation  $\{x, x\}$ .

<sup>6</sup> Another reason for introducing discount factor  $\tau$  is to make sure that precision *ex post* can never be as good as precision *ex ante* and that cooperation *ex post* can never be as good as cooperation *ex ante*.

Denote by  $\sigma_F$  the cooperation strategy entailing playing distribution  $\Delta$  on the equilibrium path and  $x$  once off the equilibrium path before resuming playing  $\Delta$ . The defection payoff for country  $i$ , assuming  $j$  cooperates, is  $U_i(\Xi|\sigma_F) = \alpha\beta(f + \delta D_i) + \alpha(1 - \beta)\delta^2 D_i + (1 - \alpha)\delta(\theta(\tau f + \delta D_i) + (1 - \theta)\delta^2 D_i)$ .

The first expression is the payoff for winning the draw. The second is the payoff for losing it, refusing to cooperate, and going into a one round of mutual defection before resuming cooperation. The third is a delayed, delegated combination of the first two with the  $\tau$  discount affecting the winning payoff.  $D^i$  is the continuation value.

A systematic definition of the formal governance strategy  $\sigma_F$  is provided below.

**Definition 1** [*formal governance strategy  $\sigma_F$  with one-round punishment*] For player  $i$ , apply agreed distribution  $\Delta$ . This means: (a) play  $x$  when Nature or third party draws a contingency favorable to  $i$  and (b) play  $y$  when Nature or third party draws a contingency favorable to  $j$ . Moreover, whenever  $\{x, x\}$  or  $\{y, y\}$  was played in the prior period, play  $x$  for one round and then return to applying the agreed distribution. For player  $j$ , merely substitute  $y$  for  $x$  in (a) and  $x$  for  $y$  in (b).<sup>7</sup>

The stage game has the payoff structure of a PD, insuring that a sufficiently low cost for cooperation  $b_{k=i,j}$  for each player delivers repeated cooperation on the equilibrium path. Hence, the following result:

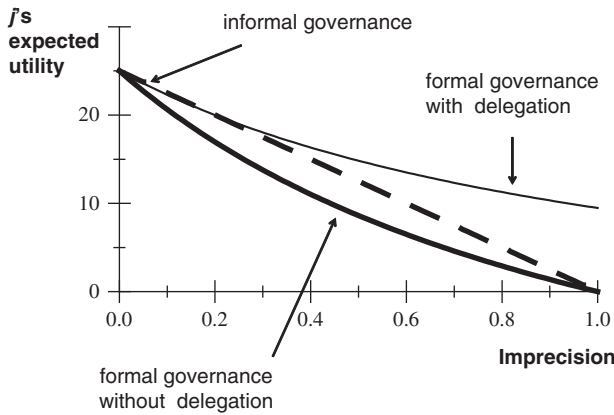
**Proposition 1** [*formal agreement with one-round punishment*] The strategy pair  $(\sigma_F|\sigma_F)$  of Definition 1 delivers a subgame perfect equilibrium for values of  $b_j$  and  $b_i$  that are below a threshold value that is strictly positive.

Proof and details are supplied in the appendix. Comparative statics shows that the highest value of the cost of cooperation  $b_{k=i,j}$  for which formal cooperation is sustainable is increasing in precision  $\alpha$  and discount factors  $\delta$  and  $\tau$ . This means that cooperation is made easier by more precision, a longer time horizon, and less politicization.

## Uncertainty of the future

Is uncertainty of the future a rationale for informal governance as it is argued in the literature (see Lipson 1991, 518; Eilstrup-Sangiovanni 2009, 207; Guzman and Meyer 2010, 197–201; Shaffer and Pollack 2010, 789;

<sup>7</sup> Note for clarification purposes that there are two cases in which  $\{x, x\}$  may occur: the first, compatible with strategy  $\sigma_F$ , is after Nature drew an imprecise contingency; the second, a defection from  $\sigma_F$ , is when a player defected by refusing to reciprocate when called for – for instance,  $i$  played  $x$  to avoid the  $-b_i$  payoff after Nature drew a contingency favorable to  $j$ .



**Figure 1** Impact of imprecision on cooperation payoff, with  $f = 4$ ,  $b_j = 1$ ,  $\beta = 0.5$ ,  $\theta = 0.5$ ,  $\tau = 0.8$ , and  $\delta = 0.9$  (complete information).

Stone 2011; Vabulas and Snidal 2013; Kleine 2013)? I show that, within the restricted compass of the present game (complete information, no outside option), the conditions for this result to obtain are unduly restrictive: the players must be of the harmony type and delegation must be unworkable.

The model above offers a ready-made way of quantifying uncertainty of the future and of simulating its toll on the act of cooperation. Uncertainty translates into imprecision and is measurable by  $1 - \alpha$ . I show in Figure 1 a simulation of how each player’s expected utility for formal cooperation is affected by a change in imprecision (attributable to a change in uncertainty). The relation (thick solid line) is clearly negative, as less precision means less utility. The reason is that every time Nature draws a contingency from the imprecise subset, the players punish each other.

One potential way of mitigating the negative impact of imprecision is to give up on punishment, and thus on obligation in the first place. Such is, in essence, what is implied by the shift from formal to informal governance: imprecision and defection are not tantamount to a breach of obligation and thus cannot be sanctioned by punishment. Basically, the parties are not expected to play  $\{x, x\}$  for one round in response to  $\{x, x\}$  or  $\{y, y\}$ . I define an informal governance strategy below. Definition 2 is like Definition 1 up to the clause on punishment, omitted here.

**Definition 2** [*informal governance strategy* ( $\sigma_i$ )] Apply the agreed distribution. This means for player  $i$ : (a) play  $x$  when Nature or third party draws a contingency favorable to  $i$  and (b) play  $y$  when Nature or third party draws a contingency favorable to  $j$ . For player  $j$ , merely substitute  $y$  for  $x$  in (a) and  $x$  for  $y$  in (b).

I use the same simulation parameters to graph the expected utility yielded by the informal governance strategy in Figure 1. The new (dashed) curve is situated above the formal (thick solid) curve at all levels of imprecision.<sup>8</sup>

However, this line of reasoning presents two caveats that call its usefulness into question. First, informal cooperation cannot sustain cooperation with PD types if defection cannot be deterred by some credible punishment threat. This is the gist of the dilemma of informal governance that was presented above: informal governance is simply not sustainable among players with PD payoffs, but can only be sustained among players with so-called ‘harmony’ payoffs, for whom  $b_{k=i,j}$  is negative.<sup>9</sup>

Perhaps, the notion of harmony is what students of informal governance allude to when they describe the members of an informal regime as being ‘like-minded’ (see footnote 1). Admittedly, this is a weak result, one that should raise a red flag among students of regimes. As Keohane (1984, 51) once put it, cooperation is not needed if countries live in harmony. I shall eventually improve on this result but not before having shown that the prospects for informal governance get even more elusive once one abandons the unrealistic assumption of complete information.

Second, informal governance is not the only format that can deal with uncertainty; formal delegation, whenever possible, is also an efficient technique. I graph a simulated value for the expected utility yielded by delegation in Figure 1. This new (thin solid) curve tops the informal (dashed) curve at high levels of imprecision.

The reader might object that my simulation did not factor in any fixed cost for the establishment of delegation. Remedying this omission would bring the curve labeled ‘formal governance with delegation’ down, perhaps below the rival curve marked ‘informal governance’. My counter-objection is that, short of endogenizing this fixed cost, any value goes, making any conclusion *a priori* indeterminate. I endogenize the cost of formalism in the next section.

### Private information

I just showed that informal governance is unsustainable between PD-type players, because it leads to mutual cheating. It is also unsustainable between harmony-type and PD-type players, as the latter takes advantage of

<sup>8</sup> Note that the two curves meet at both extremes: in the case where the agreement is fully precise, because there is no punishment, and in the case where the agreement is one-hundred percent ambiguous, because there are no positive payoffs. Also note that the gap between the two curves would deepen if the punishment protocol, instead of being a one-round defection, were harsher.

<sup>9</sup> For the record, the one-round-punishment strategy of Definition 1 is subgame perfect between harmony types or between a harmony and a PD type.

the former. Clearly, the risk of exploitation puts a premium on harmony types to choose their partners with circumspection.

