International Theory (2012), 4:1, 133–163 © Cambridge University Press, 2012
doi:10.1017/S1752971911000194

How uncertainty about outside options impedes international cooperation

JOHANNES URPELAINEN
Assistant Professor, Department of Political Science, Columbia University, New York, NY, USA

E-mail: ju2178@columbia.edu

International cooperation often requires costly policy adjustments. States may worry, however, that such adjustments weaken their outside options, and thus reduce their bargaining power. How does uncertainty about the effects of policy adjustments on outside options influence the depth of cooperation that states can achieve? My game-theoretic analysis shows that uncertainty about outside options is an obstacle to deep cooperation. If states agree on deep cooperation, they have to compensate vulnerable states with weak outside options for their losses. Under uncertainty, states that are not vulnerable have an incentive to falsely claim that they are vulnerable (i) to avoid a side payment or (ii) to obtain compensation for being vulnerable. The result holds even if the added value of deep cooperation would be large enough to fully compensate the losers. In equilibrium, the more vulnerable state sometimes offers a side payment to the less vulnerable one. More broadly, the analysis reveals a new international cooperation problem and provides a new rationale for costly signaling mechanisms and delegation to international organizations.

Keywords: international cooperation; bargaining; incomplete information; contract theory; power asymmetry

International cooperation consists of reciprocal, mutually profitable policy adjustments (Keohane 1984). The policy adjustments needed for cooperation may, however, increase a state's dependence on continued cooperation. If trade liberalization makes companies invest in costly export facilities, a reversal to autarky will be very costly because these export facilities are no longer useful. Similarly, when each member of a military alliance specializes in the production of certain armaments to capitalize on economies of scale, these countries will be unable to defend themselves if alliance cooperation break down. Owing to this problem, cooperation often fails because state *A* fears that it may become dependent on, and vulnerable to exploitation by, state *B* (Yarbrough and Yarbrough 1992; McLaren 1997; Cooley and Spruyt 2009; Lake 2009a; Rector 2009).

In a perfect world, this problem would be solved as follows. First, if the collective benefits of deep cooperation are large enough, the winners would compensate the losers through side payments or issue linkage (Sebenius 1983; Carrubba 1997; Barrett 2001). Second, if the cost to the losers is too high, states would deliberately choose shallow cooperation and reduce the vulnerability of disadvantaged members to exploitation (Yarbrough and Yarbrough 1992; Ikenberry 1998; Bradley and Kelley 2008; Cooley and Spruyt 2009; Rector 2009). In each case, the mode of international cooperation, deep or shallow, would maximize joint surplus (Keohane 1984; Koremenos, Lipson, and Snidal 2001).[1]

In reality, the mode of international cooperation emerges from deeply divisive negotiations among rational egoists. Similar to crisis bargaining, states have incentives to exploit private information for distributional gain (Fearon 1995; Wagner 2000). The resulting bargaining may restrain the scope of admissible solutions or even produce a cooperation failure. How does uncertainty about *ex post* outside options influence the depth of international cooperation? In this article, I offer a game-theoretic analysis of this problem. The analysis offers a twofold contribution to international cooperation theory. First, it reveals a previously unseen information revelation problem. Second, it uncovers a new strategic benefit of several common features of international institutions, such as costly signaling mechanisms (Kydd 2000) and delegation to third parties (Abbott and Snidal 1998).

In the model, two states can engage in international cooperation. They can choose between 'deep' and 'shallow' modes of international cooperation. While shallow cooperation has limited effects on a state's outside options, deep cooperation requires relational investments that may endogenously undermine a state's bargaining position by reducing the value of its outside options (McLaren 1997; Lake 2009a; Rector 2009). Following Keohane and Nye (1977), a state can either be 'vulnerable' or 'nonvulnerable' to international cooperation: the outside options of nonvulnerable states are influenced, but not nearly as much as those of vulnerable states.[2] The problem, then, is that nonvulnerable states may exploit vulnerable states due to asymmetric changes in their outside options.

[1] In this article, the depth of international cooperation is defined in terms of the size of the policy adjustments that states must implement (Lake 2009a; Rector 2009). It is *not* defined in terms of institutional design, such as delegation to an international organization (Abbott and Snidal 1998; Koremenos, Lipson, and Snidal 2001).

[2] This is a slight modification of the sensitive-vulnerable distinction that they proposed. Their primary interest was in highlighting the fact that while some states are only sensitive to international shocks, others are vulnerable because they cannot adjust. My emphasis is on vulnerability to renegotiation in a bargaining setting, so a more pronounced distinction between vulnerable and nonvulnerable states is more appropriate.

I begin by demonstrating that concerns regarding outside options impede deep cooperation between nonvulnerable and vulnerable states when (i) the true frequency of nonvulnerable states in the population of all states is high and (ii) asymmetric deep cooperation effectively undermines the bargaining position of vulnerable states. Next, I characterize the consequences of establishing a simple information revelation mechanism, such as a negotiation forum that allows states to formally communicate with each other regarding the form and nature of future cooperation (Morrow 1994; Fearon 1998).

I find that a simple information revelation mechanism does not allow deep cooperation between nonvulnerable and vulnerable states. To see why, suppose instead that a nonvulnerable–vulnerable pair will engage in deep cooperation, but the vulnerable state is compensated for its losses. Let state *A* be nonvulnerable and consider its incentive to reveal information. If state *B* is also nonvulnerable, state *A* should falsely claim vulnerability to obtain compensation. If state *B* is vulnerable, state *A* should also falsely claim vulnerability to avoid giving compensation.

Why is deep international cooperation so difficult? While previous scholarship has emphasized enforcement issues and distributional conflict, the present article adopts a different approach (Krasner 1991; Downs, Rocke, and Barsoom 1996; Fearon 1998). Even if international cooperation is enforceable and side payments costless, uncertainty about *ex post* outside options complicates asymmetric deep cooperation when the effects of relational investments are asymmetric. The result holds even if deep international cooperation produces a very large joint surplus, so that under complete information it would be easy to compensate the losers for mutual gain. The analysis also shows that if states with different vulnerabilities do engage in cooperation, the more vulnerable state sometimes offers a side payment to the less vulnerable state.

These findings have notable implications for the design of international institutions (Keohane 1984; Koremenos, Lipson, and Snidal 2001). If information revelation is an important rationale for international institutionalization, as Morrow (1994) and Fearon (1998) maintain, it is important to understand how information revelation mechanisms influence the nature of cooperation. When deep cooperation is very difficult to attain due to concerns about exploitation, simple information revelation mechanisms can allow vulnerable states to cooperate with each other, whereas asymmetric deep cooperation remains unfeasible due to incentives to misrepresent.

Can states overcome the limitations of simple information revelation mechanisms? After formally characterizing the strategic problem, I discuss the merits of two possible solutions to it. First, *costly signaling* mechanisms

could allow vulnerable states to reveal their type (Fearon 1997). However, the feasibility of this solution depends on the existence of actions that carry a low cost for vulnerable states but a high cost for nonvulnerable states. Given that nonvulnerable states are generally better able to implement policy adjustments than vulnerable states, the existence of such signaling mechanisms remains in doubt. Second, *delegation* to an international organization could mitigate the renegotiation problem, and thus dispose of the need to reveal information in the first place (Abbott and Snidal 1998; Bradley and Kelley 2008).[3] However, this solution is only available when the cooperation problem is such that states can fully delegate policy adjustments to an international organization. Even though the GATT/WTO has a dispute resolution mechanism, for example, it cannot directly force member states to adjust their trade policies (Keohane 1986; Bagwell and Staiger 1999).

While scholars have recognized that international institutions admit various forms of *ex post* bargaining, from dispute settlement in the shadow of international courts to informal governance by major powers, this issue has not been investigated from the perspective of *ex ante* bargaining over the nature of cooperation (Rosendorff 2005; Stone 2008; Rector 2009; Gilligan, Johns, and Rosendroff 2010). The present article shows how this can be done.

International cooperation and uncertainty about *ex post* outside options

In this article, I examine the distributional consequences of *international cooperation*, or mutually profitable policy adjustments, such as tariff reductions or capital account liberalization (Keohane 1984; Downs, Rocke, and Barsoom 1996). By *deep cooperation*, I refer to major policy adjustments that are difficult to reverse, and thus weaken the outside options of each state. By *shallow cooperation*, I refer to minor policy adjustments that can be reversed somewhat more easily, so that they have less forceful effects on outside options. In trade negotiations, for example, deep cooperation could comprise tariff reductions that destroy entire industries, while shallow cooperation would represent more modest tariff reductions (McLaren 1997; Rector 2009).

The key difference between deep and shallow cooperation is that the former increases a state's dependence on continued cooperation more than the latter. When international cooperation requires (partially) irreversible

[3] Under delegation, I also include the design of voting rules that reduce the importance of raw bargaining power (Zamora 1980; Maggi and Morelli 2006).

investments, differences in asset specificity may produce asymmetric effects on the outside options of states relative to the status quo ante.[4] In turn, these asymmetries will endogenously change the bargaining outcome. If policy adjustments make one state more dependent on international cooperation than the other, the top dog may exploit the asymmetry of *ex post* outside options to demand a renegotiation of cooperation rules. The newly dependent state cannot reject this demand because the cost of cooperation failure is too high *ex post* (Yarbrough and Yarbrough 1992; Cooley and Spruyt 2009; Lake 2009a; Rector 2009).[5]

Asset specificity plays an important role in international cooperation theory. Yarbrough and Yarbrough (1992) argue that in trade cooperation the nature of irreversible adjustments and investments are a central determinant of contractual governance. McLaren (1997) formally demonstrates that trade liberalization may entrap small states into dependence because they undergo irreversible adjustments. Lake (2009a) argues that the asymmetric dependence of major and minor powers on security arrangements drives the design of military alliances. Rector (2009) argues that federations can help states achieve 'contrived symmetry', and thus avoid renegotiation problems between unequal partners due to asymmetric changes in outside options. Cooley and Spruyt (2009) develop a full theory of incomplete contracting in international relations based on the idea of asset specificity, and the contracting problems that result from the endogenous bargaining effects thereof.

