

---

# Multilateralism, Bilateralism, and Exclusion in the Nuclear Proliferation Regime

Daniel Verdier

---

**Abstract** I use the nuclear proliferation regime to show that dyadic diplomacy is not necessarily incompatible with the building of a multilateral regime; bilateralism is not the opposite of multilateralism, but an efficient component thereof. Although this point will not be new to most students of institutions, no general rationale has so far been offered on the complementarity of bilateral and multilateral diplomacy. Starting from a characterization of proliferation as the result of a large number of prisoner's dilemmas played out between states engaged in local dyadic rivalries, I demonstrate that it is possible for the superpowers to design an optimal mix of threats and bribes in which states with low compliance costs join the regime on the terms of the multilateral treaty alone; states with intermediate compliance costs need additional customized incentives, delivered through bilateral agreements; and states with high compliance costs are not only left out of the regime but also punished for nonparticipation. I draw a few comparative statics that I systematically test on Nuclear Proliferation Treaty (NPT) membership data. I discuss the applicability of the model to the currency, trade, and aid regimes.

---

The nuclear proliferation regime has design features that are interesting to the institutionalist research agenda.<sup>1</sup> A key feature is the complementarity of a multilateral instrument—the Nuclear Proliferation Treaty (NPT)—with superpower bilateral diplomacy. The United States and the Soviet Union did more than negotiate a text that was agreeable to other countries; they also bribed and threatened some of their respective clients into signing the NPT. This study suggests that the NPT regime, along with several other important regimes, is neither bilateral nor multilateral, but a combination of both.

The bilateral-multilateral dichotomy is a staple of the International Relations literature. It is prominently featured in the grand debate pitting constructivists

My thanks to Alex Downes, Jing Han, Massimo Morelli, Alex Thompson, Ching-Jen Sun, the editors of *International Organization*: Lisa Martin, Emanuel Adler, and Louis Pauly, three anonymous reviewers for comments, and T. Marion Anderson and Byungwon Woo for research assistance. Research was financed by a grant from the Mershon Center for International Security Studies.

1. See Koremenos, Lipson, and Snidal 2001.

against realists. Ruggie<sup>2</sup> characterizes the last seventy years as a progressive shift away from bilateral and hegemonic regimes, in which relations among states are compartmentalized into dyads and obligations are specific to each dyad, toward multilateral regimes, characterized by equal treatment and universal participation. Realists such as Mearsheimer<sup>3</sup> counter that the spread of international institutions changes nothing in the way states have been interacting for centuries, for these institutions merely reflect the more powerful states' calculations.<sup>4</sup> Using a more fine-grained approach to institutional reality, students of the trade and aid regimes have found that multilateral and bilateral obligations actually coexist, but in neither case has this finding led to a formalization that would make it possible to predict the actual proportion of multilateral and bilateral instruments.<sup>5</sup> The institutionalist literature in general holds that multilateral institutions perform better in issue areas involving bargaining and coordination, while bilateral institutions do better in areas involving enforcement through retaliation, but this literature stops short of arguing that the two instruments are necessary complements.<sup>6</sup> In this article, I show that both bilateral and multilateral instruments are optimally used in the nonproliferation regime, and for reasons other than the bargaining-enforcement dichotomy.

I offer a general rationale for the complementarity of multilateral and bilateral diplomacy. Regimes are like contracts in which a group of founders purchase a good from signatory states in exchange for a price and/or the nonimposition of a sanction. Such contracts can be multilateral, offering uniform terms across eligible participants, or they can be bilateral, customizing the offer to reflect each state's peculiar circumstances. The multilateral strategy has the advantage of saving on transaction costs—there is only one deal and it is the same for everyone—but has the drawback of being expensive: participants are offered an incentive that is calculated to be sufficient to elicit the participation of the individual state that is burdened with the highest cost of compliance with the proposed regime. In contrast, the bilateral strategy allows the founders to save on resources by giving each state the incentive it needs to participate and no more. However, the bilateral strategy offers drawbacks as well, multiplying transaction costs, since a brand new contract has to be written for each new participant. Consequently, I argue, for states with a low cost of compliance, a multilateral instrument should suffice,

2. Ruggie 1992.

3. Mearsheimer 1994–95.

4. On the importance of relative power, see also Krasner 1991. For an argument that the U.S. hegemonic decline has caused a shift away from multilateralism to mini- and bilateralism, see Yarbrough and Yarbrough 1992.

5. See Rodrick 1996; and Milner 2006 on aid; and Pahre 2001 on trade.

6. See Oye 1986; and Koremenos, Lipson, and Snidal 2001, 783. See also Keohane 1984, 90; and Martin 1992. The bargaining-enforcement dichotomy is not the only rationale thought to favor multilateralism. Snidal (1991) argues that increasing the number of states mitigates the bargain's zero-sum properties. Fearon (1998, 298) cites multilateral bargains as a solution to the hold-out problem. Taking exception with the bargaining-enforcement dichotomization are Pahre 1994; and Lohmann 1997.

whereas for states with a higher cost of compliance, this multilateral instrument should be supplemented with bilateral deals.

A second notable feature of the NPT regime is nonparticipation. It took many years after the signing of the treaty in 1968 for most countries to join and to date four states still have not. Recent work on the causes for restricted membership point to the severity of the enforcement problem, in contrast with the severity of the distribution problem, which, instead, is thought to correlate with expanded membership.<sup>7</sup> I show that a severe distribution problem is a primary cause for restricted membership. States with a high cost of compliance are left out of the regime and, in some circumstances, even punished for not participating.

I model the NPT regime as a contract between nuclear and non-nuclear-weapon states. The former want the latter to forfeit the right to develop nuclear weapons, which the latter prefer to keep in order to face rivals in local arms races. I formally demonstrate the existence of an optimal combination of three instruments: the multilateral treaty itself, bilateral incentives in the form of security guarantees offered by the superpowers, and sanctions inflicted on nonparticipants. I then show, in accordance with the argument, the existence of a statistical relation between a country's marginal cost of compliance, the basket of positive and negative bilateral incentives that it received, and the decision to join the multilateral regime.

Last, because the model is custom-made for the nuclear proliferation regime, I assess its generalizability to currency, trade, and aid regimes. These cases provide a realistic assessment of the model's boundary conditions.

The article is organized as follows. I first present and model the nonproliferation regime. I then draw the comparative statics and confront the model against what is known of the nuclear nonproliferation regime. Last, I contrast the present approach with existing empirical studies of the NPT regime and sketch possible extensions of the model to the other regimes.

## The Nuclear Proliferation Regime

This section describes the three-instrument structure—multilateral, bilateral, and exclusion—of the proliferation regime. The parties to the regime belong to two distinct groups: on the one side are the countries forming the nuclear cartel, on the other are the non-nuclear-weapon states. The cartel initially included the United States, Britain, and the Soviet Union; France and China were given an open invitation to join the cartel as nuclear powers; the non-nuclear countries were all the other countries. Then the first three nuclear powers co-opted all exporters of nuclear fuel or technology within the so-called Nuclear Suppliers Group (NSG) to enforce

7. I am referring to conjectures M1 and M3 in Koremenos, Lipson, and Snidal 2001, 783–85. See also Oye 1986 on enforcement; and Snidal 1991 on distribution.

the treaty, with the result that many non-nuclear countries embraced the nuclear cartel countries' preferences, somewhat blurring the boundaries between roles.<sup>8</sup>

The regime arms the cartel with three incentives to deter non-nuclear countries from pursuing nuclear arms, the first being a carrot in the form of easy access to future scientific nuclear technology. The NPT explicitly provides for signatory states to gain access to the scientific spin-offs and technological by-products of nuclear explosions at production price, free of research and development costs.<sup>9</sup> A second incentive is the stick against nonparticipation, of which the effect is to worsen a nonmember's reservation value by threatening to block access to fissionable material and technology for the production of nuclear material for peaceful purposes.<sup>10</sup> A third component is the stick against cheating, required to combat moral hazard. The cartel cannot directly observe whether or not signatories comply, but has access only to imperfect reports drawn up every year by the International Atomic Energy Agency (IAEA) inspectors. To deter cheating, the treaty threatens to refer to the UN Security Council signatories who would be caught in breach of compliance.<sup>11</sup>

Noncartel members fall into three groups. A first group joined because the mix of carrots and sticks featured in the treaty provided just the right incentive to renounce the nuclear option. A well-known case is Sweden. Swedish premier Tage F. Erlander was very clear about the link between his country's renunciation of the nuclear program and the NPT. In 1963, he informed the United States that Sweden "had the possibility of developing its own weapons rather quickly and quite easily."<sup>12</sup> He added that there was no desire to do so, but that it would depend on whether the Federal Republic of Germany acquired nuclear weapons. He also hoped that "a great power agreement to end tests and put the lid on Nth [sic] countries will spare the Government from facing this issue."