But what happens when types are private information? For instance, the United States cannot be certain that the Russian government prefers a non-nuclear to a nuclear Iran, no more than Russia can count on future US Congresses not to wish Poland to have sufficient access to nuclear deterrence to undercut Russian geopolitical influence in Eastern Europe. Is there a way for these two countries to signal their types to each other? Sadly, the answer is in the negative: there is nothing that either country can do to signal its type within the context of informal governance.

Signaling requires an actually incurred cost, either *ex ante* or *ex post* (see Fearon 1997), a feature that is more characteristic of treaty making than of informal governance. Possible *ex-ante* costs of treaty making include negotiations and ratification costs, while *ex-post* costs derive from the necessity of delivering on one's commitment whenever called on to do so by the other parties to the contract. Few if any of these costs are incurred in informal governance, where agreements are negotiated by a handful and left unratified, and where governments make cooperation strictly dependent on their own circumstances. The result is that informal governance is uninformative; aspiring regime members cannot use information signaling to circumvent the problem of exploitation.<sup>10</sup>

The main purpose of this section is to show that private information alone (without outside option) stunts informal governance and vindicates formalism. A secondary purpose is to endogenize the institutional costs of formalism – negotiation and ratification. Recall that these costs have a distinct part to play in institutionalist thought – to allow governments to make credible commitments – and thus should not be left to chance. The rest of the section is technical.

Assume that initiator  $i$  is a known harmony type ( $b_i < 0$ ), whereas Nature draws player  $j$ 's type randomly from a distribution function  $F(b_j)$ , with  $b_j$  continuously and uniformly distributed over interval  $[-B, B]$  and  $0 < B < f$ . Only  $j$  knows its type;  $i$  only knows the distribution function. Uninformed  $i$  plays first, choosing *ex-ante* cost,  $m \geq 0$ ; a positive value means the choice for a formal treaty, while a zero value means informal governance. Informed  $j$  plays second; it can reject the proffered agreement format, in which case the game is over and the payoffs are  $(0,0)$ ; or it may accept the format, in which case the game proceeds as earlier, with Nature drawing from the set of contingencies and the two players playing simultaneously

<sup>10</sup> Surely, the defecting government may still incur audience and reputation costs, but these are indirect, undependable, and thus low-powered incentives in comparison to direct punishment threats.

and repeatedly. A potential equilibrium is full revelation of  $j$ 's type, in which country  $i$  chooses an *ex-ante* cost that enables it to sort possible  $j$  types between the two categories of, first, cooperators (low positive or negative  $b_j$ ) who select into the regime and cooperate, and, second, exploiters (high  $b_j$ ) who select out of the regime. Which of formal and informal governance is better able to achieve this revelation equilibrium?

The difference between the two forms of governance in this new version of the game boils down to two features: the formal treaty has an *ex-ante* cost  $m > 0$  and a one-round punishment while the informal game has an *ex-ante* cost of zero and provides for no punishment in response to defection.

Proposition 2 states that formal governance achieves a revelation equilibrium.

**Proposition 2** [*revelation through negotiation and ratification costs*] *The strategy pair according to which, on one side, initiator  $i$  offers  $m = m^*$  and then plays the game of formal cooperation featured in Proposition 1 and, on the other side, respondent  $j$  accepts the agreement if  $b_j \leq \hat{b}_j$  and then plays the game of formal cooperation of Proposition 1, but rejects the agreement otherwise, is a perfect Bayesian Nash equilibrium.*

The proof and specific value of  $m^*$  and  $\hat{b}_j$  can be found in the appendix. The intuition behind the result goes like this. All harmony types (with  $b_j < 0$ ) accept the regime and cooperate. In addition, a minority of PD types with a positive, yet low, cost of cooperation (positive and low  $b_j$ ) also accept and cooperate on account of the punishment mechanism. In contrast, high  $b_j$  types systematically reject the agreement. The initiator country  $i$  can separate cooperators from defectors by offering a formal agreement that comes with a carefully chosen *ex-ante* negotiating cost  $m$ . The intuition behind the working of the selection mechanism is that the *ex-ante* cost brings down all payoffs, with the result that the countries with already high costs of cooperation prefer not to accept the proffered arrangement in the first place.

This result obtains when the choice offered by the initiator is between a formal regime and no regime. What happens when the choice is between an informal regime (with no negotiating and ratifying costs  $m$ ) and no regime? In this case, all  $j$  types accept the agreement, but while those with a negative cost of cooperation (harmony types) cooperate, those with a positive cost (PD types) defect. In the end, information is revealed here too, but not in a way that would guard the initiator from being exploited until the end of time by a PD-type  $j$ ; exploitation is part of the equilibrium.

Note that forever exploitation occurs despite the fact that the Harmony type may quit at any time (or even choose not to play). Yet it will not,



because exploitation is better than nothing (recall that  $b_i < 0$ ). Of course, cooperation would be better still.

This exploitative equilibrium can hardly be considered as a case of informal governance. Exploitation does not conform to any notion of cooperation, formal or informal, but seems more characteristic of unilateralism. Moreover, nobody in reality ever gets exploited until the end of time. This is because repeated exploitation typically extols a long-term price on the cooperator, a price that is not reflected in the simple payoffs featured in the game. Amending the model in a way that would reduce the future payoffs of a player every time this player were the only one to cooperate would place a limit on how long this player would be willing to be exploited – basically, until it has nothing left to lose. The informal strategy cannot support long-term cooperation.

In sum, information transmission is not the solution to the problem of exploitation that cooperators face in informal governance. We are entitled to conclude at this point of the reasoning that the prospects for informal governance are rather bleak; if there are any positive benefits to informal governance, they arise neither from contract enforcement nor informational gains. The saving grace of informal governance lies elsewhere.

### Outside option

Consider the following analogy with human relations. Couples have two ways of institutionalizing their relation: marriage, a legal procedure, and informal relationship, unsanctioned by the law. Marriage is celebrated for protecting the interests of the weaker part of the couple – the woman in a traditional western society where the man alone was entitled to earn an income, either member of the couple in our times. Informal relationships, except for registered alternatives to marriage, typically are not protected by the law, and yet they exist and are not necessarily a source of exploitation of one partner by the other. The main reason, I believe, derives from the outside option that an informal relationship preserves for each partner, providing each side with a reservation value that has the effect of deterring or terminating abuse by the other side.

Formal and informal relationships differ in how they cope with private information. In marriage, private information requires a costly sorting mechanism, lest one's unhappy choice of partner yield long-term grief. It traditionally takes the form of an expansive diamond ring, a lavish ceremony and pre-booked honeymoon, and, last but not least, the prospect of a costly divorce for those who commit too lightly. In contrast, little information is required to start or end an informal relationship. Learning does take place as the relationship evolves and it is on the basis of such

learning that partners may elect to try their luck with an unknown quantity rather than continuing with their present partner. But the nature of that learning, unlike that that is necessary in a formal relationship, is imprecise in the sense that it does not matter whether the failure of the relationship with one's current partner owes to personal divergence or poor circumstances to the extent that the experience was on average inferior to what one may expect from starting over with an unknown quantity. Learning in an informal relationship need not be very informative.

What is true of couples seems also to be true of countries. The initiator of an informal regime may fall prey to long-term exploitation by a PD type only if it has no outside option. Introducing new possible partners raises the reservation value from zero to a positive number and makes informal cooperation sustainable over the long run. Thus amended, the no-obligation, no-punishment strategy becomes an apt description of the notion and practice of informal governance.

Is it the existence of an outside option that makes informalism possible, as I argue here, or is it the existence of informalism that makes the search for an outside option possible? Fang and Ramsay (2010, 194) opt for the latter when they argue that the rigid pre-1914-alliance-system's increased search-costs weakened each alliance initiator's leverage – Russia and Austria-Hungary, respectively, for the Triple Entente and the Triple Alliance. While I agree with the short-term implications of this logic, I also feel that it misses the bigger picture. The reason for why the agreements were formal in the first place is because there were no outside options when Russia and Austria negotiated their respective alliances. This point is further developed in the empirical section below.

I reformulate the informal governance strategy of Definition 2 by introducing the outside options. As above, I assume that initiator  $i$  is a harmony type ( $b_i < 0$ ). Respondent  $j$ 's type, in contrast, is only known of himself. I also assume an outside option to be available in every round.