My model adds uncertainty about *ex post* outside options. Each state is either *nonvulnerable* or *vulnerable* to deep cooperation. Deep cooperation has smaller effects on the outside options of nonvulnerable than vulnerable states. Each state's type is private information, so cooperation may require information revelation. The key to understanding the model is to recognize that the uncertainty pertains only to outside options. Both nonvulnerable and vulnerable states obtain the same benefits from successful cooperation, so the only difference between these two types is how much their outside options weaken as a result of the decision to cooperate.[6]

[4] For asset specificity in contract theory, see Grossman and Hart (1986).

[5] This conceptualization of the depth of cooperation is orthogonal to institutional design. Depth of cooperation pertains to policy adjustments, as opposed to the nature of decision making, such as delegation to international courts (Haftel and Thompson 2006; Bradley and Kelley 2008) or the use of voting rules (Zamora 1980; Maggi and Morelli 2006). While this delimitation prevents me from fully considering such questions as 'legalization' or 'insulation' (Goldstein, Johns, and Rosendroff 2000; Koremenos, Lipson, and Snidal 2001; Voeten 2008), it allows me to separately analyze depth of cooperation and institutional design.

[6] In reality, a state's payoff from cooperation is often positively correlated with the weakening of outside options. For example, small states both need cooperation and are more

This form of incomplete information resembles the notion of 'uncertainty about preferences' in rational design (Koremenos, Lipson, and Snidal 2001). Most of the literature focuses on uncertainty about the state of the world, so the theoretical logic of uncertainty about *ex post* outside options – the consequences of stopping previously initiated cooperation – is underdeveloped. International cooperation theorists agree that many observable indicators of asset specificity exist: large countries are generally less vulnerable to renegotiation than small countries (Lake 2009b; Rector 2009), diverse and developed economies are less vulnerable than 'banana republics' (Grieco 1982), and so on. At the same time, there are several reasons to believe that complete information is also not a particularly realistic assumption.

First, uncertainty about *ex post* outside options may relate to the availability of other cooperation partners. If states *A* and *B* consider cooperation, state *A*'s ability to renegotiate the deal depends crucially on whether it could instead cooperate with state *C*. If state *A* does not have alternative cooperation partners, any policy adjustments that it implements to facilitate cooperation are futile if cooperation with state *B* fails. But if state *A* has other cooperation partners, the policy adjustments that facilitate cooperation with state *B* can also help achieve cooperation with state *C*. For example, if exporters in state *A* invest in new production facilities and state *B* suddenly raises tariffs, this is less of a problem when state *A* can divert the exports to state *C*.

The strategic problem of arms control measures on the Korean peninsula may prove instructive. If North and South Korea agreed to cooperate on arms control, the effect of reduced military expenditures would depend critically on the willingness of their main allies – China and the United States, respectively – to continue investing in their defense in emergency situations. If one of the two countries can expect more support from its main ally than the other, any arms control measures between the two would have asymmetric effects on their outside option (military conflict). Regarding North Korea, for example, Scobell (2004, v) concludes that 'North Korea … seems destined to remain heavily dependent on China for morale support and material assistance'. For South Korea, it would also be particularly hard to estimate the extent to which North Korea can rely on China's support given the secretive nature of Chinese and North Korean foreign policy. Thus, uncertainty characterizes the bargaining effects of potential arms control measures on the Korean peninsula.

vulnerable to exploitation than large states (McLaren 1997). Yet this correlation is never perfect, so analyzing uncertainty about *ex post* outside options while holding the gains from cooperation constant is substantively meaningful.

Second, uncertainty about *ex post* outside options may pertain to the consequences of the required policy adjustments. Even though state *B* understands that state *A* must implement policy adjustments for cooperation, the effect of such policy adjustments on state *A*'s outside options is not necessarily clear. In trade cooperation, exporters and importers in state *A* have private information on the cost and reversibility of investments in production facilities, while the government of state *A* is best capable of evaluating the importance of their political support (Downs and Rocke 1995; Rosendorff 2005). Similarly, the military of a state probably usually has private information on the state's ability to defend itself without an alliance.

Such uncertainty is a common problem in defense cooperation, as illustrated by Tucker's (1991, 112–18) analysis of the failure of the European air fighter project, 1983–85. Economies of scale and a history of successful defense cooperation gave a clear economic rationale for a joint technology program, especially Franco-German collaboration. Several high-level meetings on initiating collaboration were held in the period of 1983–85, but the two regional leaders failed to agree on the distribution of labor and gains. Emboldened by the West German aircraft industry's success in the last decade, Bonn demanded an equal partnership. The French, however, were unwilling to offer such terms to West Germany without a substantial side payment. One of the key reasons cited was the possibility that the program would allow the growing West German aircraft industry to develop an independent ability to produce fighter aircraft to Third World markets, and thus undercut France's position in the global defense markets. This would, of course, not be a problem so long as France and West Germany continued to cooperate on fighter production. But the French worried that West Germany's *ex post* outside option would improve so much that it would '"take the technology and run" to improve [its] competitive positions in Third World export markets' (Tucker 1991, 115). The model that I develop emphasizes the importance of such uncertainty concerning *ex post* outside options, and thus situates this outcome within a larger class of cooperation problems.

Even in successful cases of cooperation, uncertainty about *ex post* outside options often looms large. Consider, for instance, the 1988 free trade agreement between the United States and Canada. While the aggregate economic gains from further trade liberalization were never in doubt, McLaren (1997, 404) notes that many commentators in Canada expressed concern about the possible effects of a trade treaty on the country's ability to bargain with the larger United States: 'an extreme dependence on US trade led to a strong demand for safeguards against potential opportunistic action by the United States'. While supporters of

the trade agreement saw little reason for concern, the opponents argued that a treaty would make Canada's economic dependence on the United States so pronounced that Washington could dictate Canada's economic and even foreign policy. As one activist opposing the treaty put it, the United States could threaten to terminate the agreement 'as a threat against Canada in order to force agreement to certain concessions … [o]nce Canada had embarked on an industrial conversions process, cancellation would simply create another disruption to our national economy'.[7] This case illustrates the possibly dangerous effects of deep cooperation on *ex post* outside options, in spite of large aggregate gains from integration. Notably, the negotiation outcome is also consistent with my equilibrium results: 'most of the concessions in the negotiations were made by Canada', or the more concerned partner (McLaren 1997, 404).

More generally, when should one expect uncertainty about *ex post* outside options? One relevant covariate is the extent of previous interactions between states within an issue area. It seems plausible that the established members of the EU would have more information about each other than recently democratized accession candidates from Eastern Europe. In this case, one would expect that the informational dynamics of the model would be more relevant for accession negotiations than for routine discussions among current member states. Similarly, one would expect that if two states begin to negotiate a trade agreement for the first time in history, such as the European Community and South Africa in 1995, only a year after the collapse of the apartheid regime, incomplete information is a major issue. But subsequent negotiation rounds between the EU and South Africa probably feature less incomplete information.

If states expect international cooperation to carry distributional implications, how can the anticipated losers be enticed to participate? One solution is to insist on deep international cooperation – despite major distributional implications – but compensate the losers up front (Carrubba 1997; Alesina, Angeloni, and Etro 2005; Barberà and Jackson 2006; Konstantinidis 2008). For example, powerful states could offer foreign aid to weak states in exchange for economic reform that allows the expansion of international business. The main advantage of this approach is that it maximizes the joint surplus from international cooperation.

Empirically, it is more common for states to simply limit the depth of cooperation (Zamora 1980; Abbott and Snidal 1998; Maggi and Morelli 2006;

[7] From a 1988 pamphlet printed in Ottawa by Marjorie M. Bowker, 'What will the free trade agreement mean to you and to Canada? An independent analysis based on the actual text of the Canada/US free trade agreement'. Cited in McLaren (1997, 404).

Bradley and Kelley 2008; Hathaway 2008; Rector 2009). If vulnerable states fear renegotiation, reducing the scope of an international agreement or the degree of international delegation can mitigate the problem. For example, if McLaren (1997) is correct in arguing that international trade cooperation amplifies dependence, an obvious solution is to reduce the breadth and depth of liberalization.

The strategic implications of uncertainty about *ex post* outside options have not been recognized in previous research on international institutions. A burgeoning literature on international unions mostly focuses on welfare maximization (Alesina, Angeloni, and Etro 2005; Barberà and Jackson 2006), while the literature on institutional contracting in international politics emphasizes the functionalist benefits of reduced transaction costs (Lake 1996; Weber 2002; Cooley and Spruyt 2009) and the importance of enforcement (Maggi and Morelli 2006; Urpelainen 2010). Another relevant literature concerns information revelation. Previous studies have shown that international institutions can help states cooperate under uncertainty about the nature of the underlying game (Morrow 1994) and emphasized the importance of 'costly signaling' in reassurance (Kydd 2000). Several scholars have also emphasized the informational role of international institutions (Steinberg 2002; Thompson 2006; Chapman 2007). However, these studies do not model the effects of uncertainty about *ex post* outside options.