Some countries would probably have given up the nuclear-weapon option for much less than offered by the NPT. While already at peace with their neighbors and entertaining no ambition in the nuclear energy field, signing the treaty would unlock access to U.S. nuclear fuel for the research reactors and other peaceful

---

8. When it was established in 1975, the NSG initially included the United States, the Soviet Union, the United Kingdom, France, Germany, Japan, and Canada, and has been since expanded to include thirty-eight more countries.

9. NPT, Article 5.

10. The treaty initially left open a loophole by which such transfers were possible if the nonsignatory state agreed to subject such material to safeguards which, being plant-specific, were not as comprehensive as those applying to signatory states. But the United States and the Soviet Union, which enjoyed a *de facto* monopoly on the supply of nuclear fuel until 1977, plugged that loophole by unilaterally requesting the same "full-scope" safeguards, thereby nullifying any advantage in not signing the NPT.

11. The nuclear-weapon states also agreed to pursue disarmament negotiations. The United States and the Soviet Union never delivered on this promise, the failure of the Comprehensive Test Ban Treaty being the latest installment. This failure has been a sticking point at every review conference. I leave it out of the analysis for now on grounds that this article does not try to account for all the politics of the NPT and pick it up again in the penultimate section when discussing interpretive approaches.

12. Cole 1997.

applications they had or were planning to build in the near future. This group included Mexico, Ecuador, Peru, Bolivia, Senegal, and Morocco—countries that were quick to ratify the NPT.

However, a second group of noncartel countries renounced nuclear weapons development only after receiving extra compensation or specific threats from the two superpowers. Countries that already benefited and kept benefiting from a superpower's security guarantee are the first to fall in this group. The North Atlantic Treaty Organization (NATO) extended the U.S. nuclear umbrella to Canada and the European allies and removed the need for indigenous nuclear weapons programs. A good illustration is Norway. Unlike Sweden, a non-NATO country that maintained a nuclear weapons program until 1972, Norway excluded that option as soon as it became a member of NATO in 1954.<sup>13</sup> East European states, with the possible exception of Romania, did not pursue nuclear weapons in exchange for the Soviet nuclear umbrella. As a result, with few exceptions, NATO and Soviet bloc European countries were ready to sign and ratify the NPT when it was offered to them.

Outside Europe, the United States had bilateral security arrangements with Japan, Australia, New Zealand, the Philippines, South Korea, and Taiwan, providing Washington with unique leverage on some of them. For instance, when South Korea announced it would buy a reprocessing facility from France, allowing it to separate plutonium from spent fuel that could be used in a nuclear bomb, the United States successfully blocked the deal by threatening to end further delivery of fuel and equipment to a U.S. reactor under construction and promising to maintain its military presence. A similar story unfolded in Taiwan in 1975–76.<sup>14</sup> Likewise, Egypt would have pursued the bomb had not it been for U.S. diplomatic initiatives, including reassurances that Israel “will not introduce” nuclear weapons into the Middle East.<sup>15</sup> Conversely, the drop in the credibility of Soviet security assurances in the late 1980s and the official withdrawal of all security guarantees by the new Russian government in 1991 prompted the acceleration of nuclear programs in North Korea, Iraq, and Iran.<sup>16</sup> U.S. bilateral diplomacy also played an important role in the denuclearization of South Africa and, consequently, the entire southern tip of the African continent. Active American mediation and financial assistance in the cases of Belarus, the Ukraine, and Kazakhstan were determinants in securing the membership of all three states in the NPT as non-nuclear-weapon states.<sup>17</sup>

Third and last, there is a final group of countries for which neither the lure of multilateralism nor bilateral wheeling and dealing were sufficient to elicit participation. Many states waited several decades before accessing to the treaty: Brazil, Argentina, Chile, South Africa and its neighbors, Cuba, and Algeria. Three states

13. Forland 1997.

14. On South Korea and Taiwan, see Yager 1985; and Albright and Gay 1998.

15. According to Levite 2002–3, 64.

16. Lessenberry 2005.

17. On South Africa, see Albright 1994; on ex-Soviet republics, see Drezner 1999.

have not signed the treaty yet—Israel, India, and Pakistan—while one, North Korea, signed it but then denounced it.

In sum, the nuclear proliferation regime presents a variety of instruments, each aimed at countries with different compliance costs: the NPT was targeted at countries with a low cost of compliance; superpowers' security guarantees and other offsetting policies supplemented the NPT for countries with intermediate costs; sanctions were imposed on a residual group of nonparticipants. I model this instrumental gradation in the next section.

## A Contractual Model

I model the nuclear proliferation regime as a contract between a principal (the nuclear cartel) and a large number of agents (the rest of the world). Although the principal-agent terminology has been mostly used in political science to characterize a delegation from, say, voter to elected representative, or from state to international organization, it is the standard terminology in the field of mechanism design for characterizing opposite roles assumed by parties to a horizontal transaction—seller and buyer, employer and employee, lender and borrower.<sup>18</sup>

I build on a model of monopolistic competition—a buyer-seller relation in a noncompetitive market under complete information. I leave the interaction among principals out of the model. The nuclear cartel is assumed to act like a single player by means of an iterated play that is left unmodelled. I first present the problem as it stands between two agents and the principal and then extend it to a large number of dyads.

### *Principal and One Dyad*

I derive the agent's utility from a model of the utility of nuclear weapons in a local arms race. States in general are motivated by a desire to influence other states, and the asymmetric possession of nuclear weapons provides them with such influence. Two agents are engaged in an arms race modeled as a prisoner's dilemma, with temptation payoff  $z - c$ , sucker payoff  $-z$ , cooperation payoff normalized to 0 and suboptimal payoff  $-c$ . The index of rivalry between the two countries is  $z$  with  $z$  an integer drawn in the  $[1, N]$  segment. At this point,  $z$  can be interpreted as an agent's compliance cost (this is the right that the principal covets). The cost for an agent of building nuclear weapons is  $c$  with  $c > 0$ .

Both  $z$  and  $c$  are common knowledge. I assume common knowledge because asymmetric information would detract from the primary interest of this study, which is the relative importance of bilateralism, multilateralism, and exclusion in the NPT regime.

18. For an introductory text to principal-agent modeling, see Laffont and Martimort 2002.

The principal values disarming by one agent at  $V > 0$  and nondisarming at 0. The principal confronts each agent separately with a contract that (1) offers payment  $a$  in exchange for an agreement to disarm, (2) threatens sanction  $d$  in the case of no agreement, and (3) threatens sanction  $s$  in the case where an agreement is reached but the agent is thought to have armed nevertheless. This three-incentive structure closely mirrors the text of the NPT, which, as was shown earlier, combines one carrot and two sticks.

The principal incurs a per-agent cost  $\varepsilon s$  for executing sanction  $s$  and a per-agent cost  $\rho d$  for executing sanction  $d$  with  $\varepsilon$  and  $\rho > 0$ . I ignore transaction costs for now, for they are of no relevance to this preliminary game.

I also assume that promises and threats are credible *ex ante*. Credibility is the result of a costly signaling game or a reputation game that is not modeled here because it is unnecessarily complex. Leaving the credibility-generating mechanism out of the game, however, is bound to lead to absurd results with respect to incentives that do not fall on the equilibrium path, such as sanctions. For instance, the principal could threaten a hellish punishment and extract compliance at no cost because she would be released from having to deliver on the sanction threat. Unless the cost of sanctioning figures on the equilibrium path, any sanction value goes. To circumvent this problem, I assume that the mere threat of sanctioning is costly. The principal incurs a per-agent cost  $\sigma s$  for threatening  $s$  and a per-agent cost  $\delta d$  for threatening  $d$  with  $\sigma$  and  $\delta > 0$ . The two threat costs capture, in much-reduced form, the *ex ante* costs that a principal must sink prior to uttering a threat in order to establish a reputation, often through the initial imposition of high-cost sanctions.<sup>19</sup>

Compliance with treaty obligations is not directly observable by the principal. The probability that an agent who cheats is not caught by the principal is  $q$ . For convenience, Appendix Table A1 lists the notation used in the model.

The principal moves first, then the agents sequentially decide to reject or accept the offer. At no loss to generality, I assume that agent 1 moves first, agent 2 second. If neither accepts, the game ends. If one rejects and the other accepts, the former is sanctioned and the other chooses whether to comply and be rewarded, or cheat and be rewarded with  $q$  probability and punished with probability  $1 - q$ . If both accept, then they simultaneously choose between complying and cheating (they do not observe each other's choice) and are either rewarded or punished accordingly. The game tree is drawn in Figure 1.