**Definition 3** [*harmony's informal governance strategy with outside option* ( $\sigma_{OH}$ )] *Apply the agreed distribution. This means: (a) play  $x$  when Nature draws a contingency favorable to  $i$  and (b) play  $y$  when Nature draws a contingency favorable to  $j$ . Then observe  $j$ 's behavior and circumstances and update beliefs about  $j$ 's type. If the updated beliefs were equal to or above the prior, move to the next period and keep interacting with  $j$ . If the updated beliefs fall below the prior, switch to unknown quantity  $h$ . Repeat these moves, with a new partner as needed.*

Allowing the initiator to interact with a first, unknown respondent, while reserving for the initiator the right to switch to a second unknown partner any time it feels like it, frees the initiator's first interaction from

exploitation, as either the first respondent cooperates and the initiator reciprocates, or that respondent defects and the initiator switches partner. For if the first respondent is caught defecting, the initiator both learns that this partner is a PD type and has an incentive to terminate the relationship.

How is the first respondent going to respond to an initiator that possesses an outside option? It all depends on whether that respondent also enjoys an outside option or not. Consider, first, the case in which the respondent has no outside option. The initiator's threat of exclusion works like a grim-trigger strategy – punishment by abandonment is for ever – much harsher than the one-round punishment, thereby ensuring participation from much more hardened PD types.

Consider now the situation in which the respondent also has one outside option available in every round of the game. Then it is obvious that some respondent types, with positive and high cooperation cost  $b_j$ , will first exploit the initiator and then happily switch to their outside option. But some other respondent types, with a positive yet low cost  $b_j$ , will be better off cooperating because their reservation value, which is the expected value of interacting with an unknown quantity, is inferior to the expected value of cooperating with the initiator, who is known to be a harmony type.

I characterize the equilibrium in which initiator and respondent each enjoy an outside option; initiator is known to be a harmony type, but the respondent and respective outside options are unknown quantity to initiator and to one another. Probability distributions, drawn contingencies, and moves are common knowledge, allowing for beliefs to be upgraded. The value of each outside option is not an arbitrarily chosen value but is endogenous to the equilibrium.<sup>11</sup>

The initiator plays strategy  $\sigma_{OH}$  of Definition 3. If the respondent is a harmony type, it plays the same strategy. If the respondent is a PD type, it plays a strategy that is conditional on the value of  $b_j$ , which I now define:

**Definition 4** [*PD-type  $j$ 's informal governance strategy with outside option ( $\sigma_{OPD}$ )*] If  $b_j < b_{j_2}$ ,  $j$ 's informal strategy is  $\sigma_{OH}$  in Definition 3, the same as the initiator. If  $b_j > b_{j_2}$ , its informal strategy is (a) to play  $x$  in response to any contingency, and (b) switch if other player played  $x$  in response to a contingency favorable to  $j$ . Repeat (a) and (b), with a new partner as needed.

It follows that informal governance is an equilibrium.

**Proposition 3** [*informal governance equilibrium*] Assuming one-sided information asymmetry between a known initiator  $i$  and an unknown

<sup>11</sup> I am not aware of any other attempt at endogenizing outside options.

quantity  $j$ , the strategy pairs  $b_j > \tilde{b}_j$ , if the unknown quantity is a harmony type, or  $(\sigma_{OH}, \sigma_{PD})$  if it is a PD type, form a perfect Bayesian Nash equilibrium.

The proof is in the appendix. The interesting finding here is that *informal governance can also elicit cooperation from PD types, provided that the initiator is a harmony type and enjoys an outside option*, irrespective of whether PD types also enjoy or not an outside option. Not all PD types qualify, to be sure, only the mild ones.

The reason for why a PD-type respondent cooperates in the absence of a formal obligation is that the respondent fears losing the initiator, known to be harmony, more so than it would fear losing its own outside option, an unknown quantity. And so if the respondent were to force the initiator into exercising its outside option, the respondent would face a bleaker future. The fear of losing the initiator acts like a credible punishment threat, high enough to lead the respondent (provided it is of a mild PD type) to cooperate.

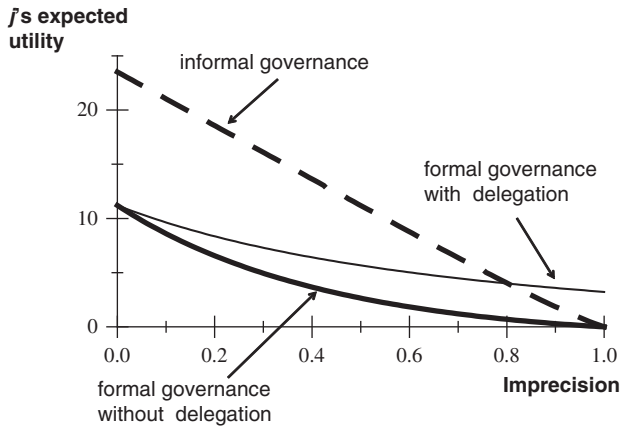
Note the irony of the situation. This outcome, beneficial to the initiator, obtains only because the initiator is known to be a harmony type; for were it unknown too, the respondent would see no difference between interacting with the initiator and its own outside option. Providing the initiator with an outside option moves it from being a sucker in a game of asymmetric information into becoming an influential partner in that very same game.<sup>12</sup>

In Figure 2, I compare and contrast the initiator's payoff for the strategies of formalism with and without delegation (both deployed in Proposition 2) and informalism (Proposition 3). The relative placement of the curves is similar to the one encountered in Figure 1, with one important difference, the fact that the two formal governance curves start way below the informal governance curve. This reflects the endogenously fixed cost  $m$ , occasioned by the screening of respondent types.

No hard proposition on the role of imprecision in the design of governance comes out of Figure 2 any more than it did from Figure 1; the relative positioning of the two top curves (informal and formal with delegation) depends on the specific values assumed by a host of parameters.

As the formal part of this paper draws to a close, it is worth restating the main finding. Formal governance is inefficient in the presence of uncertainty or disagreement, two contexts that generate imprecision in the writing of the agreement. Imprecision hurts cooperation. Imprecision can be reduced

<sup>12</sup> Separate calculations (unreported here) permitted me to establish that it does not make any difference how many outside options each player disposes of, provided it is the same number for everyone and that number is one or more.



**Figure 2** Impact of imprecision on the initiator’s payoff, with  $f = 4$ ,  $b_i = -1$ ,  $\beta = 0.5$ ,  $\theta = 0.5$ ,  $B = 3$ ,  $\tau = 0.8$ , and  $\delta = 0.9$  (private information).

through delegation but at the price of *ex-ante* fixed costs, which have to be paid to make commitments credible in a context of private information. Yet the fact that formal governance is inefficient or costly does not automatically make informal governance a better prospect because it is beset with the risk of cheating. Informal cooperation is possible and preferable to formal governance only when cheating is contained. This is so in two cases. First, between two known harmony types. Second, between a known harmony type and an unknown mild PD type, provided that outside options be made available.

We are now in measure to provide a precise content to the notion of ‘like-mindedness’, on which so much of the literature relies to support informal governance.

**Definition 5** *Like-minded are, first, the harmony types and, second, the mild PD types that are able to sustain cooperation with a harmony type that can avail itself of an outside option.*

In sum, the model points to a correlation between the availability of an outside option and informal governance in mild PD game structures. I illustrate this correlation in the next section by looking at the modern history of security regimes.

\*

Although power is not my concern, it enters the present model in several ways, with contradictory effects on governance. First, as asymmetric

outside options. Asymmetry works against the side that is shortchanged; having no outside option for a respondent when the initiator enjoys one arms the latter with a unilateral grim-trigger punishment option. The results reached by the power literature carry over in the sense that greater power increases the domain of informal governance.

But power also enters the model as a country's bargaining power, the fact that initiator  $i$ 's share of favorable calls is  $\beta$  while that of respondent  $j$  is  $1-\beta$ . While the two simulations presented earlier assumed equal bargaining power ( $\beta = 0.5$ ) as an approximation of the Nash bargaining solution,  $\beta$  usually is an endogenous variable, reflecting both the asymmetry in outside options and the bargaining protocol. In the extreme setup where the initiator has a better outside option or makes take-it-or-leave-it offers, the outcome can be perverse: the initiator, despite being a harmony type, is going to want to appropriate the whole surplus (offering  $\beta = 1$ ), with the result of driving away all PD types and indulging in the exploitation of other harmony types.<sup>13</sup> This is true irrespective of the formal or informal nature of the regime. In such a case, power asymmetry might reduce the domain of governance.