The model

In the model, states $i = A, B$ negotiate on international cooperation. Each state is either 'nonvulnerable' or 'vulnerable'. Nonvulnerable states can engage in deep cooperation by implementing such domestic policy adjustments that their outside options are only slightly weakened. If vulnerable states are to engage in deep cooperation, however, they must implement domestic policy adjustments that greatly reduce the value of their outside options. Uncertainty surrounds the type of each state.

Since I focus on information revelation, I apply the revelation principle from mechanism design (Myerson 1979; Banks 1990). If some negotiation protocol allows information revelation, one may also construct a direct mechanism wherein each state simply says 'nonvulnerable' or 'vulnerable' to produce an identical equilibrium outcome.

Sequence of moves. Consider the following moves:

(1) Each state i says 'nonvulnerable' or 'vulnerable'.
(2) Each state i proposes a level of cooperation $x_i \in \{\phi, S, D\}$ and a side payment $t_i \in [0, \infty)$.

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

(3) The game continues as follows:

(a) If either state i rejects cooperation, $x_i = \phi$, or the two governments disagree on the form of cooperation or the side payment, $x_A \neq x_B$ or $t_A + t_B \neq 0$, the game ends.
(b) Otherwise a round of Nash bargaining over cooperation benefits ensues. In the Nash bargaining game, each state i issues a demand $d_i \in [0, 1]$. If $d_A + d_B \leq 1$, bargaining succeeds. If $d_A + d_B > 1$, bargaining fails.

States can first reveal information. Subsequently, both states must agree on international cooperation, or the status quo prevails. The statement $x_i = S$ is interpreted as state i demanding shallow cooperation, while the statement $x_i = D$ implies that state i demands deep cooperation. The statement $x_i = \phi$ implies that state i rejects cooperation.

They can also use side payments to redistribute the gains from cooperation. If international cooperation follows, the two states must bargain over the distribution of gains.

The side payment can be a monetary transfer or an issue linkage (Tollison and Willett 1979; Sebenius 1983). Here t_i denotes the side payment to state i from state j. Thus, t_A is the side payment from A to B, while t_B is the side payment from B to A. If $t_A + t_B \neq 0$, cooperation fails because the two states disagree on the side payment. With $t_A + t_B = 0$, state A receives a side payment whenever $t_A > 0$ and gives one whenever $t_A < 0$.[8]

Payoffs. The outcome of the game depends on the proposals for level of cooperation (x_A, x_B). If $x_i = \phi$ for at least one state i or $x_i \neq x_j$, international cooperation fails.[9] Without loss of generality, the payoff to each state $i = A, B$ is normalized to zero. As Keohane (1984) argues, international cooperation is a voluntary exchange between sovereign states, so each state i can prevent it simply by saying no. In that case, the status quo prevails.

Suppose instead that $x_A = x_B = S$. Now both states i propose shallow international cooperation, S, so that neither state is required to invest substantial resources prior to the cooperation act. Without loss of

[8] The side payment does not carry a transaction cost for simplicity, but adding such a cost would not compromise any of the main results. I also abstract away from the problem of credibly committing to side payments. The credibility of side payments may itself cause a commitment problem, and thus form an obstacle to deep cooperation. Notably, such complications are consistent with the general thrust of my argument.

[9] If $x_i = S$ and $x_i = D$, it would be equally plausible to opt for shallow cooperation. This change would not influence the equilibrium analysis below, as the focus is on the possibility that both states demand deep cooperation, $x_A = x_B = D$.

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

generality, I assume that the collective benefit from shallow cooperation is worth 1 payoff unit. Below, I will discuss how states bargain to divide this payoff.

Suppose now that $x_A = x_B = D$. Both states i propose deep cooperation, D, so that they must incur a substantial investment cost. Net of the investments, I assume that the collective benefit from deep cooperation is $B > 1$ units. This assumption states that deep cooperation produces more surplus than shallow cooperation.[10]

Following the renegotiation literature, I assume that the states engage in a single round of Nash bargaining to distribute the gains following the exogenous revelation of Bayesian types (Nash 1950; Grossman and Hart 1986; Koremenos 2001). The technical details of the Nash bargaining game are fully described in the mathematical appendix. To solve the bargaining game, I apply the Nash Bargaining Solution (Nash 1950; Binmore, Rubinstein, and Wolinsky 1986). The Nash Bargaining Solution is a natural equilibrium concept: it captures the notion that changes in outside options influence the bargaining position of states. In equilibrium, each state receives exactly $\frac{1}{2}$ of the cooperation benefit over the payoff that it obtains from the non-cooperative outside options.[11] This is why outside options influence the equilibrium distribution of gains.

If the value of the outside option for state i is Q_i and the cooperation benefit is V, the payoff to state i will be

$$Q_i + \tfrac{1}{2}\left(V - Q_A - Q_B\right). \tag{1}$$

In this expression, $V - Q_A - Q_B$ is the net value of cooperation that the two states divide between themselves. It depends on the outside option: as the value of the outside option, Q_i, increases by one unit, the bargaining payoff to state i increases by $\frac{1}{2}$ units. Changes in outside options, therefore, have effects on the bargaining outcome (Grossman and Hart 1986; Rector 2009).

What are the outside option payoffs Q_i? Under shallow cooperation, they are fixed at $\underline{X} < 0$. To reduce notation, they do not depend on a state's type: the outside options of both nonvulnerable and vulnerable states are weakened somewhat. If the states agree on deep cooperation, I assume that the relational investments disproportionately weaken the value of the outside option for a vulnerable state (McLaren 1997; Rector 2009;

[10] I will also discuss informally the possibility that $B < 1$ in the main text.

[11] If cooperation does not produce a collective benefit, each state i obtains the status quo payoff. In the present model, this outcome is not possible on the path of play because the cooperation surplus is always positive.

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

Table 1. Equilibrium cooperation payoffs (notwithstanding side payments). Of these payoffs, all are positive except $\frac{1}{2}\left(B-X^{\text{nonv}}+X^{\text{vul}}\right)$

	Foreign type			
	Shallow cooperation		Deep cooperation	
Own type	Vulnerable	Nonvulnerable	Vulnerable	Nonvulnerable
Vulnerable	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}B$	$\frac{1}{2}\left(B-X^{\text{nonv}}+X^{\text{vul}}\right)$
Nonvulnerable	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}\left(B+X^{\text{nonv}}-X^{\text{vul}}\right)$	$\frac{1}{2}B$

Slapin 2009). Specifically, suppose the status quo payoff Q_i is X^{nonv}, where $X^{\text{nonv}} < \underline{X}$, for a nonvulnerable state but X^{vul}, where $X^{\text{vul}} < X^{\text{nonv}}$, for a vulnerable state.

I assume $X^{\text{vul}} + \frac{1}{2}(B-2X^{\text{vul}}) > 0$ but $X^{\text{vul}} + \frac{1}{2}(B-X^{\text{vul}}-X^{\text{nonv}}) < 0$, so that vulnerable states prefer *not* to engage in deep cooperation with nonvulnerable states. Vulnerable states may hesitate to engage in deep cooperation because they worry that nonvulnerable states will renegotiate the agreement, exploiting the asymmetric changes in the outside options. For ease of exposition, these payoffs are summarized in Table 1.

Information. While each state i is aware of its own type, vulnerable or nonvulnerable, it ascribes a commonly known prior probability $p^{\text{nonv}} \in (0, 1)$ to the other state j being nonvulnerable. The prior probability p^{nonv} can thus be thought of as the frequency of nonvulnerable states in the unobservable population of states. Empirically, p^{nonv} could be estimated based on observable indicators such as country size. If $p^{\text{nonv}} \to 1$ or $p^{\text{nonv}} \to 0$, prior information is relatively accurate; if $p^{\text{nonv}} \to \frac{1}{2}$, it is very inaccurate.

Strategies. A strategy for state i has two elements. First, it maps her Bayesian type into a proposed cooperation level x_i and side payment t_i. Second, upon exogenous revelation of Bayesian types, it maps them, and the agreed upon cooperation level, into a demand d_i in the Nash bargaining game. In the equilibrium analysis, I focus on pure strategies.

Analysis: complete information

To characterize the equilibrium conditions, I begin with the simple case of complete information. Under complete information, the communication stage can be safely ignored, as there is no uncertainty. As a simplifying equilibrium refinement, I minimize the net side payment, $|t_i^*|$. All proofs can be found in the mathematical appendix.

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

In the complete information version of the game, the type of each state $i = A, B$ is commonly known. What about the bargaining payoffs? Consider first shallow cooperation. We have

$$\underline{X} + \tfrac{1}{2}(1-2\underline{X}) = \tfrac{1}{2} > 0 \tag{2}$$

for each state i. Substantively, shallow cooperation allows each state to obtain a positive payoff with certainty because asymmetric weakening of outside options is not possible.

Consider now deep cooperation. We have

$$X_i + \tfrac{1}{2}(B-X_i-X_j) \gtreqless 0. \tag{3}$$

This expression is positive, except when a vulnerable state encounters a nonvulnerable state. Deep cooperation with a nonvulnerable state is harmful to a vulnerable state because exploitation is possible.