A strategy for the principal specifies the  $(a, d, s)$  regime she proposes. A strategy for agent 1 is the mapping  $(a, d, s) \rightarrow \{A, R\}$  and  $(a, d, s) \times \{A\} \times \{A, R\} \rightarrow \{C, S\}$  specifying for every possible regime whether to accept or reject and, after accepting and for every possible decision of agent 2, whether to comply or

19. Admittedly, the per-agent costs, sunk each time, are but a fairly remote approximation of the reputational costs, sunk once and *ex ante*. Both types of cost, however, accomplish the desirable goal of ruling out outlandish threats. The reputation model is worked out in Kreps and Wilson 1981.

shirk. Similarly, a strategy for agent 2 is the mapping  $(a, d, s) \times \{A, R\} \rightarrow \{A, R\}$  and  $(a, d, s) \times \{A, R\} \times \{A\} \rightarrow \{C, S\}$ . The solution concept is the subgame perfect Nash equilibrium.

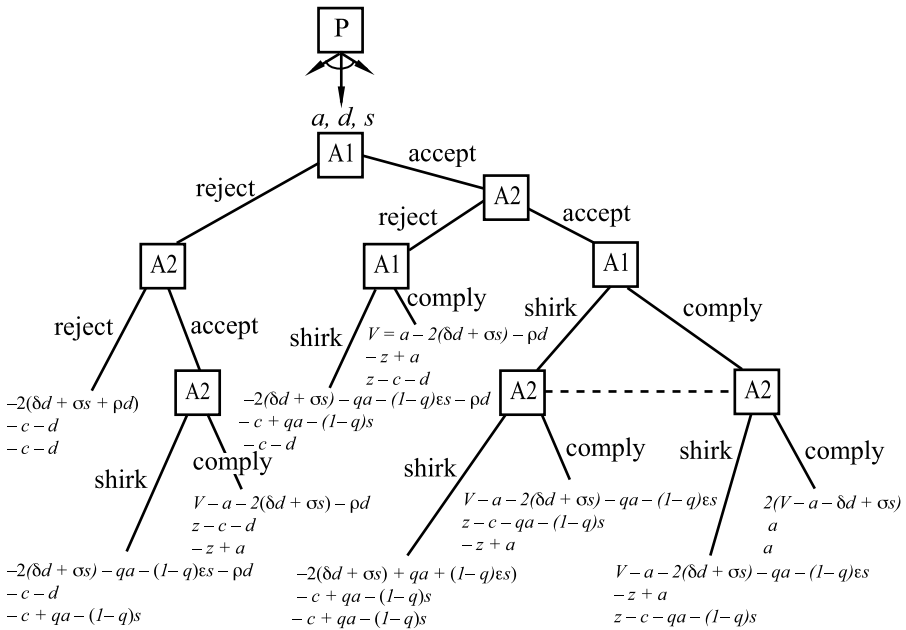


FIGURE 1. Game tree for the principal-and-one-dyad game

The principal has a clear goal—nonproliferation—which her control of the agenda enables her to achieve by following a simple strategy: offer a regime that leads the agents successively to accept the offer and comply with the obligation of disarming. To enable such result, the principal's choice of design must satisfy a certain number of constraints. A first set of constraints, called "participation constraints," insures that each agent prefers "accept" and "comply" to "reject." Depending on what the other agent is doing, there are two participation constraints for each agent. A first participation constraint insures that agent  $i$  prefers "accept" and "comply" to "reject" given that agent  $j$  rejects, with  $i, j = 1, 2$ . Formally,  $-z + a \geq -c - d$ . In other words, the agent's payoff for acceptance and compliance, which is equal to sucker payoff  $-z$  plus transfer  $a$ , must be greater than or equal to the agent's payoff for rejecting the contract, which is equal to the suboptimal payoff  $-c$  plus the harm caused by the sanction for not participating  $-d$ .

A second participation constraint insures the same result given that  $j$  accepts and complies. Formally,  $a \geq z - c - d$ , with, on the left side, the cooperation



payoff (normalized to zero) plus transfer  $a$ , and, on the right side, the temptation payoff  $z - c$  plus the sanction payoff for not participating  $-d$ . This second constraint is mathematically identical to the first constraint.

In addition to insuring that agents participate, the principal must also insure that they comply. Two constraints, called “incentive constraints,” must be satisfied to insure that result. A first incentive constraint requires that  $i$  prefers “accept” and “comply” to “accept” and “shirk” given that  $j$  rejects. Formally,  $-z + a \geq -c + qa - (1 - q)s$ . The left side is the same as above. The right side features the suboptimal payoff  $-c$ , the transfer  $a$ , which the cheating agent receives with  $q$  probability, and sanction payoff  $-s$ , which he receives  $1 - q$  of the time. A second incentive constraint insures the same result given that  $j$  accepts and complies. It is mathematically identical to the first incentive constraint and need not be repeated.

Three additional constraints are required to prevent carrots from changing into sticks and sticks into carrots:  $a, d$ , and  $s \geq 0$ .

Solving the game is tantamount to solving a constrained maximization problem—program  $P^1$ —in which the principal chooses the mix of incentives  $a, d$ , and  $s$  so as to maximize her utility for disarming the two agents subject to the five aforementioned constraints:

$$P^1 = \begin{cases} \max_{a,d,s} U_P = 2(V - a - \delta d - \sigma s) \\ \text{such that } a \geq z - c - d, \\ a \geq \frac{z - c}{1 - q} - s, \\ a \geq 0, \\ d \geq 0, \\ s \geq 0. \end{cases}$$

The principal’s payoff  $U_P$  is, for each agent, the gain for disarming  $V$  minus the paid transfer  $a$ , further diminished by the aggregate cost of threatening sanctions  $\delta d + \sigma s$ .

The solution to the problem yields a unique equilibrium for each configuration of parameters  $V, \delta, \sigma, z, c$ , and  $q$ . More precisely, depending on the relative marginal costs of the incentives that are featured on the equilibrium path, respectively unity for rewarding,  $\delta$  for threatening to sanction nonparticipation, and  $\sigma$  for threatening to sanction cheating, the solution features either a reward-only regime, a regime with sanctions only, a regime with only one sanction, or any mix thereof. Since the determination of the incentive structure of a particular regime—the particular mix of sticks and carrots—is not the primary goal of this article, I only solve the game for the incentive structure that most closely con-

forms to the reality of the nuclear proliferation regime: a mixed-incentive structure or  $a, d$ , and  $s > 0$ . The solution to the principal-and-one-dyad game is stated in the following proposition.

*Proposition (1): The principal-and-one-dyad game has a family of mixed incentive equilibria of the form  $a^* = z - c - \tilde{d}$ ,  $s^* = \frac{z-c}{1-q} q + \tilde{d}$ , for any fixed  $\tilde{d} \geq 0$  and for  $1 = \delta + \sigma$ . Both agents accept the contract and disarm.*

This result is proven in the Appendix. Note that the equilibrium is calculated for  $a$  and  $s$  as a function of  $d$ . This is because the mixed-incentive equilibrium is underspecified and can only be determined for two choice variables as a function of the third.<sup>20</sup> I chose to make  $d$  the determinant for reasons that will become apparent in the empirical section. Also note that both the reward and the two sanctions are present in the equilibrium because the reward has the same marginal cost as the two sanctions added together. This knife-edge condition does not necessarily mean that the mixed incentive equilibrium is a zero-probability event. More likely, it represents the unresolved nature of the debate on the relative efficiency of carrots and sticks, a debate that is fueled by the fact that different interest groups often have opposite stakes in this debate.

The result in proposition (1) exhibits both intuitive and counterintuitive causal connections. On the intuitive side, first, both transfer  $a$  and sanction against cheater  $s$  increase with the cost of compliance  $z$ , suggesting that countries that find it hard to part with nuclear weapons invite a higher price to join in and a steeper sanction to comply. Equally intuitive is that both incentives decrease in  $c$ , the opportunity cost of disarming. Also, a rise in nonparticipation sanction  $d$  logically commands a lower transfer  $a$  to participants; splurging on the sticks allows the principal to save on the carrot—carrot and sticks are interchangeable.

Less intuitive is that a rise in the sanction against nonparticipation,  $d$ , commands a higher sanction against cheating,  $s$ . This is because a rise in  $d$  brings into the regime members with higher compliance costs, who would not have joined otherwise, and therefore are harder to deter from cheating. Expanding membership to tough types without further repressing cheating is cause for cheating. This conclusion is in line with the link between “deep cooperation” and the need for enforcement underlined by Downs, Rocke, and Barsboom<sup>21</sup> as well as the simultaneous widening of European integration and deepening of its judicial system.<sup>22</sup>

20. The fact that  $d$  cannot be directly solved for reflects my attempt to reproduce the content of the NPT regime. One incentive is redundant; the member states could have obtained a similar result with only two incentives.

21. Downs, Rocke, and Barsboom, 1996.

22. Burley and Mattli 1993.

*Principal and "N" Dyads*

I now expand the model to a large number of dyads. I make the simplifying assumption that each country is part of only one dyad. The assumption captures the main idea that arms races are local, in the sense that Brazil worries about Argentina getting the bomb a lot more than North Korea or Iran. Note, though, that the assumption leads one to ignore arms races that are not strictly dyadic but involve three or more countries—Chile also worries about Brazil and Argentina getting the bomb.