### Historical illustrations

I survey the main security regimes that have existed in the last two centuries. I restrict the evidence to security regimes because they best match an unstated assumption of the present model: that actors are state centric. I divide my survey of security regimes into three categories: territorial settlements, crisis prevention and management, and unilateral signaling.

#### *Territorial settlements*

Territorial settlements that end wars – peace treaties – are always formal. The Napoleonic wars were settled by the Treaty of Vienna in 1815, the Crimean War by the Treaty of Paris in 1856, the Russo-Turkish War of 1877 by the 1878 Treaty of Berlin; another Treaty of Berlin, that of 1885, divided Africa between colonial powers. World War I was settled by the Versailles Treaty and successive war reparation conferences. The United States settled World War II by means of the Paris Peace Treaties in 1946 with Italy, Romania, Hungary, Bulgaria, and Finland, and separate treaties with Japan in 1951, Austria in 1955, and Germany in 1990. The Postdam Agreement on the occupation of Germany and territorial changes made to

<sup>13</sup> Ways of avoiding this corner solution include the introduction of search costs (for outside options) or the choice for a less lopsided, though not fully equitable, bargaining protocol (such as Rubinstein's alternating offer); see Muthoo (1999).

Poland, which was executed as a communiqué by the three victors of World War II, was not an exception to this formalistic pattern because it involved no cooperation, not even nominal, between victor and vanquished.

According to the model, peace treaties are formal because they deal with highly distributive issues. Territorial settlements are a direct application of the notion of national sovereignty, according to which territorial claims between countries are mutually exclusive. The inherent conflict disqualifies informal governance in favor of legal governance.

### *Crisis prevention and management*

Besides peace treaties, cooperation also involves crisis prevention and management – alliances and collective security, arms control, military intervention, and export control and interdiction. These are areas in which goals are sufficiently convergent to justify cooperation. Given moderate conflict, the model predicts that legalization should be a reverse function of the availability of outside options. More specifically, two dynamics should be empirically observable:

- (1) the disappearance of outside options makes informalism untenable;
- (2) the appearance of outside options makes formalism relatively costlier and, in the case where it was found to be too costly, abandoned in favor of informalism.

A broad determinant of the number of outside options is the degree of polarization in the international system – whether cleavages among great powers overlap along rigid coalitions or are crisscrossing and opening new alignment possibilities. A great power enjoys more outside options in a depolarized system, where a country's bargaining power is enhanced by the presence of other possible negotiating partners, than in a polarized system, where each negotiator's hand is weakened by the lack of a fallback position. Therefore, the model predicts that with respect to issues of crisis prevention and management, formalism should be more common in polarized systems, of which bipolarity is an extreme case, while informalism should be mostly found in depolarized systems.<sup>14</sup>

During the 19th century, the international system moved from one extreme to another. Depolarized and quasi-harmonious under the Concert of Europe, the system first transitioned into a phase of short-lived, yet changing, coalitions with the Crimean War, and then, starting with the signing of the Triple Alliance in 1882, it gradually settled into the two rigid

<sup>14</sup> The 'threat' vs. 'risk' distinction offered by Wallander and Keohane (1999, 24) strikes me as similar to the present one between 'high' and 'low' degree of polarization.

camps that fought each other in World War I. Along with this secular rise in polarization, the model suggests that we should observe a shift from less to more formalism. And so we do.

The Concert of Europe provides a good illustration of the conjunction of depolarization and informalism. The Concert was a summit meeting of the heads of the five great powers, who gathered to manage the affairs of the European continent in the wake of the Congress of Vienna (see Mitzen 2013). The Concert had a limited distributive content that reflected the fact that its purpose was to maintain the territorial settlement between the four winners and a restored French dynasty – a ‘benign shared hegemony’ in Schroeder’s (1994, 6) words. Having no covenant defining goals or secretariat interpreting them, it was imprecise. Meetings were *ad hoc*, usually called in response to a crisis.

The 1848 revolution, the resurgence of territorial rivalries, and, most importantly, the emergence of Germany shattered the existing balance of power by turning France and Russia into long-term rivals of Germany. For two decades, Bismarck sought to contain the emerging realignment by signing formal, yet short-lived, alliance treaties (the Three Emperor’s League in 1872 and 1881 and the 1887 secret Reinsurance Treaty with Russia). Abandoning conference diplomacy, chancelleries resorted to treaties with limited duration, calling for periodic renegotiations, as a way of securing enough flexibility to exercise eventual outside options (see Koremenos 2005). Soon, outside options vanished altogether when a threatened Britain – threatened by Germany’s naval buildup – renounced its old policy of ‘splendid isolation’ to side with France and Russia. The two camps dug themselves into deeply formal treaty making (Triple Alliance of 1882, Franco-Russian Alliance of 1893, the French and Russian ententes with Britain), with texts spelling out precisely who was expected to do what under which circumstances and sequence of events.<sup>15</sup> The formal alliances were not created to commit their respective members to fight on each other’s side, something they would automatically do given the absence of any other option, but to have allied military staffs coordinate strategy and better prepare for the eventuality of war against a predictable enemy. Formalism merely reflected the disappearance of outside options and the need to secure support in the face of threat.

In sum, as the European system progressively shifted from the club of 1815 to the two coalitions of 1907, the number of outside options enjoyed by each great power with respect to crisis management dropped from three (if one of the five powers pursued a negotiation with another, it still

<sup>15</sup> For details on the alliance dynamics that preceded World War I, see Snyder (1984).



disposed of three outside options) to zero, while the instruments of crisis management changed from informal conference diplomacy to long-term formal treaty making.

The war did not change the underlying power rivalry and the same coalitions ended up forming again 20 years later. The only differences were valiant, yet futile, attempts at legalizing collective security with the League of Nations first, the 1928 Kellogg-Briand Pact then, and the 1945 United Nations last, on which I expand in the next section.

The last 60 years have taken the world on a reverse trajectory, unfolding from the bipolar confrontation of the Cold War to the less polarized current environment. Along with this semi-secular decline in polarization, the model suggests that we should observe a shift from less to more informality, which, again, we do.

Bipolarity is an extreme case of polarization, in which friends and foes are known and outside options non-existent. As noted by Waltz (1979, 168, 170), 'In a bipolar world uncertainty lessens and calculations are easier to make...who is a danger to whom is never in doubt'. Bipolarity reduced outside options for everyone, not just the superpowers, as liberal economies feared the communist threat and communist regimes feared capitalist imperialism. By the mid-50s, about 50 countries had signed a long-term formal mutual assistance guarantee agreement with either the United States or the Soviet Union.<sup>16</sup>

Moreover, caught in a race for influence, the two superpowers sought to divide the world into two blocs, each with exclusive membership. John Foster Dulles inveighed against the immorality of neutralism. Russian leaders described neutralists as dupes of capitalist countries. It is not until 1961 that, under the leadership of Tito, Nehru, Nasser and a few other leaders of the developing world, 25 nations formed the 'non-aligned movement'. Nevertheless, not a single Latin American country, except for Cuba, dared to attend the meeting after the US ambassador reminded Brazilian President Quadros that Brazil was a 'committed nation' under the Rio Pact (see Hershberg 2007, 376). And although non-aligned countries meant not to sign formal alliance agreements, nor informal ones for that matter, but preferred to entrust their security to the balance of power, the logic of the bipolar system placed a premium on formalism, in direct contradiction with the declared principles of the movement (see Bebler 1975, 290; Appadorai 1981, 8). A good instance is the treaty of non-aggression that India signed

<sup>16</sup> These agreements were the North Atlantic Treaty Organization (NATO), the Rio Pact, the Warsaw Pact, the Australia New Zealand United States treaty, South East Asia Organization, Central Eastern Treaty Organization and US bilateral agreements with the Philippines, South Korea, Taiwan, and Japan.

with the Soviet Union in 1971 in response to the long-term assistance provided by the United States to Pakistan and the growing Sino-Soviet rift.