What about the proposed cooperation levels? To find the equilibrium at this stage, it suffices to consider how a nonvulnerable–vulnerable pair behaves: symmetric pairs have no incentive to deviate from deep cooperation, $x_A^* = x_B^* = D$. The joint surplus from international cooperation is B if $x_A^* = x_B^* = D$ and 1 if $x_A^* = x_B^* = S$. With $B > 1$, deep cooperation should be selected whenever possible. The side payments t_A^*, t_B^* must be such that both the nonvulnerable and vulnerable type prefer international cooperation. What is the optimal arrangement such that neither state prefers to exit?

Proposition 1. Consider complete information. Let state A be nonvulnerable and state B vulnerable. In equilibrium, deep cooperation, $x_A^* = x_B^* = D$, ensues. The side payment to the vulnerable state is $t_B^* = -\frac{1}{2}(B + X^{\text{vul}} - X^{\text{nonv}}) > 0$.

Given deep cooperation, the payoff from bargaining to a vulnerable state is negative. Thus, a side payment must be given to the vulnerable state as compensation.

Under complete information, the conventional wisdom on surplus maximization holds. If the cost of shallow cooperation is too high, the nonvulnerable state can simply compensate the vulnerable state. For example, European states could agree on economic integration while compensating recalcitrant members for their losses (Carrubba 1997).[12]

[12] If the cost to the nonvulnerable state were higher than the benefit to the vulnerable state, so that $B < 1$, shallow cooperation would follow. No side payment would be necessary, so $t^* = 0$. For example, many regional trading arrangements induce relatively minor changes in trade policies and do not contain legally binding dispute resolution provisions (McCall Smith 2000).

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

Analysis: incomplete information

I now consider the incomplete information version of the game. Following Morrow (1994), different equilibria of the game can be thought of as alternate institutional arrangements for negotiations. My solution concept is the perfect Bayesian equilibrium in pure strategies.[13] For each state type, an equilibrium of the game is a strategy vector containing a statement ('nonvulnerable' or 'vulnerable'), a proposed level of cooperation x_i^*, a side payment offer t_i^*, and an action in the Nash bargaining game. All strategies must be optimal given equilibrium beliefs, and equilibrium beliefs must accord with behavior on the equilibrium path. The game has two important classes of equilibria that are considered in the main text, and some additional equilibria are presented in the mathematical appendix.

I assume that the states cannot freely renegotiate cooperation upon full information revelation, except as specified in the bargaining game. In other words, they can credibly commit to the mechanism. This assumption implies that states must possess some commitment capacity despite international anarchy (Axelrod and Keohane 1985). If states also suffered from credibility problems in regard to the mechanism itself, the problem of achieving deep cooperation would be aggravated.

Babbling equilibria. In babbling equilibria, states cannot condition international cooperation on prior communication. They decide on cooperation based on their prior beliefs. For example, states could negotiate a preferential trading agreement 'under anarchy', in the absence of institutionalized information revelation (Axelrod and Keohane 1985). As Morrow (1994, 403) writes, a 'babbling equilibrium is analogous to the absence of a working regime … both players must proceed on their own'.

Separating equilibria. In separating equilibria, states engage in discussions that allow them to reveal information. As a concrete illustration, Morrow (1994, 411) provides the case of the Med Plan, an international institution to reduce pollution in the Mediterranean, where an 'epistemic community' of scientists from different nations helped governments to communicate information regarding their preferences. Similarly, Steinberg (2002) argues that institutionalized negotiations, and consensus voting in particular, within the GATT and the World Trade Organization could be thought of as an information revelation mechanism.

When multiple equilibria exist, I use an *ex ante* selection criterion, so that the expected payoff to a state (prior to type assignment) is maximized. While not uncontroversial, this assumption appears plausible given that

[13] The game also has equilibria in mixed strategies, but they do not produce any additional substantive insight.

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

international institutions help states achieve good bargaining outcomes in the long run (Morrow 1994; Fearon 1998). Assuming that states are sometimes nonvulnerable, and sometimes vulnerable, within an issue area, they have incentives to design institutions that maximize expected payoffs.

To foreshadow, I find that deep cooperation between vulnerable and nonvulnerable states is possible without information revelation when each state is vulnerable with high probability and deep cooperation does not weaken a vulnerable state's outside options much. Otherwise, only pairs of nonvulnerable states can engage in deep cooperation, and pairs of vulnerable states must choose shallow cooperation instead.

Babbling equilibrium

Any game with cheap talk has 'babbling' equilibria (Crawford and Sobel 1982; Morrow 1994). Babbling equilibria capture the possibility that states cooperate despite failure of information revelation. In these equilibria, side payments are precluded, $t_A^* = t_B^* = 0$, because they have no effect on incentives because they cannot be conditioned on messages.

The game has three babbling equilibria that involve some cooperation by all states.[14] First, all states propose deep cooperation, $x_i^* = D$. Second, nonvulnerable states propose deep cooperation, $x_i^* = D$, while vulnerable states propose shallow cooperation, $x_i^* = S$. Third, all states propose shallow cooperation, $x_i^* = S$. I omit the third equilibrium because it does not allow any deep cooperation.[15]

Nonvulnerable states have nothing to lose from deep cooperation, so they are always willing to propose it. Under a babbling equilibrium, two possibilities must therefore be considered. First, nonvulnerable and vulnerable states are both willing to propose deep cooperation without information. In this equilibrium, each state i always proposes $x_i^* = D$. Second, only nonvulnerable states are willing to engage in deep cooperation without information (while vulnerable states engage in shallow cooperation). In this equilibrium, nonvulnerable states play $x_i^* = D$ and vulnerable states play $x_i^* = S$.[16]

What is required for both nonvulnerable and vulnerable states to engage in deep cooperation? The equilibrium payoff is unambiguously lower for the vulnerable state given the asymmetric changes in outside

[14] Equilibria in which some states do not cooperate at all are in weakly dominated strategies, so I omit them throughout.

[15] It can also be excluded as Pareto-inefficient whenever B and $1-p^{\text{priv}}$ are low enough.

[16] In addition, equilibria exist such that all states engage in shallow cooperation or no state ever cooperates. I omit a discussion of such equilibria, as they do not shed light on the possibility of deep cooperation.

options from deep cooperation. If a vulnerable state can propose deep cooperation, a nonvulnerable state can also do so. Thus, it suffices to examine the conditions under which the 'weakest link', the vulnerable state, has no incentive deviate. In equilibrium, a vulnerable state can expect the following payoff:

$$p^{\text{nonv}}\left(X^{\text{vul}} + \tfrac{1}{2}\left(B - X^{\text{vul}} - X^{\text{nonv}}\right)\right) + (1 - p^{\text{nonv}})\tfrac{1}{2}. \tag{4}$$

In this expression, p^{nonv} is the probability that it faces a nonvulnerable state and $1-p^{\text{nonv}}$ is the probability of facing another vulnerable state. The first term of this expression is the (negative) expected payoff from cooperation in case the foreign state is nonvulnerable, and exploitation follows. The second term is the (positive) expected payoff given that the foreign state is also vulnerable.

This expression is a positive value as long as (i) p^{nonv} is low enough or (ii) X^{vul} is not too negative. When this expression is negative, only nonvulnerable states can engage in deep cooperation in a babbling equilibrium. Thus, the critical determinants are (i) the probability that a state is nonvulnerable and (ii) the effect of deep cooperation on the vulnerable state's outside options. Information revelation, then, is only an issue if vulnerable states have much to worry about. If these conditions fail to hold, deep cooperation can occur without information revelation or side payments.

Separating equilibrium

In separating equilibria, the strategy of state i must be such that it truthfully says 'nonvulnerable' if it is nonvulnerable – otherwise it says 'vulnerable'. For this equilibrium, if both states i say either 'nonvulnerable' or 'vulnerable' in equilibrium, so that the situation is symmetric, each state i proposes deep international cooperation, $x_i^* = D$, and there is no side payment, $t_A^* = t_B^* = 0$. If states truthfully reveal information, and neither has a bargaining advantage, they should engage in deep cooperation.[17]

For a nonvulnerable–vulnerable pair, my equilibrium refinement implies that the smallest admissible side payment is selected.[18] In equilibrium, the side payments could go either way. Nonvulnerable states could compensate vulnerable states (compensation for losers), or alternatively vulnerable states could compensate nonvulnerable states (might over right). I let t^{vul} denote the net side payment from the nonvulnerable to the vulnerable state. Correspondingly, let $t^{\text{nonv}} = -t^{\text{vul}}$ denote the side payment from the vulnerable to the nonvulnerable state.

[17] Technically, I reject equilibria in weakly dominated strategies.

[18] This restriction is not needed for the main results, but it pins down a unique equilibrium.

Table 2. Payoffs from asymmetric cooperation in different separating equilibria

	Value	
Payoff	Asymmetric deep	Asymmetric shallow
$V^{\text{vul,nonv}}$	$X^{\text{vul}} + \frac{1}{2}\left(B - X^{\text{vul}} - X^{\text{nonv}}\right) < 0$	$\frac{1}{2}$
$V^{\text{vul,vul}}$	$\frac{1}{2}B > 0$	$\frac{1}{2}$
$V^{\text{nonv,vul}}$	$X^{\text{nonv}} + \frac{1}{2} < (B - X^{\text{vul}} - X^{\text{nonv}}) > 0$	$\frac{1}{2}$
$V^{\text{nonv,nonv}}$	$\frac{1}{2}B > 0$	$\frac{1}{2}$

The 'asymmetric deep' column gives the equilibrium payoff when a nonvulnerable–vulnerable pair engages in deep cooperation. The 'asymmetric shallow' column gives the equilibrium payoff when they engage in shallow cooperation.