Dyads differ with respect to the rivalry index,  $z$ . For the sake of simplicity, I assume that the dyads are distributed discretely on  $z$ ; they are numbered according to the intensity of rivalry  $z_k$  with  $k$  varying from 1 to  $N$ , such that  $z_1 = 1$ ,  $z_2 = 2, \dots, z_N = N$ . Hereon,  $z_k$  may be interpreted as an agent's marginal cost of compliance: it is equal to the marginal value that each agent attributes to a prize equally valued at 1.

The incentive structure of the regime again is of the mixed type.<sup>23</sup>

The principal's choice variables are twofold: whether to exclude or include, and, in the case of inclusion, whether to do so only by means of a multilateral treaty or also with bilateral instruments. The two types of instruments differ in two respects. First, a multilateral treaty provides all signatories with the same contract irrespective of their compliance costs. It is analogous to the law of one price in market economics. In contrast, a bilateral instrument allows the principal to offer each agent his own compliance cost. The bilateral approach is equivalent to market segmentation practiced by monopolists.

Second, the transaction costs for the multilateral approach are lower than for the bilateral approach. Transaction costs is a basket category, initially coined by economists to explain why market relations, understood as a contract between two independent individuals, might fail to maximize these individuals' joint utility.<sup>24</sup> Here, I adopt a restrictive definition of transaction cost as the contractual costs incurred in the process of making a deal. Aggregate transaction costs are lower for the multilateral approach to regime building than for the bilateral approach because there are scale economies in bargaining and dealmaking.<sup>25</sup> I model this feature by

23. A complete solution would require that all incentive structures be solved for simultaneously—a technically feasible, yet quite unnecessary task for the study of a single regime. Note, however, that a more comprehensive solution along these lines does not change the equilibrium results for the mixed incentive structure. What it does is to identify the conditions of application of the mixed incentive structure in relation to other structures. One of these conditions was calculated separately and is offered in the Appendix and proposition (1). Moreover, all empirical predictions are generated through the comparative statics method, which does not formally require the identification of boundary conditions.

24. Williamson 1975. T-costs are causes for market failures. Depending on the type of solution that is prescribed to the market failure, they fall into two families: opportunism, asset specificity, uncertainty, and the inability to make a credible commitment cause failures that are solved by delegation or hierarchy; bargaining costs, scale economies, policy externalities, coordination, and the holdout problem cause failures that find a solution by moving from bilateral to multilateral negotiations.

25. For an argument in this direction, see Keohane 1984, 90. See also conjecture C3 in Koremenos, Lipson, and Snidal 2001, 788.

assuming that any deal, bilateral or multilateral, costs constant  $T$  to process, with  $T$  positive.

From these two differences, it follows that the principal matches instrument with agent's compliance cost: (1) the multilateral treaty is calibrated to the low and intermediate marginal-cost agents, (2) bilateral instruments are directed to intermediate cost agents, and (3) exclusion targets high cost agents. To see this, first imagine a situation with no transaction cost. The principal only uses bilateral contracts. This is easily seen in Figure 2A, in which I map the cost to the principal on the vertical axis as a function of the agent's compliance cost  $z$  on the horizontal axis, with integer  $v > 1$  a point on the  $z$  axis. To keep the graph two-dimensional, it is assumed that the only incentive that is available to the principal is the transfer  $a(z)$ ;  $s(z) = d(z) = 0$ . The cost of wooing countries with compliance cost less or equal than  $v$  by means of a multilateral contract is equal to the  $va(v)$  rectangle, because the multilateral approach forces the principal to pay the transfer it pays to  $v$  to all agents left of  $v$ . In contrast, by means of bilateral contracts alone, this same cost would be equal to  $\sum_1^v a(z)$ , the triangle situated below the curve, which is only half the size of the rectangle. The triangle above the curve is a surplus that I refer to as the "agent surplus" in direct analogy with the "producer surplus" of market economics.<sup>26</sup>

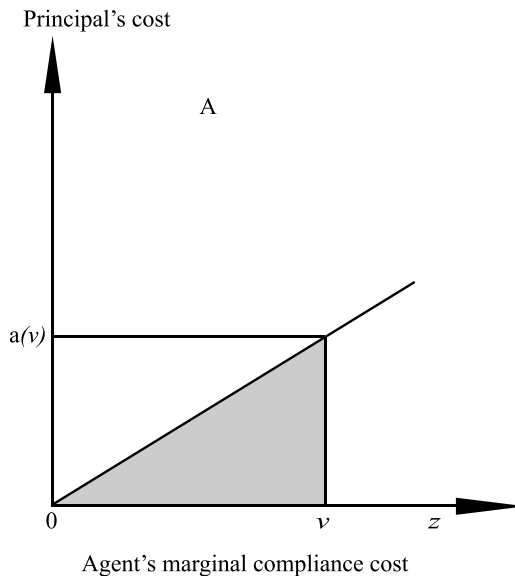


FIGURE 2A. *Impact of transactions costs*

26. The producer surplus is the amount that producers benefit by selling at a market price that is higher than they would be willing to sell for.

With a transaction cost  $T$  significantly different from 0, however, the  $v$  bilateral contracts could be costlier than a multilateral contract. This case is represented in Figure 2B, where the cost of the multilateral contract is the same as before plus the transaction cost incurred once ( $va(v) + T$ ), while the cost of the bilateral contracts is the former triangle augmented with the transaction cost incurred  $v$  times ( $\sum_1^v a(z) + vT$ ).

Finally, there is a higher value of  $v$  past which the bilateral approach becomes the less costly of the two again. Indeed, it is easy to verify that as  $v$  increases, so does the relative size of the  $c b a(v)$  triangle relative to the  $cde$  triangle. Therefore, from the principal's perspective, the presence of transaction costs makes the one-treaty approach more efficient than the bilateral approach for low-compliance-cost countries, but less efficient for high-compliance-cost countries.

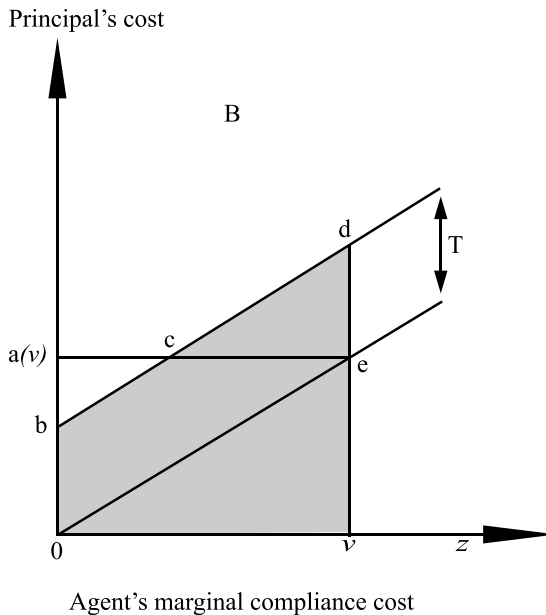


FIGURE 2B. *Impact of transaction costs*

Although I have assumed so far that every country is included, this need not be so. The principal would like to exclude countries with excessively high compliance costs. She gains nothing from an agent's nonparticipation, she may actually lose something if she sanctions them; and yet she is nevertheless better off excluding than including these countries in the regime.

Combining these different insights yields the solution template drawn in Figure 3 (also drawn under the assumption that  $s(z) = d(z) = 0$ ). On the left-hand

side of the graph, in the  $(1, x]$  segment, the agent's cost is sufficiently low that it makes sense to seek nonproliferation by offering a reward and by doing so through a single generic contract, minimizing transaction costs but providing a surplus to all the agents to the left of  $x$ . In the middle part of the graph, in the  $(x, y]$  interval, each agent's compliance cost is too high in relation to the fixed transaction cost for overlooking the surplus. Rather than offering a beefier multilateral contract, the principal tops off the existing multilateral contract with bilateral contracts customized to each member of the interval. Finally, there is a third group of agents on the right-hand side of the graph, the  $(y, N]$  segment, with costs of compliance that are so high that it is not in the interest of the principal to even try to elicit their compliance. To put it succinctly, the regime should obey the following generic condition:

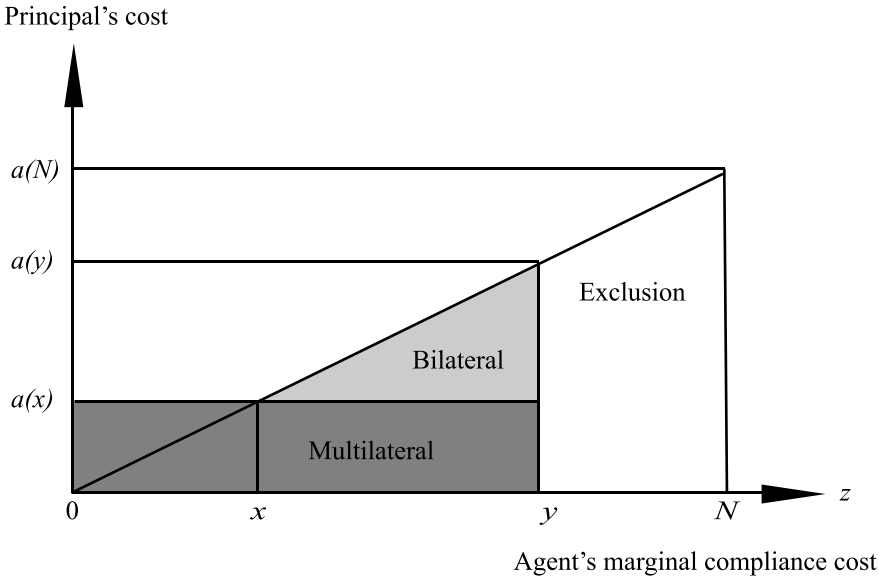


FIGURE 3. Solution template

*Lemma 1: If  $x = z_k$ , with  $k \in (1, N]$ , is defined as the level of dyadic rivalry that makes the principal indifferent between giving or not giving bilateral incentives in addition to multilateral incentives to the agents in dyad  $k$ , and if  $y = z_l$ , with  $l \in (1, N]$ , is defined as the level of dyadic rivalry that makes the principal indifferent between including and excluding the agents in dyad  $l$ , the regime must satisfy  $1 \leq x \leq y \leq N$ .*

The analogy with consumer economics assumes that the incentive is a reward. What if, instead, the principal merely threatens to sanction noncompliance? The analogy still works provided that one is ready to think of sanctions as negative prices. For instance, Santa Lucia, a country better known for its beaches than its desire to pursue nuclear weapons, would comply with the proliferation ban even if she were given negative price  $-s$  or, to put it more intuitively, even if she were asked to pay  $s$  in addition to complying. The fact that she is not asked anything besides compliance means that she saves  $s$ . Generally speaking, positive incentives set the agent's reservation value to 0, whereas negative incentives set the reservation value below 0.

Solving the  $N$ -dyad game allows for the principal to choose cutpoints  $x$  and  $y$  that maximize her aggregate utility while simultaneously satisfying the constraints necessary to have  $y$  dyads take the mixed-incentive contract and comply. Formally, it means for the principal to solve program

$$P^N = \begin{cases} \max_{x, y} U_P = 2xg(x) - T + 2 \sum_{x+1}^y (g(z) - T) - 2\rho\tilde{d}(N - y) \\ \text{with } g(z) = V - (z - c - \tilde{d}) - \sigma \left( \frac{z - c}{1 - q} q + \tilde{d} \right) - \delta\tilde{d} \text{ for any } z \in (1, N] \\ \text{and such that } 1 \leq x \leq y \leq N. \end{cases}$$

Function  $g(z)$  is the principal's payoff for extending a mixed contract to a country with war intensity  $z$ . It is constructed by substituting the equilibrium values of proposition (1) into the principal's utility function previously defined in  $P^1$ . The first two terms of the principal's utility function  $U_P$  are the principal's net gain for the multilateral treaty with the first  $x$  dyads: she gives to each of the  $x$  dyads of that group the same treatment as to the dyad with the highest marginal cost of compliance in that group,  $z = x$ , and pays only one transaction cost. The summation term is the principal's net gain for the  $y - x$  dyads with which she signs  $2(y - x)$  bilateral treaties with as many transaction costs.<sup>27</sup> The last term is the cost to the principal of excluding  $N - y$  dyads from the regime.

The problem affords different types of solutions depending on the values taken by  $x$  and  $y$ . I do not solve for all of them, but focus instead on the one that corresponds to the NPT regime—the mixed instrument regime (not to be confused with the mixed incentive structure also present in the NPT regime), featuring a multilateral treaty, numerous bilateral agreements, and the exclusion of a handful of countries. This is the case for which  $1 < x < y < N$ . That equilibrium is stated in proposition (2) and proven in the Appendix.

27. To build the summation term, I took advantage of the the mathematical identity between offering each of the the  $y - x$  dyads (1) a multilateral treaty and a supplementary bilateral treaty, or (2) a bilateral treaty with incentives that subsume those of the multilateral treaty.

*Proposition (2): There exists a subgame perfect Nash equilibrium in which the principal offers*

1. *transfer*  $a_M^* = x^* - c - \tilde{d}$  *and sanction threat*  $s_M^* = \frac{x^* - c}{1 - q} q + \tilde{d}$  *to agents with rivalry index*  $z \in [1, x^*]$ ; *agents accept and comply;*
2. *transfer*  $a^*(z) = z - c - \tilde{d}$  *and*  $s^*(z) = \frac{z - c}{1 - q} q + \tilde{d}$  *to agents with rivalry index*  $z \in [x^*, y^*]$ ; *agents accept and comply;*
3. *no transfer but inflicts sanction*  $\tilde{d}$  *on agents with rivalry index*  $z \in [y^*, N]$ ; *with*

$$\frac{1 - q}{1 - q + \sigma q} T \leq x^* \leq \frac{1 - q}{1 - q + \sigma q} T + 1 \text{ and } \frac{(V - T + (1 - \delta - \sigma + \rho)\tilde{d} + c(1 + \frac{\sigma q}{1 - q}))(1 - q)}{1 - q + \sigma q}$$

$$- 1 \leq y^* \leq \frac{(V - T + (1 - \delta - \sigma + \rho)\tilde{d} + c(1 + \frac{\sigma q}{1 - q}))(1 - q)}{1 - q + \sigma q}, \text{ for any given } \tilde{d}$$

$$\in (0, y^* - c), \text{ and for } \max \left\{ 0, V - (N - c) \left( \frac{1 - q(1 - \sigma)}{1 - q} \right) + \tilde{d}(1 - \sigma - \delta + \rho) \right\}$$

$$< T < \frac{1}{2} V + \frac{1}{2} c \left( \frac{1 - q(1 - \sigma)}{1 - q} \right) + \frac{1}{2} \tilde{d}(1 - \sigma - \delta + \rho).$$

Clause (1) summarizes the multilateral instrument, clause (2), the multilateral together with the bilateral instruments, and clause (3), the conditions of exclusion. Incentives are higher for agents who receive bilateral instruments. Note also that the equilibrium obtains for a specific family of parametric configurations, which one may characterize as intermediate values of transaction cost  $T$  expressed as a function of the other parameters. A  $T$  that is too small takes the wind out of the multilateral instrument, whereas a very large  $T$  prices out all types of instruments.

## Conjectures and Empirics

The central idea of this article is that the NPT regime has three components—multilateral, bilateral, and exclusion—and that the relative importance of each component is a function of both compliance and transaction costs. Although I showed in the first section that a cursory sorting of countries matched this three-tiered structure, the result could either be spurious or statistically insignificant. To raise confidence in the belief that the model captures the actual causal mechanisms behind this sorting, one needs to draw more predictions and test as many of those as possible against historical reality. Therefore, I use the model to draw a series of conjectures. The model's comparative statics—eight total—are calculated in the Appendix; four of them are tested below.

First, proposition (1) predicts that a higher rivalry index implies a higher compliance cost and thus a higher risk of exclusion. Although fairly intuitive, the prediction conflicts with Koremenos, Lipson, and Snidal's "M3 conjecture," according to which "inclusive membership increases with the severity of the distribution prob-



lem.”<sup>28</sup> The M3 conjecture draws on Snidal’s<sup>29</sup> work on relative and absolute gains; it rests on the assumption that states are concerned about power gaps that might develop between them as a result of a commonly but asymmetrically beneficial endeavor, like trade for instance. Trading with other states, Snidal argues, may mitigate this concern—hence the drive for inclusion. Although the present model clearly deals with a distributive issue, the relative gains concern is absent from the principal-agent relation, yielding an opposite prediction as to the impact of distribution on inclusion. Hence, the first conjecture:

*Conjecture 1. Rivalry: A higher rivalry between members of a dyad and thus a higher marginal cost of compliance implies a greater risk of exclusion.*

Second, the costlier the development of nuclear arms is to the agents (the higher  $c$ ), the more inclusive the regime is. This stands to reason, the same way as it stands to reason that a buyer and a seller of a good are more likely to reach an agreement if the seller’s value for the good is low. The monetary cost of pursuing a weapons program cannot be directly observed, but since it matters only in relation to the wealth of the potential agent, it is easily proxied by the latter. Hence the second conjecture:

*Conjecture 2. Affordability: The greater affordability of nuclear weapons raises the risk of exclusion.*

Third, the regime is more inclusive if it is costlier to the principal to sanction nonsignatories. In such a case, the sanctioner tries to reduce the pool of nonparticipants who have to be sanctioned by signing in more countries. The hypothesis applies to the marginal cost  $\rho$ , and in some cases to the unitary cost  $\tilde{d}$  (see Appendix for details). If the commitment to sanctioning nonparticipation frays, then initially-agreed-upon transfers will no longer suffice to prevent more agents from opting for nonparticipation. One should observe the following conjecture:

*Conjecture 3. Sanctioning Capacity: A rise in the sanctioner’s sanctioning cost of nonparticipation increases the risk of exclusion.*

Fourth, a rise in monitoring inefficiency  $q$  also leads to a rise in the fraction of excluded countries; this is because deterring cheating calls for a higher transfer, making inclusion costlier and thus exclusion cheaper. This is a standard result, though it is usually phrased differently: the more severe the moral hazard problem is, the more centralized the monitoring device has to be to hold membership constant.<sup>30</sup> Hence, the last conjecture:

28. Koremenos, Lipson, and Snidal 2001, 784.

29. Snidal 1991.

30. This is conjecture C1 in Koremenos, Lipson, and Snidal 2001, 787.

*Conjecture 4. Monitoring Capacity: A rise in monitoring efficiency decreases the risk of exclusion.*

Four more comparative statics should be mentioned despite the fact that they cannot be tested on the NPT case. First, a principal with a higher value for nonproliferation excludes fewer countries (a rise in  $V$  yields a rise in  $y$ ). Since I see no reason why the superpowers' valuation for nonproliferation would have changed over time, the conjecture is not testable on the NPT case. Second, a rise in the cost of threatening sanctions makes inclusion costlier to the principal, thereby increasing the number of excluded states (a rise in  $\sigma$  or  $\delta$  yields a drop in  $y$ ). It is difficult, however, to entertain the idea that reputation costs change over time. Last, a rise in transaction costs yields a drop in the use of bilateral instruments (a rise in  $T$  yields a rise in  $x$  but drop in  $y$ ). Although central to the model, this prediction can only be tested on a sample of regimes. I do not do this systematically but offer illustrations in the penultimate section.

I now look at the evidence for the four highlighted conjectures.

### *Rivalry and Affordability*

I test the rivalry and affordability conjectures by correlating the principal's decision whether or not to include a country in the regime with that country's security characteristics and wealth. A nonparticipant being a nonsignatory, the dependent variable is NPT RATIFICATION, coded 1 for ratification and 0 for nonratification.

Ratification is a unidirectional decision: once made, it is typically not, with one exception, taken back. Moreover, countries that ratified did it over a period of thirty-plus years (see Figure 4), during which many countries' security position changed. I use the technique of event history modeling that allows one to focus on the spell of time before ratification occurs and study the respective impact of cross-national and time-varying covariates.<sup>31</sup>

A first set of independent variables seek to capture the rivalry index ( $z$ ). A key implication of the local arms race is that countries that feel threatened by their neighbors should be less eager to give up their right to acquire nuclear weapons. I assume that the sense of insecurity reflects the history of conflict. I alternatively use two measures of a country's history of conflict. A first measure HOSTILITY is the sum of the Correlates of War (COW) indices of hostility levels with contiguous states averaged over the prior twenty years.<sup>32</sup> A second measure, ENDURING

31. Box-Steffensmeier and Jones 2004.

32. The variable is CWHOST1; it is described as "Relevant hostility level reached by CCode1 in a MID vs. CCode2 in this year (0 = No hostility [no MID], 1 = No militarized action, 2 = Threat to use force, 3 = Display of force, 4 = Use of force, 5 = War)." Neighbors are contiguous on land or within 400 miles over water. Neighbors are unweighted. Weighting each neighbor with CAP\_2, COW national capabilities index for CCode2, made no difference. Reference is to Bennett and Stam 2000.

RIVALRY, is participation in an enduring rivalry because most militarized disputes occur in the context of enduring rivalries.<sup>33</sup>

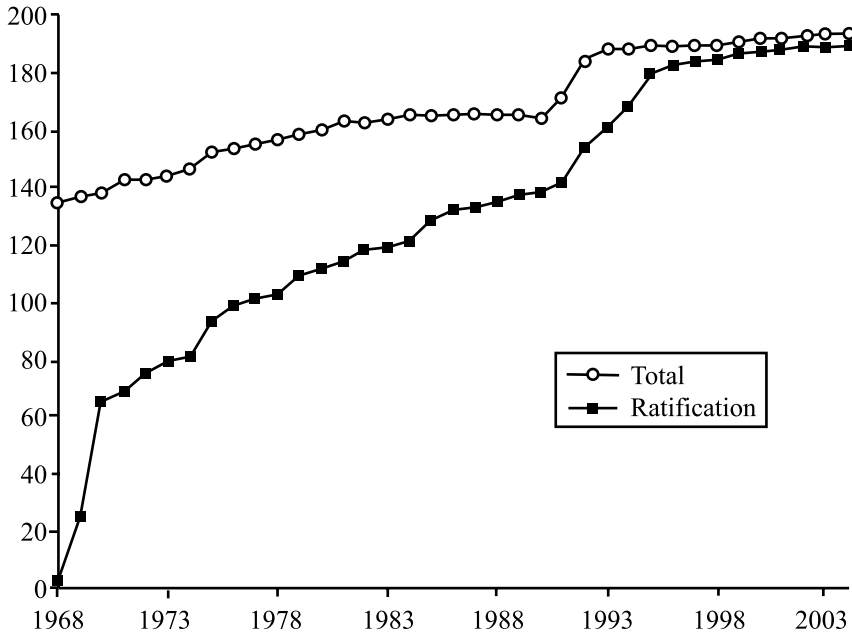


FIGURE 4. Number of ratifications and total number of countries

A country's sense of insecurity may also reflect the regime type of its neighbors. Building on the democracy and war literature, I venture that being in the neighborhood of democracies provides a country with a greater sense of security than if it were surrounded with autocracies. The variable *DEMOCRATIC NEIGHBORS* is a weighted average of the democracy score of a country's neighbors. I used national capabilities as weight to select out information on a country's neighbors that are smaller than the country and unlikely to represent a military threat.<sup>34</sup>

33. I am using Singh and Way's dichotomous recoding of Bennett's coding of enduring rivalries. Enduring rivalry is "any dyad in which six MID's occurred over a period of twenty years with a maximum of fifteen-year gap between any two disputes" (Bennett 1998, 1214); see Singh and Way 2004, 869.

34. The variable is built on Polity IV's *POLITY2* variable described as "Democracy Score for CCode1 (DEMOC - AUTOC)." Reference is Jagers and Gurr 1995. Neighbors are land contiguous or within 400 miles over water. Democracy scores are weighted with the COW variable *CAP\_2*, described as "National Capabilities Index for CCode2." Both variables were obtained through EUGene (Bennett and Stam 2000).

The modeling of the local arms race as a prisoner's dilemma suggests that the worst outcome is for one side to give up nuclear weapons unilaterally, for then the other side has an advantage in not doing so. This suggests that a country would not join the NPT regime unilaterally, but in coordination with its neighbors. One observes instances of this logic in Western Europe and in the Southern tip of Africa. I created the variable *NEIGHBOR'S RATIFICATION*, which for a given country, identifies all its contiguous countries, determines whether or not they have ratified the NPT, and weighs that score with their relative power.<sup>35</sup>

The presence of a security threat may not lead a country to acquire nuclear weapons if the threat is offset by a security guarantee from a nuclear power (variable *a*). Security guarantees figure as one of the most effective bargaining tools that the United States and the Soviet Union used in their respective bilateral efforts to stop proliferation. I build the variable *SUPERPOWER'S DEFENSE GUARANTEE*, which identifies all the countries toward which the United States or the Soviet Union/Russia had a formal defensive obligation during the year of observation.<sup>36</sup> I am aware that many of the guarantees that are included in the dataset were not granted bilaterally but through a collective security treaty—NATO for example. I treat them as bilateral instruments nevertheless because the model does not recognize a regional level different from the bilateral and multilateral levels. There is also the fact that NATO—the most “collectivist” of the regional arrangements—is not a symmetric alliance: with the notable exception of the American superpower, no single member country is indispensable to the survival of NATO, not even Britain or France.

Not quite a security guarantee, but a source of U.S. leverage nevertheless, is *U.S. AID PER CAPITA*, the total amount of economic and military aid provided by this country and calculated as a proportion of the population of the receiving country.<sup>37</sup>

The second independent variable is the affordability of a nuclear weapon program (*-c*), proxied by the wealth and size of the country: The larger and wealthier a country, the easier and more desirable the acquisition of nuclear weapons. I use total *POPULATION* to proxy for size and *GDP PER CAPITA* to proxy for level of development.<sup>38</sup>

I round off the test with several control variables representing possible alternative accounts of a government's decision to ratify the NPT. A first control variable is peer pressure. Constructivists argue that, as more countries uphold the regime, the regime gains in authority and nonparticipants are hard-pressed to join. An easy measure of peer pressure is the cumulative number of members. I call this variable *TOTAL RATIFICATION*.