The limited fluidity of the Cold War also accounted for the extant formalism between the two great powers, as evidenced by the Anti-Ballistic Missile (ABM) Treaty, the Strategic Arms-Limitation Treaties, the Intermediate-Range Forces Treaty, and several others that, although signed after 1989, were initiated during the Cold War.<sup>17</sup> The two superpowers were not like-minded partners, but self-declared enemies. Despite a common interest in slowing down the arms race, lack of trust forced them to back up their commitments with the formal apparatus of international law. An essential component of each agreement was, according to Eilstrup-Sangiovanni (2009, 211), 'strong verifiability measures to ensure that cheating would be caught'.

The collapse of the Soviet Bloc and the move to a multipolar world brought a radical change: outside options multiplied and informalism thrived. It is not that the formal treaties that had been concluded under the Cold War were denounced or replaced by informal agreements. Surely, a few of them such as the Warsaw Pact and the ABM treaty were terminated, the Biological Weapons Convention was gutted, while the NPT was sidelined by its informal rival, the PSI.<sup>18</sup> The overwhelming majority of alliance treaties and arms control agreements were continued or, as in the case of the START agreement, renegotiated. But the emphasis shifted over time from alliance and arms control, issues that lent themselves to treaty making, to export control and the fight against terrorism, issues that seem better handled informally. While there had been only one informal export control regime under the Cold War, the Coordinating Committee for Multilateral Export Controls (CoCom), three CoCom-like regimes were created in the final years and after.<sup>19</sup>

The shift in emphasis reflected a change in the nature of the threat. From targeted, state based, and all-inclusive, threats after the Cold War became diffuse, involving non-state actors, and issue specific. Along, changed the number of outside options, for threat and outside option are dual images of

<sup>17</sup> Discussion on the 1992 Treaty on Open Skies had started under Eisenhower. Negotiations on the 1992 START1 agreement had started in 1982. Discussions on 1993 Chemical Weapons Convention began in 1968 while discussions on the 1990 Treaty on Conventional Armed Forces in Europe (imposing troop ceilings for NATO and Warsaw Pact forces in Europe) were initiated in 1972.

<sup>18</sup> The Warsaw Pact was replaced in 1992 by a downsized Collective Security Treaty Organization, a treaty of mutual assistance in case of aggression between Russia and the ex-Soviet republics.

<sup>19</sup> The Australia Group on the export of chemical and biological weapons, the Wassenaar Arrangement on dual-use technology, and the Missile Technology Control Regime.

each other: a threat originates from a foe while an outside option comes from a potentially friendly third party. Non-existent in the bipolar system, outside options emerged in an *ad hoc* and issue-specific way. The military coalition that the United States put together in the Iraq war no longer was NATO based, but a so-called ‘coalition of the willing’, aptly described by Patrick (2010, 32) as a ‘hub and spoke’ arrangement founded on bilateral deals between the United States and a large and heterogeneous group of countries’, in which not every partner was asked to make the same contribution (see also Bowie, Haffa, and Mullins 2003). The rationales that were provided by administration insiders and observers alike directly invoked the notion of greater choice: Richard Haas, the then-State Department’s director of policy planning, spoke of ‘à la carte multilateralism’, while an observer wrote ‘White House says the U.S. is not a loner, just choosy’ (see Shanker 2001).

It is in this context that the two ex-superpowers started cooperating on specific issues of mutual concern without the need for hard legal apparatus. The Cooperative Threat Reduction program provided informal assistance to ex-Soviet bloc countries to reduce and safeguard their strategic arms and nuclear weapon stockpiles. The already-mentioned MTCR, Wassenaar Arrangement, and PSI are other instances of choice in favor of informalism, reflecting the possibility for ex-enemies to be of one mind on some issues.<sup>20</sup>

The PSI deserves special attention. Founded during the Cold War, the nuclear non-proliferation regime, which revolved around the NPT and a complementary organization, the 1974 Nuclear Suppliers Group, became deadlocked. The Bush administration seized the opportunity made available by the end of the Cold War to create a new forum, free from past controversies and memberships, and see it patronized by Russia and other important players. This positive response has allowed the United States so far to keep the PSI free from the basic trappings of an international organization – secretariat, voting procedures, building, funding, internal memory, issue linkage, and seemingly unkept past promises. Although the Bush administration pursued the initiative approach in several other issue areas, the format, however, has turned out to be more useful for sharing information than to validate enforcement.<sup>21</sup>

In sum, periods of intense polarization (post-German unification Europe, Cold War) have tended to coexist with a formal restlessness that is not as intensely felt in periods of depolarization (Concert, present era). To the

<sup>20</sup> See Eilstrup-Sangiovanni (2009) for a more comprehensive coverage.

<sup>21</sup> Other initiatives include the 2001 Global Health Security Initiative on bioterrorism, the 2006 Global Initiative to Combat Nuclear Terrorism, and the 2006 Global Initiative on Sharing All Influenza Data.

extent that depolarization is a valid proxy for the emergence of outside options, the historical narrative shows that informal governance is more likely to be found useful by great powers when their bargaining power is strengthened by the presence of outside options.

### *Formalism as signal*

A reason to choose formalism over informalism is to signal that one does not value one's outside option. Such is especially the case when the outside option is unilateralism. Although in theory the choice between unilateralism and multilateralism does not necessarily involve a choice between informalism and formalism because the multilateral option could be as informal as the unilateral one, it turns out that, in practice, the two dimensions overlap, thus providing us with additional evidence on the correlation between informalism and outside option.

Both the League and the United Nations were instances of the desire of American administrations to signal to the world that, by embracing a formal process, the United States no longer saw isolationism as a valuable outside option. American isolationism rested on the belief that the United States could protect itself against war by withdrawing within itself and paying regard only to its frontiers. This belief was first falsified by submarine warfare and, rather than adopting the equally falsified idea that limited alliances would suffice to prevent war, the Wilson administration opted for collective security, whose grace was to have never been tried before. Collective security required the creation of a formal institution, headed by an Assembly and a Council. Thirty years later, World War II drove home the same lesson that neither diplomacy nor US isolationism, consummated in 1919 by the Senate rejection of the League, were credible options. It was also believed that the League's multiple failures, including that of not having been consulted to prevent World War II, could be remedied by redistributing decision-making power to the Great Powers and making their decisions binding (see Sobel 1994).

The same argument has been made in the context of European and Asian regional integration. Peters (1999) argues that Germany, following the fall of the Berlin Wall in 1989, successfully pressed for the transformation of the informal Conference on Security and Cooperation in Europe into the formal Organization thereof (OSCE) to prove that its reunification was not pursued in the country's narrow self-interest and that it would remain committed to multilateralism rather than regain unilateral freedom of action. The same argument was made many times to account for Germany's eagerness to support European integration in the first place by joining the European Coal and Steel Community (see Gillingham 1991, 364). In the

same vein, China's initial entry in the informal ASEAN (Association of Southeast Asian Nations) Regional Forum and its subsequent acceptance of an admittedly timid level of institutionalization and somewhat more intrusive agenda issues had the positive effect of signaling to its neighbors its endorsement of the status quo and its willingness to learn to work with others (see Johnston 1999, 316).

More broadly, Voeten (2001) and Thompson (2006) proposed that securing the endorsement of an international institution is a costly process that functions as a credible signal of willingness to work with other countries.

## Conclusion

The efficiency approach to informal governance stumbles on a dilemma: informal governance makes cooperation easier than formal governance in the face of institutional costs, but by skipping obligation and punishment, informal governance condones cheating, with the effect of undercutting the sustainability of cooperation. The high risk of cheating led some of the literature to circumscribe the scope of informal governance to like-minded individuals, thinking in harmony or almost, but with shaky microfoundations.

The paper seeks to provide the missing microfoundations. Resting on the three pillars of legalization – obligation, precision, and delegation – the model innovates in several respects. First, it offers a way of operationalizing the notion of imprecision, deliberate or unexpected, as an imperfect partition of expected contingencies. Second, by introducing private information, the model endogenizes the value of the negotiation and ratification costs, which are seen by the literature as causes for abandoning formalism. Last, by equitably distributing outside options to all protagonists – with outside option a concept borrowed from the neorealist strand of regime analysis – the model endogenizes the notion of outside option.

The main substantive result is that unknown PD types are willing to cooperate informally with known harmony types. The risk of exploitation is hedged by the outside option. Not only is the capacity to exercise an outside option the essence of informal governance, but, by operating like an implicit punishment strategy, implicit in the sense that the strategy requires no prior definition of, or agreement on, the terms of cooperation, it leads like-minded types, formally defined as mild PD types, to prefer cooperation over defection with a harmony type.