Let $V^{x,y}$ denote the cooperation payoff – not including the side payments – to a state of type y when it encounters a state of type x and the two states announce, truthfully or not, different types. Intuitively, this expression characterizes the payoffs in case the two states make different announcements in the beginning of the game. For example, $V^{\text{nonv,vul}}$ denotes the payoff to a vulnerable state from encountering a nonvulnerable state, and one of them says 'nonvulnerable', while the other announces 'vulnerable'. I sometimes have to use this generic expression instead of the exact payoff because multiple equilibrium candidates exist, and the payoff varies across them. The value of $V^{x,y}$ in different contingencies is summarized in Table 2.

Vulnerable state incentives. When is state i willing to truthfully reveal its type? When state i selects a statement, it has yet to learn the other state's type. To begin with, suppose state i is actually vulnerable. If it truthfully says 'vulnerable', its expected payoff is

$$p^{\text{nonv}}\left(V^{\text{vul,nonv}} + t^{\text{vul}}\right) + (1-p^{\text{nonv}})\tfrac{1}{2}B. \tag{5}$$

The first term is the expected payoff from encountering a nonvulnerable state, while the second term is the expected payoff when the foreign state is also vulnerable. The first term may or may not be positive, depending on the side payment, whereas the second term is always positive because two vulnerable states benefit from deep cooperation.

Misrepresentation yields

$$p^{\text{nonv}}\left(X^{\text{vul}} + \tfrac{1}{2}\left(B - X^{\text{vul}} - X^{\text{nonv}}\right)\right) + (1-p^{\text{nonv}})\left(V^{\text{vul,nonv}} - t^{\text{vul}}\right). \tag{6}$$

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

Again, the first term captures the possibility that the foreign state is nonvulnerable, while the second term captures the possibility that it is vulnerable. Note, however, that the first term is now always negative, while the second term's sign is ambiguous. The first term is negative because by pretending to be nonvulnerable a vulnerable state would basically engage in deep cooperation if the foreign state is nonvulnerable. This would ultimately allow exploitation.

The vulnerable state has no incentive to misrepresent whenever the side payment to the nonvulnerable state is not too high. Unless the nonvulnerable state must offer a large side payment, it has no incentive to misrepresent information. By falsely claiming nonvulnerability, it would either (i) forgo compensation or (ii) engage in deep cooperation with a nonvulnerable state. Neither option, of course, is beneficial to a vulnerable state.

Nonvulnerable state incentives. What about the nonvulnerable state? Telling the truth yields

$$p^{\text{nonv}}\tfrac{1}{2}B + (1-p^{\text{nonv}})\left(V^{\text{nonv,vul}}-t^{\text{vul}}\right). \tag{7}$$

The first term reflects the payoff from facing another nonvulnerable state, and the second term the payoff from facing a vulnerable state. Pretending to be vulnerable yields

$$p^{\text{nonv}}\left(V^{\text{nonv,nonv}} + t^{\text{vul}}\right) + (1-p^{\text{nonv}})\left(X^{\text{nonv}} + \tfrac{1}{2}\left(B-X^{\text{nonv}}-X^{\text{vul}}\right)\right). \tag{8}$$

The first term is the payoff from facing a nonvulnerable state, and the second term is the payoff from facing a vulnerable state.

Deep asymmetric cooperation. Is deep cooperation, $x_A^* = x_B^* = D$, possible for a nonvulnerable–vulnerable pair? With $X^{\text{vul}} + \frac{1}{2}(B-X^{\text{nonv}}-X^{\text{vul}}) < 0$, the vulnerable state must be compensated for deep cooperation. As shown in the appendix, the nonvulnerable state must therefore lie: the only effect of claiming strength is to lose value of the side payment, t^{vul}, regardless of the type of the other state. Only vulnerable (nonvulnerable) states obtain (give) side payments in equilibrium. By truthfully announcing nonvulnerability, a nonvulnerable state incurs a loss of t^{vul} without any corresponding benefits. Thus, truthful information revelation is *not* possible when nonvulnerable–vulnerable pairs are expected to engage in deep cooperation.

Shallow asymmetric cooperation. I have so far found that shallow asymmetric cooperation is necessary for information revelation by nonvulnerable states. Consider now $x_A^* = x_B^* = S$ instead, so that asymmetric cooperation is shallow, and a side payment t^{vul}. From conditions (7) and (8),

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

the incentive compatibility condition for a nonvulnerable state is derived in the mathematical appendix:

$$t^{\text{nonv}} \geq \frac{(1 + p^{\text{nonv}})\left(\frac{1}{2}X^{\text{vul}} - \frac{1}{2}X^{\text{nonv}} - 1\right) - \frac{1}{2}(B-1)}{(2p^{\text{nonv}} - 1)}. \tag{9}$$

To incentivize the nonvulnerable state to reveal its type, the vulnerable state must compensate the nonvulnerable state.

Consider the intuition. First, recall that side payments are not feasible for symmetric state pairs, so any transfers are only possible for a nonvulnerable–vulnerable pair. When asymmetric cooperation is shallow, it is *not* necessary to compensate the vulnerable state because the equilibrium payoffs are $\frac{1}{2}$. Thus, from an *ex post* perspective it would be enough to offer a zero payoff. However, what about the nonvulnerable state's incentive to reveal information? It need not be compensated at this point, but the only way the nonvulnerable state can be induced to reveal its nonvulnerable type is to promise a side payment for doing so. In a separating equilibrium, the nonvulnerable state must thus obtain a side payment from the vulnerable state. This accords with the outcome of the 1988 trade negotiations between the United States and Canada, with Canada making most of the concessions even though it had much more potential to lose from the deal than the much larger United States.

According to expression (9), a nonzero side payment is needed whenever nonvulnerable states are the minority, $p^{\text{nonv}} < \frac{1}{2}$. Whenever they are the majority, $p^{\text{nonv}} \geq \frac{1}{2}$, a zero side payment is also admissible because the nonvulnerable state has no incentive lie in any case. When p^{nonv} is high, a nonvulnerable state understands that it will probably meet another nonvulnerable state. In this case, it prefers to truthfully reveal its type because falsely claiming vulnerability would prevent mutually profitable deep cooperation.

The vulnerable state must be willing to make the side payment whenever it does face a nonvulnerable state. It gains $\frac{1}{2}$ from shallow cooperation in equilibrium, so we need $t^{\text{nonv}} \leq \frac{1}{2}$. If this condition is not met, the vulnerable state would rather give up shallow cooperation than provide the side payment. In total, we need

$$\frac{1}{2} \geq t^{\text{nonv}} \geq \frac{(1 + p^{\text{nonv}})\left(\frac{1}{2}X^{\text{vul}} - \frac{1}{2}X^{\text{nonv}} - 1\right) - \frac{1}{2}(B-1)}{(2p^{\text{nonv}} - 1)}. \tag{10}$$

The following proposition thus holds.

Proposition 2. Consider incomplete information. The separating equilibrium exists if and only if condition (10) holds for some t^{nonv}.

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

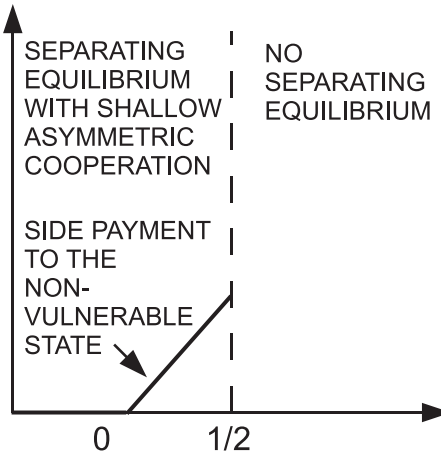

Figure 1 Existence of a separating equilibrium with shallow asymmetric cooperation and the side payment to the nonvulnerable state (y-axis) as a function of $\frac{(1+p^{\text{nonv}})\left(\frac{1}{2}X^{\text{vul}}-\frac{1}{2}X^{\text{nonv}}-1\right)-\frac{1}{2}(B-1)}{2p^{\text{nonv}}-1}$ (x-axis).

For a nonvulnerable–vulnerable pair, shallow cooperation occurs, $x_A^* = x_B^* = S$. The side payment is from the vulnerable to the nonvulnerable state:

$$t^{\text{nonv}} = \max\left\{\frac{(1+p^{\text{nonv}})\left(\frac{1}{2}X^{\text{vul}}-\frac{1}{2}X^{\text{nonv}}-1\right)-\frac{1}{2}(B-1)}{(2p^{\text{nonv}}-1)}, 0\right\} \geq 0.$$

Why is deep cooperation so difficult to attain in a separating equilibrium? Suppose a nonvulnerable state untruthfully claims vulnerability. Then, either the nonvulnerable state avoids the side payment (foreign state is vulnerable) or the nonvulnerable state actually *obtains* a side payment (foreign state is nonvulnerable). Regardless of what the foreign state's type is, the nonvulnerable state benefits from misrepresentation. In these circumstances, information revelation requires more than the information revelation mechanism in focus. The existence of a separating equilibrium with shallow asymmetric cooperation, as well as the equilibrium side payment, are illustrated in Figure 1.