35. I use as weight the COW variable *CAP\_2*, described as “National Capabilities Index for CCode2.”

36. I built this variable using the *DEFENSE* variable of the Directed Dyad-Year ATOP dataset (Leeds 2005), featuring the United States and the Soviet Union/Russia as “State A.”

37. See the Web site of the United States Agency for International Development at <http://www.usaid.gov>.

38. Both variables are from Gleditsch 2002. Version 4.1 of the data is available at <http://privatewww.essex.ac.uk/~ksg/exptradegdp.html>. Accessed 14 October 2007.

A second control variable is the end of the Cold War, which, according to realists, took the wind out of global regimes. I create the dichotomous variable 1991–2002, coded “1” during the corresponding years and “0” otherwise.

Moreover, Singh and Way<sup>39</sup> found an occasional role for regime type, the idea that a DEMOCRACY is less likely to pursue nuclear arms and thus more likely to join the nonproliferation regime.<sup>40</sup>

I estimate four Cox proportional hazards models on various combinations of independent variables. The proportional model is preferable to a parametric model in the absence of a particular theory about the way time affects the causal impact.<sup>41</sup> The universe is all sovereign countries in a given year, with the exclusion of the five nuclear powers recognized in the NPT. The results are presented in Table 1.

TABLE 1. *Cox models of NPT ratification*

<i>Variables</i>	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
<i>Security risk</i>				
HOSTILITY	−0.065	−0.082*		
RIVALRY			−0.45**	−0.53**
DEMOCRATIC NEIGHBORS	0.030	0.043**	0.045**	0.04**
NEIGHBORS' RATIFICATION	0.43	0.47*	0.51*	0.42*
<i>Bilateral guarantees</i>				
SUPERPOWER'S DEFENSE GUARANTEE	0.31	0.69***	0.69***	0.53***
RIO		−0.958***	−0.980***	−0.837**
U.S. AID PER CAPITA	−6.68	−7.38	−8.55*	0.16***
<i>Affordability</i>				
GDP PER CAPITA	$−4 \times 10^{-5}***$	$−5 \times 10^{-5}***$	$−4 \times 10^{-5}***$	$−4 \times 10^{-5}***$
POPULATION	$−6 \times 10^{-5}**$	$−9 \times 10^{-6}**$	$−7 \times 10^{-6}**$	$−6 \times 10^{-6}**$
<i>Controls</i>				
TOTAL RATIFICATION 1991–2002	−0.007*	−0.008*	−0.010**	−0.007
DEMOCRACY	1.22***	1.01***	1.00***	0.84**
	0.006	0.003	0.003	—
<i>Log likelihood</i>	−537.9	−532.5	−531.9	−607.5
<i>Chi-squared</i>	82.99***	91.49***	87.07***	216.45***
<i>Number of countries</i>	142	142	142	157
<i>Number of ratifications</i>	136	136	136	151
<i>Time at risk</i>	1173	1173	1173	1246

Notes: Coefficients are estimates for Cox proportional hazards model; robust standard errors, adjusted for clustering by country, are used to calculate the *p* values for two-sided tests. GDP = gross domestic product. \*\*\* *p* < .01; \*\* *p* < .05; \* *p* < .10.

39. Singh and Way 2004.

40. See footnote 34 for definition and origins of the democracy variable.

41. See Box-Steffensmeier and Jones 2004, 88. Although three variables—RIVALRY, RIO, and POPULATION—fail the Grambsch and Therneau (1994) test of the proportional hazard assumption, the model as a whole passes it.

The results for the baseline model, listed in column (1), seem weak at first sight—none of the security risk and bilateral guarantee variables are statistically significant. A closer examination of the data identifies as the cause for the unexpected result the Rio Pact countries: being a member of the Rio Pact makes a country less likely to sign and ratify the NPT than not being a member. Controlling for the Rio Pact in the other columns brings the security and guarantee variables in line with the present argument. There is one apparent exception to this generalization, the variable *U.S. AID PER CAPITA*, which in models 1, 2, and 3 is wrongly signed. Once again, a closer look at the data reveals that this result is not robust, but instead the artifact of collinearity with the *DEMOCRACY* variable. Dropping the latter leads the variable *U.S. AID PER CAPITA* to be positively and significantly correlated with the dependent variable (model 4). A possible reason for this correlation is the greater likelihood for U.S. aid to be distributed to democratic than authoritarian regimes.<sup>42</sup>

Models 2 to 4 underscore the importance of two sets of variables: the raw security risk, whether measured by hostility or enduring rivalry, and the security guarantees with the exception of the Rio Pact. I have no ready-made explanation for the Latin American exception, other than referring to the love-hate relation that exists between the United States and its traditional zone of influence. The coefficients on the security and guarantee variables are significantly different from zero. Their estimated impact on the dependent variable is also substantial. The entries of Table 2 represent the percentage change in the baseline hazard rate for a given change in each explanatory variable of model 4. For example, a country involved in an enduring rivalry with a neighbor has a hazard rate for signing the NPT that is 41 percent lower than a similar country involved in no enduring rivalry. In contrast, a formal security guarantee from a superpower other than through the Rio Pact increases the hazard rate by 71 percent. U.S. financial aid also plays a role, though a more limited one: a one-standard deviation increase from the mean yields a puny 4 percent increase in the chance of ratifying the treaty. These results make it clear that exclusion is a function of rivalry between neighboring states and thus of compliance costs with the principal's goal of nonproliferation. Koremenos, Lipson, and Snidal's<sup>43</sup> idea that distributive problems at large can be solved by increasing membership is not observable in the nuclear proliferation regime.

In accordance with predictions, the economic variables have a meaningful negative impact on the odds of ratifying the NPT, confirming the affordability hypothesis.

Among the control variables, the cumulative ratification index has no noticeable effect on the rate of ratification in model 4 and is wrongly signed in models 1 to 3, and thus does not support the view that states wish to conform. Only the end of the Cold War is significantly and substantively correlated with ratification, though

42. Dropping the *U.S. AID PER CAPITA* (or any other) variable does not make *DEMOCRACY* significant.

43. Koremenos, Lipson, and Snidal 2001.

**TABLE 2.** *Percentage change in hazards for statistically significant variables (Model 4)*

<i>Variables</i>	<i>Percentage change</i>
RIVALRY ( <i>dummy</i> )	−41
DEMOCRATIC NEIGHBORS ( <i>from none to all</i> )	+4
SUPERPOWER' DEFENSE GUARANTEE ( <i>dummy</i> )	+71
RIO ( <i>dummy</i> )	−57
U.S. AID PER CAPITA ( <i>one SD increase from mean</i> )	+4
1991–2002 ( <i>dummy</i> )	+131
GDP PER CAPITA ( <i>one SD increase from mean</i> )	−35
POPULATION ( <i>one SD increase from mean</i> )	−23

*Notes:* The change is calculated according to the formula:  $[(\exp(\beta \cdot X_2) - \exp(\beta \cdot X_1)) / \exp(\beta \cdot X_1)] \cdot 100$ , with  $\beta$  the estimated coefficient,  $X_1$  and  $X_2$  respectively the mean and the mean augmented by a standard deviation for continuous variables, and 0 and 1 for dummy variables. SD = standard deviation.

the sign belies gloomy neorealist predictions—ratification received a boost after 1990. Rather than sapping the basis for the nonproliferation regime, the end of the Cold War led to a revival of the NPT. I further investigate the reasons for this revival in the next section.

### *Sanctioning Nonparticipants and Monitoring Participants*

I now present anecdotal and systematic evidence for conjectures 3 and 4 related to the principal's ability to sanction ( $\rho$ ) and monitor ( $1 - q$ ). During the period 1977–91, the founders lost their capacity to sanction nonparticipants and monitor cheating. They did not offset the deficiency in this incentive by making the other incentives more powerful. As a result, the nonproliferation regime exhibited a loss in efficiency; it stopped attracting participants. In 1991, in contrast, the founders rebuilt this capacity; they also strengthened the IAEA safeguards and verification regime, resulting in greater participation.

The surprise explosion of an Indian device in 1974 led the founders to open their initial export cartel to other suppliers of nuclear fuel and technology. They founded the Nuclear Suppliers Group (NSG) and established the Trigger List, a list of material and equipment, the export of which they would regulate. From then on, however, the NSG fell into a state of, in one of his past chairman's words, "relative inactivity."<sup>44</sup> The supplier cartel broke up, divided between the three

44. Strulak 1993, 3.

Anglo-Saxon countries (United States, Canada, and Australia), who held near-monopolist control on the supply of uranium, and France, Germany, Switzerland, and Belgium, who were willing to fill the demand for the construction of fuel facilities—uranium enrichment and plutonium reprocessing plants.<sup>45</sup> Cooperation within the cartel came to an abrupt halt. The group did not meet for thirteen years. During that period, no change was made to the Trigger List, despite a pressing need for regular updating and extension to keep up with new technologies.