A quick survey of security regimes during the last two centuries helped illustrate the role of outside options. Most alliances, coalitions, and other groupings set up to prevent and manage eventual crises, issues that are typically not as distributive as territorial settlements, were shown to be

either formal or informal according to the availability of outside options. Proxying the latter variable by the degree of polarization of a power structure, I showed that formalism closely tracked a rise in polarization of the international system both at the turn of the 20th century and during the Cold war. Other periods, such as the early part of the 19th century and the current period, in contrast, show a correlation between depolarization and informalism. The shift from the bipolar system of the Cold War to the current multipolar power distribution offers a vivid example of the connection made by recent American administrations between the greater choice of partners and the declining utility of treaties.

Finally, shifting the emphasis away from structure toward agency, I further illustrated the central prediction of the model by pointing to the common state strategy of choosing formal multilateral settings to signal disregard for one's outside option.

In the end, informalism is more than a remedy to the costs and rigidities of treaty making. The internal discipline that guards members against opportunism radically differs between the two modes of governance. In formal arrangements, it is the official threat of suspending cooperation that makes cooperation sustainable. In informal arrangements, it is the casual threat of walking out of a relation, provided that it is backed up with the existence of an outside option, that makes cooperation possible.

## Acknowledgments

The author thanks Bryan Coutain, Xinyuan Dai, Christina Davis, Mette Eilstrup-Sangiovanni, Justin Leinaweaver, William Minozzi, Alexander Thompson, Byungwon Woo, the editors, and three anonymous reviewers for their insightful comments on previous drafts. An earlier version was presented at the Political Economy of International Organizations Conference held at Villanova University on 26–28 January 2012.

## References

- Abbott, Kenneth W., and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54(3):421–56.
- Abbott, Kenneth W., Robert O. Keohane, Andrew Moravcsik, Anne-Marie Slaughter, and Duncan Snidal. 2000. "The Concept of Legalization." *International Organization* 54(3): 401–19.
- Appadorai, Arjun, and John Gillingham. 1981. "Non-Alignment: Some Important Issues." *International Studies* 20(1–2):3–11.
- Bebler, Anton. 1975. "Security Aspects of Non-Alignment." *International Studies* 14(2): 289–302.
- Boese, Wade. 2003. "The Proliferation Security Initiative: An Interview with John Bolton." *Arms Control Today*, December 2003, p. 37.

- Bowie, Christopher J., Robert P. Hafta Jr., and Robert E. Mullins 2003. "Future War: What Trends in American's Post-Cold War Military Conflicts Tell us about Early 21st Century Warfare." Northrop Grumman Analysis Center Papers. Accessed October 20, 2014. <http://www.northropgrumman.com>.
- Dehousse, Renaud. 1997. "Regulation by Networks in the European Community: The Role of European Agencies." *Journal of European Public Policy* 4(2):246–61.
- Downs, George W., and David M. Rocke 1995. *Optimal Imperfection! Domestic Uncertainty and Institutions in International Relations*. Princeton: Princeton University Press.
- Downs, George W., David M. Rocke, and Peter N. Barsoom 1996. "Is the good news about compliance good news about cooperation?" *International Organization* 50(3):379–406.
- Eilstrup-Sangiovanni, Mette. 2009. "Varieties of Cooperation: Government Networks in International Security." In *Networked Politics: Agency, Power, and Governance*, edited by Miles Kahler, 194–227. Ithaca: Cornell University Press.
- Fang, Songying, and Kristopher W. Ramsay 2010. "Outside Options and Burden Sharing in Nonbinding Alliances." *Political Research Quarterly* 63:188–202.
- Fearon, James D. 1997. "Signaling Foreign Policy Interests." *Journal of Conflict Resolution* 41(1):68–90.
- Galbraith, Jean. 2003. "The Bush Administration's Response to the International Criminal Court." *Berkeley Journal of International Law* 21(3):683–702.
- Gillingham, J. 1991. *Coal, Steel, and the Rebirth of Europe, 1945–1955: The Germans and French from Ruhr Conflict to Economic Community*. Cambridge: Cambridge University Press.
- Gruber, Lloyd. 2000. *Ruling the World: Power Politics and the Rise of Supranational Institutions*. Princeton: Princeton University Press.
- Guzman, Andrew T., and Timothy L. Meyer 2010. "International Solf Law." *Journal of Legal Analysis* 2(1):171–225.
- Helleiner, Eric, Stefano Pagliari, and Hubert Zimmermann. 2010. *Global Finance in Crisis: The Politics of International Regulatory Change*. London: Routledge.
- Hershberg, James G. 2007. "'High-Spirited Confusion': Brazil, the 1961 Belgrade Non-Aligned Conference, and the Limits of an 'Independent' Foreign Policy during the High Cold War." *Cold War History* 7(3):373–88.
- Johns, Leslie. 2012. "Courts as Coordinators: Endogenous Enforcement and Jurisdiction in International Adjudication." *Journal of Conflict Resolution* 56(2):257–89.
- . 2014. "Depth versus Rigidity in the Design of International Trade Agreements." *Journal of Theoretical Politics* 26(3):468–95.
- Johns, Leslie, and Krzysztof Pelc J. 2009. "Forth Depth versus Rigidity in the Design of International Trade Agreements." *Journal of Theoretical Politics*.
- Johnston, Alastair I. 1999. "The Myth of the ASEAN Way? Explaining the Evolution of the ASEAN Regional Forum." In *Imperfect Unions: Security Institutions over Time and Space*, edited by Helga Haftendorn, Robert O. Keohane, and Celeste A. Wallander, 287–324. Oxford: Oxford University Press.
- Josselin, Daphné. 2009. "Regime Interplay in Public-Private Governance: Taking Stock of the Relationship Between the Paris Club and Private Creditors Between 1982 and 2005." *Global Governance* 15(4):521–38.
- Kahler, Miles, and David A. Lake 2009. "Economic Integration and Global Governance: Why So Little Supranationalism?" In *The Politics of Global Regulation*, edited by Walter Mattli, and Ngaire Woods, 242–76. Princeton: Princeton University Press.
- Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton: Princeton University Press.

- . 2001. "Governance in a Partially Globalized World." *American Political Science Review* 95(1):1–14.
- Keohane, Robert O., and Joseph S. Nye 1974. "Transgovernmental Relations and International Organizations." *World Politics* 27(1):39–62.
- Kerwer, Dieter. 2006. "Governing Financial Markets by International Standards." In *New Modes of Governance in the Global System: Exploring Publicness, Delegation and Inclusiveness*, edited by Koenig-Archibugi, Mathias, and Michael Zu(e)rn, 77–100. New York: Palgrave Macmillan.
- Kleine, Mareike. 2013. "Knowing your Limits: Informal Governance and Judgement in the EU." *Review of International Organizations* 8(2):245–64.
- Koremenos, Barbara. 2005. "Contracting Around International Uncertainty." *American Political Science Review* 99(4):549–65.
- . 2008. "When, What, and Why do States Choose to Delegate?" *Law and Contemporary Problems* 71(1):151–92.
- Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 1999. "The Rational Design of International Institutions." *International Organization* 55(4):761–99.
- Krasner, Stephen D. 1991. "Global Communications and National Power: Life on the Pareto Frontier." *World Politics* 43(3):336–66.
- Lipson, Charles. 1991. "Why are some International Agreements Informal?" *International Organization* 45(4):495–538.
- Lipson, Michael. 2005/2006. "Transgovernmental Networks and Nonproliferation: International Security and the Future of Global Governance." *International Journal* 61(1):179–98.
- Milgrom, Paul R., and John Roberts. 1992. *Economics, Organization and Management*. Englewood Cliffs, NJ: Prentice Hall.
- Mitzen, Jennifer. 2013. *Power in Concert: The Nineteenth-Century Origins of Global Governance*. Chicago: University of Chicago Press.
- Mosley, Layna. 2009. "Private Governance for the Public Good? Exploring Private Sector Participation in Global Financial Regulation." In *Power, Interdependence, and Nonstate Actors in World Politics*, edited by Helen V. Milner, and Andrew Moravcsik, 126–46. Princeton: Princeton University Press.
- Muthoo, Abhinay. 1999. *Bargaining Theory with Applications*. Cambridge: Cambridge University Press.
- Patrick, Stewart. 2010. "'The Mission Determines the Coalition': The United States and Multilateral Cooperation after 9/11." In *Cooperating for Peace and Security*, edited by Bruce D. Jones, Shepard Forman, and Richard Gowan, 20–44. Cambridge: Cambridge University Press.
- Peters, Ingo. 1999. "The OSCE and German Policy: A Study in How Institutions Matter." In *Imperfect Unions: Security Institutions Over Time and Space*, edited by Helga Hafendorn, Robert O. Keohane, and Celeste A. Wallander, 195–222. Oxford: Oxford University Press.
- Raustiala, Kal. 2002–2003. "The Architecture of International Cooperation: Transgovernmental Networks and the Future of International Law." *Virginia Journal of international Law* 43(2):2–92.
- Reinhardt, Eric. 2001. "Adjudication without enforcement in GATT disputes." *Journal of Conflict Resolution* 45(2):174–95.
- Rosendorff, B. Peter, and Helen V. Milner 2001. "The Optimal of International Trade Institutions: Uncertainty and Escape." *International Organization* 55(4):829–58.
- Shaffer, Gregory C., and Mark A. Pollack 2010. "Hard vs. Soft Law: Alternatives, Complements, and Antagonists in International Governance." *Minnesota Law Review* 94(3):706–99.