This counterintuitive result explains why states often deliberately limit the scope of international cooperation. The alternative compensation strategy is not credible because it leads all states to claim that they will suffer from deep cooperation. States must either choose shallow cooperation, or find alternative solutions such as costly signaling and delegation.

If asymmetric shallow cooperation does occur, the side payment is from the vulnerable to the nonvulnerable state. While the side payment is unnecessary at this point, it is necessary to give the nonvulnerable state an *ex ante* incentive to reveal information. From a strategic perspective, such side payments are also rather realistic. While it may seem counterintuitive that a vulnerable state would compensate a nonvulnerable state, recall that *shallow cooperation* follows. Thus, the fact that the nonvulnerable state would have been in a stronger bargaining position under *deep cooperation* has no consequences for the resulting bargaining round.[19]

Implications for international institutions

In this section, I examine the broader implications of my findings. First, I explore the effect of information revelation on the expected level of international cooperation. Surprisingly, simple information revelation mechanisms may sometimes reduce the probability of cooperation. Second, I examine more sophisticated institutional designs to address the problem. While both costly signaling and delegation to international organizations can help, each suffers from notable limitations.

Information revelation and international cooperation

Deep cooperation is sometimes possible for all states even in the babbling equilibrium. Specifically, condition (4) must hold. In these circumstances, creating a simple information revelation mechanism paradoxically *reduces* the probability of international cooperation. To see why, recall that nonvulnerable–vulnerable pairs cannot engage in deep cooperation under information revelation.

Proposition 3. If condition (4) holds, information revelation mechanisms reduce the probability of deep cooperation.

This result qualifies Fearon's (1998) claim that international institutions generally facilitate bargaining by reducing incomplete information. When deep cooperation is not completely unattainable under anarchy, states may prefer to avoid international institutions that allow or require information revelation.

It does not follow, however, that all states prefer the babbling equilibrium. The babbling equilibrium sometimes allows nonvulnerable states to exploit

[19] A possible counterargument is that a level playing ground should prevent asymmetric side payments due to equal bargaining power. If the two states are indeed unable to commit to side payments, the separating equilibrium with shallow asymmetric cooperation would unravel. This possibility is entirely consistent with the substantive argument, as it underscores the difficulty of information revelation under uncertainty about *ex post* outside options.

vulnerable states under deep cooperation. While nonvulnerable states prefer not to design simple information revelation mechanisms, vulnerable states benefit from them because they can thus avoid deep cooperation with nonvulnerable states. A distributional conflict exists.

What if condition (4) fails, so that vulnerable states have no incentive to engage in deep cooperation without information revelation? Both nonvulnerable and vulnerable states may now have a mutual interest in information revelation. The outcome is not changed in the case of asymmetry, except that the nonvulnerable state obtains a side payment. In the case of symmetry, however, information revelation enables deep cooperation between the states.

Proposition 4. If condition (4) does not hold, information revelation mechanisms increase the probability of deep cooperation.

When deep cooperation is difficult under anarchy, information revelation mechanisms are useful because they allow vulnerable states to engage in deep cooperation with each other. In a babbling equilibrium, vulnerable states fail to engage each other because they worry that the foreign partner may actually be nonvulnerable (and thus capable of exploitation). Information revelation mechanisms allow vulnerable states to provide each other with credible evidence that they are not capable of exploitation.

Designing international institutions

So far, I have focused on analyzing the strategic problem in a basic strategic environment. In reality, states may also use more sophisticated institutional design features. In this section, I discuss informally two institutional designs, namely costly signaling and delegation, that could help states achieve deep cooperation between nonvulnerable and vulnerable states.

Recall that my focus here is on uncertainty. Thus, solutions such as 'contrived symmetry' or gradual and sequential policy adjustments are hardly useful in and of themselves (Cooley and Spruyt 2009; Rector 2009; Urpelainen 2011). The strategic problem pertains to information revelation: nonvulnerable states are unwilling to reveal their type, so any solutions that are based on assumed information vulnerability are useless.

Costly signaling. Given that nonvulnerable states have incentives to misrepresent, a costly signaling mechanism should allow a vulnerable state to send a signal that is too costly for a nonvulnerable state to send (Fearon 1997; Kydd 2000). If such a costly signaling mechanism is available, an equilibrium of the modified game would exist wherein (i) vulnerable states send a costly signal, while (ii) nonvulnerable states do not. By contrast, it is not possible to devise a costly signaling scheme wherein the nonvulnerable state sends the costly signal: in equilibrium,

side payments from asymmetric deep cooperation must be given to the vulnerable state, so nonvulnerable states would always have an incentive to mimic vulnerable states.

The costly signaling solution has some merit, but it depends on a problematic premise: a costly signal must exist that is *easier* for vulnerable states to send than nonvulnerable. Without this 'single crossing property', no separating equilibrium can exist in a costly signaling model. Given that nonvulnerable states are better equipped to engage in cooperation than the vulnerable states, finding such signals is not easy. A vulnerable state could implement domestic policies that trigger a response by third parties, such as international financial markets. But even then, one could worry that nonvulnerable states will mimic such domestic policies and fake vulnerability. Given these problems, the transaction costs of developing a functioning costly signaling scheme might be high.

Delegation. Another solution to the information revelation problem is delegation. If states can delegate policy implementation to an international organization, deep cooperation will not cause renegotiation problems: states cannot renegotiate contractual arrangements when they are being implemented by an international organization (Abbott and Snidal 1998; Ikenberry 2000). Thus, nonvulnerable states would no longer have incentives to withhold information regarding their type. In addition, an international delegation scheme could also facilitate the transfer of side payments, and thus produce a 'double dividend' for the concerned states.

Institutional voting rules offer a similar solution (Zamora 1980; Steinberg 2002; Maggi and Morelli 2006). States can delegate collectively to themselves by developing formal decision rules, such as supermajority or unanimity. If designed in an appropriate fashion, they may mitigate the renegotiation problem. Thus, they would have similar effects to international delegation to a genuinely external party.

While international delegation holds promise, it also suffers from two important limitations. First, international delegation is unhelpful when states are ultimately responsible for implementing domestic policies. Unless an international organization exists that is so powerful as to dictate domestic policy implementation, international delegation does not help. Even powerful institutions such as the GATT/WTO and the EU ultimately rely on state decisions to comply with multilateral rules (Tallberg 2002; Bagwell and Staiger 2005). In this sense, delegation may be more helpful for problems such as multilateral foreign aid (no domestic policy implementation required beyond initial contributions) than trade liberalization or pollution abatement (international delegation is useless unless states can credibly commit to domestic policies).

Second, international delegation is itself vulnerable to forms of renegotiation. As Fleck and Kilby (2006) have shown, even relatively powerful international organizations such as the World Bank are vulnerable to unilateral influence by powerful states. Stone (2008) finds that this is also true of the International Monetary Fund. These observations imply that while international delegation may sometimes mitigate the renegotiation problem, the fact remains that powerful states can exploit their bargaining capabilities by acting informally through the international organization that is supposed to prevent renegotiation. Unless the international organization can be perfectly 'insulated' from unilateral influence – something that is relatively difficult to achieve according to an empirical survey by Haftel and Thompson (2006) – international delegation cannot completely eradicate the renegotiation problem.

Interestingly, Rector's (2009, 135–48) analysis of the failure of East African regional integration provides an example of the limitations of delegation strategies. The regional hegemon, Kenya, had proposed to the smaller Uganda and Tanganyika that federal institutions would reduce its dominance, and thus prevent exploitation. However, the Ugandan government had little confidence in these solutions. As Rector (2009, 145) argues, Uganda did not believe that a bicameral legislature with a house and a senate would offset Kenya's bargaining gains while '[a] more subtle source of concern was the composition of the federal bureaucracy: if the civil service was oriented toward serving the urban industrial area around Nairobi, it would further enhance Kenya's exit option'.

In principle, international organizations could also directly provide information to states (Abbott and Snidal 1998). However, this solution requires that the international organization can learn strategically sensitive information concerning states' *ex post* outside options. Given the constraints that states place on the autonomy of international organizations (Haftel and Thompson 2006), these conditions are difficult to meet.

Conclusion

I have investigated how uncertainty about the effects of deep cooperation on outside options influences bargaining. Under uncertainty, it is difficult for states to credibly reveal information about their vulnerability. Shallow cooperation is useful to mitigate exploitation threats, even if side payments are costless and deep cooperation would allow states to realize joint gains.

These results have important implications for the design of international institutions. For one, if deep cooperation is unattainable under anarchy, even simple information revelation mechanisms can help vulnerable states achieve deep cooperation with each other, whereas asymmetric deep

cooperation remains difficult. Conversely, if deep cooperation is possible under anarchy, information revelation mechanisms may reduce the prior probability of deep cooperation. While such solutions as costly signaling or delegation to international organizations may help in some circumstances, I have also found that they are hardly a panacea.

While my formal model explicitly models the renegotiation problem, it omits several issues that hold promise as future extensions. One key issue is the role of domestic politics. As Martin (2000) and Dai (2005), domestic political institutions allow some states – especially democracies – to credibly commit to foreign policies. In the presence of domestic commitment devices, problems of renegotiation and exploitation may lose importance. The role of domestic politics thus offers a fruitful avenue for future research.