- 2011. “Hard versus Soft Law in International Security. University of Minnesota Law School.” Legal Studies Research Paper Series. Research Paper No. 11–13.
- Schroeder, Paul W. 1994. *The Transformation of European Politics, 1763–848*. New York: Oxford University Press.
- Shanker, Thom. 2001. “White House Says the U.S. Is Not a Loner, Just Choosy.” *New York Times*, July 31.
- Slaughter, Anne-Marie. 2004. *A New World Order*. Princeton: Princeton University Press.
- Snyder, Glenn H. 1984. “The Security Dilemma in Alliance Politics.” *World Politics* 36(4): 461–495.
- Sobel, Russell S. 1994. “The League of Nations Covenant and the United Nations Charter: An Analysis of Two International Constitutions.” *Constitutional Political Economy* 5(2): 173–92.
- Speyer, Bernhard. 2009. “Governing Global Financial Markets – Basel I and II: The Role of Non-State Actors.” In *Global Governance and the Role of Non-State Actors*, edited by Gunnar Folke Schuppert, 101–16. Baden-Baden: Nomos.
- Steinberg, Richard H. 2002. “In the Shadow of Law or Power? Consensus-Based Bargaining and Outcomes in the GATT/WTO.” *International Organization* 56(2):339–74.
- Stone, Randall W. 2011. *Controlling Institutions: International Organizations and the Global Economy*. Cambridge: Cambridge University Press.
- Thompson, Alexander. 2006. “Coercion through IOs: The Security Council and the Logic of Information Transformation.” *International Organization* 60(1):1–34.
- Vabulas, Felicity, and Duncan Snidal. 2013. “Organization Without Delegation: Informal Intergovernmental Organizations (IIGOs) and the Spectrum of Intergovernmental Affairs.” *Review of International Organizations* 8(2):193–220.
- Verdier, Pierre-Hugues. 2009. “Transnational Regulatory Networks and Their Limits.” *Yale Journal of International Law* 34(1):114–72.
- Victor, David G. 2011. *Global Warming Gridlock: Creating More Effective Strategies for Protecting the Planet*. Cambridge: Cambridge University Press.
- Voeten, Erik. 2001. “Outside Options and the Logic of Security Council Action.” *American Political Science Review* 95(4):845–58.
- von Stein, Jana. 2008. “The International Law and Politics of Climate Change: Ratification of the United Nations Framework Convention and the Kyoto Protocol.” *Journal of Conflict Resolution* 52(2):243–68.
- Wallander, Celeste A., and Robert O. Keohane 1999. “Risk, Threat, and Security Institutions.” In *Imperfect Unions: Security Institutions over Time and Space*, edited by Helga Haftendorn, Robert O. Keohane, and Celeste A. Wallander, 21–47. Oxford: Oxford University Press.
- Waltz, Kenneth N. 1979. *Theory of International Politics*. Reading, MA: Addison-Wesley.
- Whytock, Christopher A. 2005. “A Rational Design Theory of Transgovernmentalism: The Case of E.U.–U.S. Merger Review Cooperation.” *Boston University International Law Journal* 23(1):1–53.

## Appendix

### *Proof of Proposition 1 (formal governance equilibrium)*

On the equilibrium path, strategy  $\sigma_F$  yields country  $j$  expected utility  $U_j(\sigma_F | \sigma_F) = (1-\beta)\alpha(f + \delta C_F) + \alpha\beta(-b_j + \delta C_F) + (1-\alpha)\delta((1-\theta)(\tau f + \delta C_F) + \theta(-\tau b_j + \delta C_F))$ , assuming that a  $1-\beta$  share of the  $f$  contingencies (and a  $1-\theta$

share as well in case of delegation) are attributed to  $j$ .  $C_F$  is the continuation value and is equal to  $U_i(\sigma_F|\sigma_F)$ , and thus  $C_F = -\frac{\alpha\beta b_j - \delta(\alpha-1)(\theta\tau b_j + f\tau(\theta-1)) + f\alpha(\beta-1)}{\delta^2(\alpha-1) + \alpha\delta(\beta-1) - \alpha\beta\delta + 1}$ . A one-period deviation (play  $x$  under any circumstances – call it  $X$ ) would yield  $U_j(X|\sigma_F) = (1-\beta)\alpha(f + \delta C_F) + \beta\alpha\delta^2 C_F + (1-\alpha)\delta((1-\theta)(\tau f + \delta C_F) + \theta\delta^2 C_F)$ . We have  $U_j(\sigma_F|\sigma_F) \geq U_j(X|\sigma_F) \Rightarrow b_j \leq \frac{f\delta(\theta\delta(1-\alpha) + \alpha\beta)(\alpha(1-\beta) + (1-\theta)\tau\delta(1-\alpha))}{((1-\alpha)\delta(1+\delta\theta) + \alpha\beta\delta + 1)(\alpha\beta + \theta\tau\delta(1-\alpha))} = \bar{b}_j$ . Note that  $\left\{\frac{\partial \bar{b}_j}{\partial \delta}, \frac{\partial \bar{b}_j}{\partial \alpha}\right\} > 0$ ,  $\left\{\frac{\partial \bar{b}_j}{\partial \theta}, \frac{\partial \bar{b}_j}{\partial \beta}\right\} < 0$ , and  $\frac{\partial \bar{b}_j}{\partial \tau} > 0$  if  $\beta > \theta$ , negative otherwise.

Off the equilibrium path, assuming there was an observed defection, punishing delays the resumption of cooperation by one period. Failure to punish is clearly worse, as it yields an inferior payoff  $-b_j$  during the first period and then postpone resuming cooperation by two periods.

*Proof of Proposition 2 (revelation)*

The timing of the revelation game is as follows: after Nature has chosen  $j$ 's type,  $i$  plays first by choosing a negotiation cost  $m$  and then  $j$  either rejects the proposed  $m$  and ends the game, or accepts it and the game carries on indefinitely, with, in each round, Nature drawing a contingency and both players simultaneously choosing an action.