Although classical international cooperation theory and rationalist institutional design have illuminated the effects of international institutions on cooperation (Keohane 1984; Koremenos, Lipson, and Snidal 2001; Gilligan, Johns, and Rosendroff 2010), this research has not adequately theorized the strategically prior bargaining stage. This article shows that a careful consideration of the bargaining stage qualifies and refines the conventional wisdom on the role of uncertainty about *ex post* outside options, and incomplete information more generally, in international cooperation.

Acknowledgments

I thank Jennifer Kavanagh, Paul Poast, Chad Rector, the anonymous reviewers, and the editors of *International Theory* for their constructive comments and advice.

References

Abbott, Kenneth W., and Duncan Snidal. 1998. "Why States Act through Formal International Organizations." *Journal of Conflict Resolution* 42(1):3–32.

Alesina, Alberto, Ignazio Angeloni, and Federico Etro. 2005. "International Unions." *American Economic Review* 95(3):602–15.

Axelrod, Robert, and Robert O. Keohane. 1985. "Achieving Cooperation under Anarchy: Strategies and Institutions." *World Politics* 38(1):226–54.

Bagwell, Kyle, and Robert W. Staiger. 1999. "An Economic Theory of GATT." *American Economic Review* 89(1):215–48.

—— 2005. "Enforcement, Private Political Pressure, and the General Agreement on Tariffs and Trade/World Trade Organization Escape Clause." *Journal of Legal Studies* 34(3):471–513.

Banks, Jeffrey S. 1990. "Equilibrium Behavior in Crisis Bargaining Games." *American Journal of Political Science* 34(3):599–614.

Barberà, Salvador, and Matthew O. Jackson. 2006. "On the Weights of Nations: Assigning Voting Weights in a Heterogeneous Union." *Journal of Political Economy* 114(2):317–39.

Barrett, Scott. 2001. "International Cooperation for Sale." *European Economic Review* 45(10):1835–50.

Binmore, Ken, Ariel Rubinstein, and Asher Wolinsky. 1986. "The Nash Bargaining Solution in Economic Modelling." *Rand Journal of Economics* 17(2):176–88.
Bradley, Curtis A., and Judith G. Kelley. 2008. "The Concept of International Delegation." *Law and Contemporary Problems* 71(1):1–36.
Carrubba, Clifford J. 1997. "Net Financial Transfers in the European Union: Who Gets What and Why?" *Journal of Politics* 59(2):469–96.
Chapman, Terrence L. 2007. "International Security Institutions, Domestic Politics, and Institutional Legitimacy." *Journal of Conflict Resolution* 51(1):134–66.
Cooley, Alexander, and Hendrik Spruyt. 2009. *Contracting States: Sovereign Transfers in International Politics*. Princeton: Princeton University Press.
Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6):1431–451.
Dai, Xinyuan. 2005. "Why Comply? The Domestic Constituency Mechanism." *International Organization* 59(2):363–98.
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.
Fearon, James D. 1995. "Rationalist Explanations of War." *International Organization* 49(3):379–414.
—— 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41(1):68–90.
—— 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52(2):269–305.
Fleck, Robert K., and Christopher Kilby. 2006. "World Bank Independence: A Model and Statistical Analysis of US Influence." *Review of Development Economics* 10(2):224–40.
Gilligan, Michael J., Leslie Johns, and Peter B. Rosendorff. 2010. "Strengthening International Courts and the Early Settlement of Disputes." *Journal of Conflict Resolution* 54(1):5–38.
Goldstein, Judith, Miles Kahler, Robert O. Keohane, and Anne-Marie Slaughter. 2000. "Introduction: Legalization and World Politics." *International Organization* 54(3):385–99.
Grieco, Joseph M. 1982. "Between Dependency and Autonomy: India's Experience with the International Computer Industry." *International Organization* 36(3):609–32.
Grossman, Sanford J., and Oliver D. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94(4):691–719.
Haftel, Yoram Z., and Alexander Thompson. 2006. "The Independence of International Organizations: Concept and Applications." *Journal of Conflict Resolution* 50(2):253–75.
Hathaway, Oona A. 2008. "International Delegation and State Sovereignty." *Law and Contemporary Problems* 71(1):1–36.
Ikenberry, G. John. 1998. "Constitutional Politics in International Relations." *European Journal of International Relations* 4(2):147–77.
—— 2000. *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order after Major Wars*. Princeton: Princeton University Press.
Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton: Princeton University Press.
—— 1986. "Reciprocity in International Relations." *International Organization* 40(1):1–27.
Keohane, Robert O., and Joseph S. Nye. 1977. *Power and Interdependence: World Politics in Transition*. Boston: Little, Brown.

Konstantinidis, Nikitas. 2008. "Gradualism and Uncertainty in International Union Formation: The European Community's First Enlargement." *Review of International Organizations* 3(4):399–433.
Koremenos, Barbara. 2001. "Loosening the Ties that Bind: A Learning Model of Agreement Flexibility." *International Organization* 55(2):289–325.
Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. "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.
Kydd, Andrew H. 2000. "Trust, Reassurance, and Cooperation." *International Organization* 54(2):325–57.
Lake, David A. 1996. "Anarchy, Hierarchy, and the Variety of International Relations." *International Organization* 50(1):1–34.
—— 2009a. *Hierarchy in International Relations*. Ithaca: Cornell University Press.
—— 2009b. "Open Economy Politics: A Critical Review." *Review of International Organizations* 4(3):219–44.
Maggi, Giovanni, and Massimo Morelli. 2006. "Self-Enforcing Voting in International Organizations." *American Economic Review* 96(4):1137–58.
Martin, Lisa L. 2000. *Democratic Commitments: Legislatures and International Commitments*. Princeton: Princeton University Press.
McCall Smith, James. 2000. "The Politics of Dispute Settlement Design: Explaining Legalism in Regional Trade Pacts." *International Organization* 54(1):137–80.
McLaren, John. 1997. "Size, Sunk Costs, and Judge Bowker's Objection to Free Trade." *American Economic Review* 87(3):400–20.
Morrow, James D. 1994. "Modeling the Forms of International Cooperation: Distribution Versus Information." *International Organization* 48(3):387–423.
Myerson, Roger B. 1979. "Incentive Compatibility and the Bargaining Problem." *Econometrica* 47:61–74.
Nash, John F. 1950. "The Bargaining Problem." *Econometrica* 18(2):155–62.
Rector, Chad. 2009. *Federations: The Political Dynamics of Cooperation*. Ithaca: Cornell University Press.
Rosendorff, B. Peter. 2005. "Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure." *American Political Science Review* 99(3):389–400.
Scobell, Andew. 2004. *China and North Korea: From Comrades-in-Arms to Allies at Arm's Length*. Strategic Studies Institute.
Sebenius, James K. 1983. "Negotiation Arithmetic: Adding and Subtracting Issues and Parties." *International Organization* 37(2):281–316.
Slapin, Jonathan B. 2009. "Exit, Voice, and Cooperation: Bargaining Power in International Organizations and Federal Systems." *Journal of Theoretical Politics* 21(2):187–211.
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. 2008. "The Scope of IMF Conditionality." *International Organization* 62(4):589–620.
Tallberg, Jonas. 2002. "Paths to Compliance: Enforcement, Management, and the European Union." *International Organization* 56(3):609–43.
Thompson, Alexander. 2006. "Coercion Through IOs: The Security Council and the Logic of Information Transmission." *International Organization* 60(1):1–34.
Tollison, Robert D., and Thomas D. Willett. 1979. "An Economic Theory of Mutually Advantageous Issue Linkages in International Negotiations." *International Organization* 33(4):425–49.

Tucker, Jonathan B. 1991. "Partners and Rivals: A Model of International Collaboration in Advanced Technology." *International Organization* 45(1):83–120.
Urpelainen, Johannes. 2010. "Enforcing International Environmental Cooperation: Technological Standards Can Help." *Review of International Organizations* 5(4):475–96.
—— 2011. "Domestic Reform as a Rationale for Gradualism in International Cooperation." *Journal of Theoretical Politics* 23(3):400–27.
Voeten, Erik. 2008. "The Impartiality of International Judges: Evidence from the European Court of Human Rights." *American Political Science Review* 102(4):417–33.
Wagner, R. Harrison. 2000. "Bargaining and War." *American Journal of Political Science* 44(3):469–84.
Weber, Katja. 2002. "Hierarchy Amidst Anarchy: A Transaction Costs Approach to International Security Cooperation." *International Studies Quarterly* 41(2):321–40.
Yarbrough, Beth V., and Robert M. Yarbrough. 1992. *Cooperation and Governance in International Trade: The Strategic Organizational Approach*. Princeton: Princeton University Press.
Zamora, Stephen. 1980. "Voting in International Economic Organizations." *American Journal of International Law* 74(3):566–608.

Mathematical appendix

This mathematical appendix has four parts. First, I solve the Nash bargaining game, and thus derive the equilibrium payoffs given in Table 1. Second, I derive the incentive compatibility conditions for a separating equilibrium. Third, I examine some alternative equilibria that are omitted in the main text. Finally, I provide proofs for all propositions in the main text.

Nash bargaining

In the Nash bargaining game, each state i selects a demand $d_i \in [0, 1]$. If $d_A + d_B \leq 1$, then state i receives her disagreement payoff Q_i and a share d_i of the net surplus $B - Q_A - Q_B$. Otherwise state i receives her disagreement payoff Q_i. In the Nash equilibria of the game, the demands d_i^* must be best responses to each other. I use the Nash Bargaining Solution to select among Nash equilibria. By standard arguments elaborated in Nash (1950), the Nash Bargaining Solution must maximize the value of

$$\left(Q_A + d_A\left(B - Q_A - Q_B\right) - Q_A\right)\left(Q_B + d_B\left(B - Q_A - Q_B\right) - Q_B\right).$$

By the efficiency requirement, it has to be the case that $d_A^* + d_B^* = 1$. Thus, we have

$$\left(Q_A + d_A\left(B - Q_A - Q_B\right) - Q_A\right)\left(Q_B + (1 - d_A)\left(B - Q_A - Q_B\right) - Q_B\right).$$

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

Differentiating with respect to d_A, we obtain the symmetric solution $d_A^* = d_B^* = \frac{1}{2}$. Thus, the equilibrium demands are $d_A^* = d_B^* = \frac{1}{2}$ and the equilibrium payoff to state i is

$$Q_i + \tfrac{1}{2}(B - Q_A - Q_B).$$

In the main text, these equilibrium payoffs are given for various combinations of Bayesian types in Table 1.

Incentive compatibility conditions

Deep asymmetric cooperation. Consider first the possibility that a nonvulnerable–vulnerable pair engage in deep cooperation, $x_A^* = x_B^* = D$. For the vulnerable state to not deviate, we need

$$p^{\text{nonv}}\left(V^{\text{vul,nonv}} + t^{\text{vul}}\right) + (1-p^{\text{nonv}})\tfrac{1}{2}B \geq p^{\text{nonv}}\left(X^{\text{vul}} + \tfrac{1}{2}\left(B - X^{\text{nonv}} - X^{\text{vul}}\right)\right) + (1-p^{\text{nonv}})\left(V^{\text{vul,nonv}} + t^{\text{vul}}\right).$$

This condition can be rewritten as

$$p^{\text{nonv}}\left(V^{\text{vul,nonv}} - Y^{\text{vul}} + t^{\text{vul}}\right) + (1-p^{\text{nonv}})\left(\tfrac{1}{2}B - V^{\text{vul,nonv}} - t^{\text{vul}}\right) \geq 0,$$

where $Y^{\text{vul}} = X^{\text{vul}} + \frac{1}{2}(B - X^{\text{nonv}} - X^{\text{vul}}) < 0$. With $V^{\text{vul,nonv}} \geq Y^{\text{vul}}$ and $\frac{1}{2}B \geq V^{\text{vul,nonv}}$, this condition is guaranteed to hold as long as $t^* \geq 0$. Thus, the incentive compatibility condition is automatically met.

Consider now the nonvulnerable state. We need

$$p^{\text{nonv}}\tfrac{1}{2}B + (1-p^{\text{nonv}})\left(V^{\text{nonv,vul}} - t^{\text{vul}}\right) \geq p^{\text{nonv}}\left(V^{\text{nonv,nonv}} + t^{\text{vul}}\right) + (1-p^{\text{nonv}})Y^{\text{nonv}},$$

where $Y^{\text{nonv}} = X^{\text{nonv}} + \frac{1}{2}(B - X^{\text{nonv}} - X^{\text{vul}}) > 0$. This condition can be rewritten as

$$p^{\text{nonv}}\left(\tfrac{1}{2}B - V^{\text{nonv,nonv}} - t^{\text{vul}}\right) + (1-p^{\text{nonv}})\left(V^{\text{nonv,vul}} - Y^{\text{nonv}} - t^{\text{vul}}\right) \geq 0.$$

When $x_A^* = x_B^* = D$ for all possible type pairs in equilibrium, we have $V^{\text{nonv,nonv}} = \frac{1}{2}B$ and $V^{\text{nonv,vul}} = Y^{\text{nonv}}$. Thus, we can rewrite the condition as

$$-t^{\text{vul}}p^{\text{nonv}} - t^{\text{vul}}(1-p^{\text{nonv}}) \geq 0.$$

Clearly, this condition cannot be met for any t^{vul} 0. Thus, this separating equilibrium cannot exist.

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

Shallow asymmetric cooperation. Given shallow cooperation between nonvulnerable–vulnerable pairs, the analysis of incentive compatibility for the vulnerable state remains intact. What about the nonvulnerable type? Again, we need

$$p^{\text{nonv}}\left(\tfrac{1}{2}B-V^{\text{nonv,nonv}}-t^{\text{vul}}\right)+(1-p^{\text{nonv}})\left(V^{\text{nonv,vul}}-Y^{\text{nonv}}-t^{\text{vul}}\right)\geq 0.$$

Given that asymmetric pairs engage in shallow cooperation, we now have $V^{\text{nonv,nonv}}=\underline{X}+\tfrac{1}{2}(1-2\underline{X})=\tfrac{1}{2}$ and $V^{\text{nonv,vul}}=\underline{X}+\tfrac{1}{2}(B-2\underline{X})=\tfrac{1}{2}$. Thus, the condition can be written as

$$p^{\text{nonv}}\left(\tfrac{1}{2}B-\tfrac{1}{2}-t^{\text{vul}}\right)+(1-p^{\text{nonv}})\left(\tfrac{1}{2}-Y^{\text{nonv}}-t^{\text{vul}}\right)\geq 0.$$

Equivalently,

$$(1-2p^{\text{nonv}})t^{\text{vul}}\leq p^{\text{nonv}}\left(\tfrac{1}{2}B+Y^{\text{nonv}}-1\right)+\tfrac{1}{2}-X^{\text{nonv}}-\tfrac{1}{2}\left(B-X^{\text{nonv}}-X^{\text{vul}}\right)\Leftrightarrow$$

$$(1-2p^{\text{nonv}})t^{\text{vul}}\leq p^{\text{nonv}}\left(\tfrac{1}{2}X^{\text{vul}}-\tfrac{1}{2}X^{\text{nonv}}-1\right)+\tfrac{1}{2}-\tfrac{1}{2}B-\tfrac{1}{2}X^{\text{nonv}}+\tfrac{1}{2}X^{\text{vul}}\Leftrightarrow$$

$$(1-2p^{\text{nonv}})t^{\text{vul}}\leq (1+p^{\text{nonv}})\left(\tfrac{1}{2}X^{\text{vul}}-\tfrac{1}{2}X^{\text{nonv}}-1\right)-\tfrac{1}{2}(B-1).$$

With $X^{\text{vul}}<X^{\text{nonv}}$, the right side is negative. Thus, we obtain the incentive compatibility condition:

$$t^{\text{vul}}\leq\frac{(1+p^{\text{nonv}})\left(\tfrac{1}{2}X^{\text{vul}}-\tfrac{1}{2}X^{\text{nonv}}-1\right)-\tfrac{1}{2}(B-1)}{1-2p^{\text{nonv}}}.$$

To obtain a similar expression for $t^{\text{nonv}}=-t^{\text{vul}}$, simply multiply both sides by -1.

Alternative equilibria

One class of separating equilibria that I have yet to consider includes the possibility that deep cooperation results if and only if both states are nonvulnerable. In this equilibrium, it is immediate to establish that even without side payments, $t_i^*=t_j^*=0$, no state has an incentive to lie. Nonvulnerable states lose the benefits of deep cooperation when they lie, and vulnerable states do not want to engage in deep cooperation with nonvulnerable states in the absence of side payments. However, this equilibrium clearly does not solve the information revelation problem at hand. When only nonvulnerable states cooperate, the prior probability of cooperation is simply $p^{\text{nonv}}\,p^{\text{nonv}}$. This probability is strictly lower than in the equilibrium wherein all pairs engage in deep cooperation except the asymmetric nonvulnerable–vulnerable pair: $p^{\text{nonv}}\,p^{\text{nonv}}+(1-p^{\text{nonv}})(1-p^{\text{nonv}})$.

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press

What about a separating equilibrium wherein deep cooperation results if and only if both states are vulnerable? This equilibrium cannot exist because nonvulnerable states would obviously have incentives to pretend that they are vulnerable. Thus, this equilibrium can be completely ignored.

Proof of Proposition 1

The payoff to a vulnerable state from deep cooperation is $X^{\text{vul}} + \frac{1}{2}(B - X^{\text{vul}} - X^{\text{nonv}}) < 0$. Thus, $t_B^* \geq -\frac{1}{2}(B + X^{\text{vul}} - X^{\text{nonv}})$ is needed or the vulnerable state B deviates from $x_B^* = D$, preventing international cooperation. The joint surplus is weakly higher under deep cooperation because $B > 1$, so deep cooperation must be chosen. ■

Proof of Proposition 2

The existence of a separating equilibrium such that $x_A^* = x_B^* = S$ is proven in the main text. For the side payment, examine condition (10) and recall that $|t^{\text{nonv}}|$ must be minimized. ■

Proof of Proposition 3

In a babbling equilibrium, we have $x_A^* = x_B^* = S$ with probability zero. In a separating equilibrium, we have $x_A^* = x_B^* = S$ with probability $2p^{\text{nonv}}\,(1 - p^{\text{nonv}})$. ■

Proof of Proposition 4

In a babbling equilibrium, we have $x_A^* = x_B^* = H$ with probability $p^{\text{nonv}}\,p^{\text{nonv}}$. In a separating equilibrium, we have $x_A^* = x_B^* = H$ with probability $p^{\text{nonv}}\,p^{\text{nonv}} + (1 - p^{\text{nonv}})(1 - p^{\text{nonv}})$. ■

https://doi.org/10.1017/S1752971911000194 Published online by Cambridge University Press