(A) I first consider the choice between having a formal regime and no regime at all. I start by identifying the  $j$  type that, in a formal regime, is indifferent between cooperating and systematically defecting – call it  $\hat{b}_j$ . It is the type for which  $U_j(\text{accept, coop} | m) = U_j(\text{accept, defect} | m)$ , that is,  $-m + C_F = -m + C_D$ , with  $C_F$  the continuation value for formal cooperation defined above and  $C_D$  the continuation value solving  $C_D = \alpha(1-\beta)(f + \delta C_D) + \alpha\beta\delta^2 C_D + (1-\alpha)\delta((1-\theta)(\tau f + \delta C_D) + \theta(\delta^2 C_D))$ , assuming that a  $1-\beta$  and a  $1-\theta$  shares of the  $f$  contingencies are attributed to  $j$ . Solving for  $\hat{b}_j$  yields

$$\hat{b}_j = \frac{f\delta(\theta\delta + \alpha\beta - \theta\alpha\delta)(\alpha(1-\beta) + \tau\delta(1-\alpha)(1-\theta))}{(\delta(1-\alpha)(\theta\delta + 1) + \alpha\beta\delta + 1)(\alpha\beta + \theta\tau\delta(1-\alpha))}.$$

Initiator  $i$  also pays ratification cost  $m$ , its strategy is thus to maximize its value for the game by choosing the smallest  $m$  that is accepted by all the  $j$  types with cooperation cost  $b_j < \hat{b}_j$  and rejected by all the others. This means that the minimization must meet the following constraint: all  $j$  types with  $b_j < \hat{b}_j$  accept the offer and cooperate, whereas all other types turn down the offer. In other words,  $\hat{b}_j$  must satisfy  $U_j(\text{reject} | m) = U_j(\text{accept, coop} | m) \Rightarrow 0 = -m + (1-\beta)\alpha(f + \delta C_F) + \alpha\beta(-\hat{b}_j + \delta C_F) + (1-\alpha)\delta((1-\theta)(\tau f + \delta C_F) + \theta(-\tau\hat{b}_j + \delta C_F))$ . Solving for  $m$  yields

$$m^* = \frac{f(\alpha(1-\beta) + \tau\delta(1-\alpha)(1-\theta))}{1 - \delta(\alpha(1-\delta)(1-\beta) + \delta(1-\theta)(1-\delta)(1-\alpha))}.$$

Comparative statics  $\frac{\partial m}{\partial \alpha} > 0$  indicates that the less imprecision, the higher the ratification costs have to be to separate cooperators ( $b_j < \hat{b}_j$ ) from defectors ( $b_j > \hat{b}_j$ ). Imprecision mitigates the importance of the ratification costs – a realistic implication of the model.

Initiator  $i$ 's expected payoff is equal to  $U_i(m, \text{coop} | \text{accept}, \text{coop}) = -m + C_G$ , with  $C_G$  the value that solves  $C_G = \beta\alpha(f + \delta C_G) + (1 - \beta)\alpha(-b_i + \delta C_G) + (1 - \alpha)\delta(\theta\tau f + (1 - \theta)\tau(-b_i) + \delta C_G)$ . I assume that  $i$  can keep searching at no cost for a partner willing to spend  $m^*$ .

(B) Consider then the choice between having, this time, an informal regime and no regime at all. In an informal regime, all  $j$  types with a cost of cooperation  $b_j < 0$  cooperate, while all others defect. Moreover, those who defect prefer to accept the regime to turning it down, with the result that all  $j$  types accept the regime and no information about their cost of cooperation is revealed.

*Proof of Proposition 3 (informal governance equilibrium)*

Informal governance with outside option is a subgame perfect Nash equilibrium. To show this, assume that  $i$  and  $j$  each have an outside option in each successive round and so do the outside options, *ad infinitum*, irrespective of how many options they exercised in prior rounds.

I first consider the initiator's move. Being a harmony type ( $b_i < 0$ ),  $i$  will never defect; it will either cooperate or, if it becomes convinced that it is facing a partner that will defect, exercise its outside option. Its expected payoff, therefore, is  $C_O = pC_1 + (1 - p)C_2$ , with  $p$  the probability that  $j, k, l, \dots$  cooperate ( $k, l, \dots$  being  $i$ 's successive possible outside options); with  $C_1 = \beta\alpha(f + \delta C_1) + (1 - \beta)\alpha(-b_j + \delta C_1) + (1 - \alpha)\delta C_1$ , the long-term payoff for having found a cooperator; with  $C_2 = \beta\alpha(\delta C_O) + (1 - \beta)\alpha(-b_i + \delta C_2) + (1 - \alpha)\delta C_2$ , the long-term payoff for having found a non-collaborator; and with  $C_O$  the long-term payoff for exercising the outside option in the case where  $j$  revealed itself to be a non-cooperator by defecting. Substituting and calculating yields a value for  $C_O$  that is a function of  $p = \frac{b_j + B_j}{2B_j}$  and thus of  $\tilde{b}_j$ , with  $\tilde{b}_j$  the  $j$  type that is indifferent between cooperating and defecting with  $i$ .

To determine  $\tilde{b}_j$ , I consider  $j$ 's calculation in its move with  $i$ .  $j$  expects from cooperation  $C_3 = (1 - \beta)\alpha(f + \delta C_3) + \beta\alpha(-b_j + \delta C_3) + (1 - \alpha)\delta C_3$ , and from defection  $D = (1 - \beta)\alpha(f + \delta D) + \beta\alpha(\delta S_j) + (1 - \alpha)\delta D$ , with  $S_j$  its continuation value in the case where his defection leads  $i$  to break the cooperation and force  $j$  to turn toward its outside option (call this outside option  $h$ ). The value of  $\tilde{b}_j$  is the  $b_j$  that makes  $C_3 = D$ , with  $D$  a function of  $S_j$ .

To calculate  $S_j$ , I consider player  $j$  in a round after it defected and broke up with  $i$  and is about to play with  $h$ . Having defected with  $i$ , a fortiori  $j$  is

expected to defect with  $h$ . The reason is that  $j$  expects less from cooperating with  $h$ , whose type is unknown, than it did with  $i$ , whose type was known to be harmony.<sup>22</sup> And so  $j$  fears losing  $h$  less than it did losing  $i$ . As a result, the  $b_j$  type that is indifferent between cooperating and defecting with  $h$  (call it  $\tilde{b}_j$ ) is lower than the  $b_j$  type that was indifferent between cooperating and defecting with  $i$  (which I called  $\tilde{b}_j$ ). And so, after  $j$  rejects the initiator's first-period offer,  $h$  updates its belief about  $j$ 's type;  $h$  believes that  $b_j$  is uniformly distributed on  $[b_j, B_j]$ . Thus, anticipating  $j$  to defect,  $h$  (and if need be, all  $j$ 's subsequent outside options) only cooperates if it is a harmony type. And so  $j$ 's expected payoff is  $S_j = \frac{1}{2}D_2 + \frac{1}{2}S'_j$  with  $D_2 = (1-\beta_2)\alpha_2(f + \delta D_2) + \beta_2\alpha_2(\delta S_j) + (1-\alpha_2)\delta D_2$ , the value for defecting with a harmony type, and with  $S'_j = (1-\beta_2)\alpha_2(\delta S_j) + \beta_2\alpha_2(\delta S_j) + (1-\alpha_2)\delta S'_j$ , the value for defecting with a PD type. Note that  $\alpha_2$  and  $\beta_2$  are the terms of the arrangement between  $j$  and  $h$ . To simplify notation, I assume all contracts to be identical:  $\alpha_2 = \alpha$  and  $\beta_2 = \beta$ ; as a result,  $D_2 = D$ . Calculating  $S'_j$  and substituting it along with  $D$  into  $S_j$  and then  $S_j$  into  $D$ , yields  $D = f\alpha(\beta-1) \frac{-2\delta + \alpha\delta + 2}{(\delta-1)(-2\delta + \alpha\delta + \alpha\beta\delta + 2)}$ . Consequently, the  $b_j$  that makes  $C_3 = D$ , is

$$\tilde{b}_j = \frac{\delta(1-\beta)\alpha f}{2(1-\delta) + \delta\alpha(1+\beta)},$$

and thus  $p = \frac{1}{2} \left( 1 + \frac{1}{B_j} \frac{\delta(1-\beta)\alpha f}{2(1-\delta) + \delta\alpha(1+\beta)} \right)$ . All there is left to do is to substitute the value of  $p$  into player  $i$ 's expected payoff for informal governance with an

outside option  $C_O = \frac{(ab_i - \alpha\beta b_i) \frac{\frac{1}{2B_j} \left( B_j + \frac{f\alpha\delta - f\alpha\beta\delta}{\alpha\delta - 2\delta + \alpha\beta\delta + 2} \right) - 1}{\alpha\beta\delta - \delta + 1} - \frac{1}{2B_j(\delta-1)} \left( B_j + \frac{f\alpha\delta - f\alpha\beta\delta}{\alpha\delta - 2\delta + \alpha\beta\delta + 2} \right) (f\alpha\beta - ab_i + \alpha\beta b_i)}{\alpha\beta\delta \frac{\frac{1}{2B_j} \left( B_j + \frac{f\alpha\delta - f\alpha\beta\delta}{\alpha\delta - 2\delta + \alpha\beta\delta + 2} \right) - 1}{\alpha\beta\delta - \delta + 1} + 1}$ .

The expression is used in the simulation reported in Figure 2.

<sup>22</sup> This is the feature that makes this game solvable.