Compounding the erosion of the export controls regime was the ongoing weakness of the IAEA verification regime. The regime was admittedly weak from the start: time, frequency, and scope of inspections had to be prearranged with each government. Still, it was expected that the discovery of cases of noncompliance would call for and result in step-by-step strengthening. The thirteen-year lack of coordination among supplier states prevented this from happening, giving a head-start to would-be rogues. Iraq's first-attempt at getting the bomb—centered around the Osirak reactor and an Italian hot cell to separate plutonium—was to have been an NPT-safeguarded and regularly IAEA-inspected facility. After Israel bombed Iraq's reactor in 1981, Iraqi leader Saddam Hussein created a clandestine weapons program, which, despite employing seven thousand people, remained hidden from Western inspectors for nearly a decade.<sup>46</sup>

From a contractual perspective, the breakup of the cartel of suppliers modified the terms of the generic contract in favor of non-nuclear weapon states. Furthermore, the superpowers were not willing to make up for the weakening of the multilateral instrument by means of an active bilateral approach. The fall of the pro-American Shah in Iran and the Soviet invasion of Afghanistan in 1979 led the United States to take a more favorable attitude toward exports of fuel to India and Pakistan.<sup>47</sup> The administration of President Jimmy Carter exempted India from sanctions, while the Reagan administration maintained full military aid to Pakistan until 1990.<sup>48</sup>

The 1977–91 years were bleak for nuclear nonproliferation. On average, fewer countries adhered to the treaty in that period. During this time, Brazil and Argentina, two nonsignatories, began their nuclear weapon programs, and South Africa, a nonsignatory, succeeded in building a nuclear device, while Libya, Iran, Iraq, and North Korea, four signatories, began cheating with their obligations under the treaty.

This dismal period came to an end in 1992, when the conjunction of long-trend effects (abundance of uranium, drying out of the market for nuclear power) and short-term emergencies (the Persian Gulf War and the subsequent discovery of unsuspected amounts of sensitive materials, equipment, and technology in Iraq)

---

45. On the nuclear fuel shortage, see Ribicoff 1976. The breakdown in international consensus is well-captured in Kaiser 1980, 2.

46. Judith Miller and James Risen. "An Iraqi Defector Warns of Iraq's Nuclear Weapons Research." *New York Times*, 15 August 1998.

47. Clausen 1993, 151, 163, 170.

48. *Ibid.*, 170.



led France and Germany to rejoin the cartel and side with the U.S. desire to tighten up the regime. The list of sensitive items was expanded and export restrictions agreed upon thousands of items. The safeguards were tightened and the verification regime was strengthened through the 1997 Model Additional Protocols.<sup>49</sup> The 1990s saw a readjustment of the terms of the regime in favor of the nuclear cartel.

As a result, more states either signed and ratified or complied. Indeed, the list of states that acceded as non-nuclear-weapon states included Argentina, Chile, and Brazil in Latin America; South Africa, Angola, Namibia, Zambia, and Zimbabwe in the southern tip of Africa; Tanzania and Algeria in the rest of Africa; and all the ex-Soviet Republics in Asia. With the possible exception of some ex-Soviet Republics, none of these countries received positive or negative incentives other than those included in, or implied by, the NPT.<sup>50</sup>

The hypothesis that the breakdown of the cartel led to a slack in joining rates is easily tested on the ratification dataset. To this end, I create the dichotomous variable 1977–91, coded “1” during the corresponding years and “0” otherwise. I merely introduce this variable in the prior data analysis, from which I remove the end of the Cold War variable 1991–2002 on the grounds that it corresponds, both statistically and substantively, with the rebirth of the cartel. Percentage changes in hazard rates for a given change in each statistically significant explanatory variable are reported in Table 3. The 1977–91 period exhibits a strong negative impact on ratification, with a level of ratification half of what it was in the prior and posterior periods. All prior results hold.

**TABLE 3.** *Percentage change in hazards for statistically significant variables (Model 4 modified)*

<i>Variables</i>	<i>Percentage change</i>
RIVALRY ( <i>dummy</i> )	–42
DEMOCRATIC NEIGHBORS ( <i>from none to all</i> )	+4
SUPERPOWER’ DEFENSE GUARANTEE ( <i>dummy</i> )	+73
RIO	–57
U.S. AID PER CAPITA ( <i>one SD increase from mean</i> )	+3
1977–1991 ( <i>dummy</i> )	–50
TOTAL RATIFICATIONS ( <i>one SD increase from mean</i> )	+5
GDP PER CAPITA ( <i>one SD increase from mean</i> )	–35
POPULATION ( <i>one SD increase from mean</i> )	–22

*Notes:* The change is calculated according to the formula:  $[(\exp(\beta \cdot X_2) - \exp(\beta \cdot X_1)) / \exp(\beta \cdot X_1)] \cdot 100$ , with  $\beta$  the estimated coefficient,  $X_1$  and  $X_2$  respectively the mean and the mean augmented by a standard deviation for continuous variables, while 0 and 1 for dummy variables. SD = standard deviation.

49. Strulak 1993.

50. On the ex-Soviet Republics, see Drezner 1999.

In sum, four of the comparative statics that were drawn from the model find some form of historical validation: the risk of exclusion increases with (1) high compliance costs, (2) high opportunity costs, (3) poor sanctioning of nonparticipation, and (4) poor monitoring of compliance. These predictions are fairly intuitive, but given that they were all drawn from the theory, their empirical validation strengthen our confidence in the ability of the theory to explain the three-tiered structure of the nuclear proliferation regime.

## Extensions

It remains to be seen what applicability the present model commands in cases other than nuclear proliferation. I try the model on the currency, trade, and aid regimes. Although the currency regime seems a straightforward application of the present model, the trade and aid regimes call for special attention because they differ from the proliferation and currency regimes in important ways: transaction costs take a different meaning in trade, whereas multilateralism can support market segmentation in aid.

### *Currency*

The role of transaction costs in regulating the relative importance of multilateral and bilateral instruments is visible in currency regimes. Bretton Woods consisted of a multilateral agreement signed in 1944, featuring a common positive incentive—the possibility to draw resources from the International Monetary Fund (IMF)—and various bilateral arrangements between national treasuries.<sup>51</sup> In contrast, the gold standard in its heyday featured no multilateral instrument proper but consisted mainly of various informal arrangements between central banks. It is only after World War I that the regime became a matter for negotiations at international conferences.<sup>52</sup>

The ratio between multilateral and bilateral components thus differed markedly between the two regimes: whereas Bretton Woods was run through the IMF according to rules consigned in the agreement, the gold standard essentially was run by central bank governors meeting one another on an ad hoc basis whenever they felt the need to. The principle of gold convertibility was commonly shared without having ever been sanctioned by an international conference because relations between governments were delegated to and mediated by financial markets, a mechanism apt to minimize transaction costs.

Today's international currency regime resembles the gold standard. There is no institutionalized multilateral regime beyond the occasional Group of 8 (G8) meet-

51. See Dominguez 1993.

52. On the origins of the gold standard, see Gallarotti 1994.

ings among the major currency countries, while the IMF's role is limited to helping peripheral countries. Most of the bidding is made in global capital markets.

An effect of market mediation was, and is, to minimize the agent surplus that characterized Bretton Woods. Countries that were capable of monetary discipline still preferred to pursue liberal growth policies on the grounds that in case of emergency they could always avail themselves of the resources that were initially put in place for countries with less flexible domestic constituencies—a choice that eventually fueled worldwide inflation. Under market mediation, national treasuries pay interest rate premiums that closely reflect their monetary policies. Although more efficient, market mediation was not available in the postwar era because banks and securities markets were strictly regulated.<sup>53</sup> The multilateral agreement helped overcome problems of coordination among treasuries, which were held responsible for the collapse of the gold standard in the interwar period.

### *Trade*

The notion of transaction costs that I have used so far is restricted to the contractual costs of bargaining usually expended in time, transport, hotel, and legal assistance, which can be substantial given the multiple rounds it often takes for each side to figure out their respective reservation value. Although present in trade negotiations too, these costs take second place to another kind of transaction costs known as externalities. Under the modern trade regime, any bilaterally-negotiated tariff reduction is extended to all other parties on the basis of most-favored-nation (MFN) treatment, potentially lessening the benefit that any country may expect from entering into a bilateral negotiation in the first place. As a result, negotiators have an interest in free riding on one another and few deals get made. The solution to externalities of such kind is a multilateral negotiation in which trade partners follow an inflexible bargaining rule such as a 10 percent cut across the board, cuts proportional to current level of protection, or what not.<sup>54</sup>

While the inflexible rule approach addresses the externality problem (in a way that parallels how multilateral bargaining addresses the transaction cost problem), at the same time it presents the drawback of forcing the negotiations into a sub-optimal outcome, since cuts no longer reflect each economy's relative comparative advantages and marginal costs, but are homogeneous across trade partners, causing an agent surplus cashed in by those who would have made the same cuts for less in return. This is the reason why, whenever externalities were tolerable, negotiations within the General Agreement on Tariffs and Trade (GATT)—and a fortiori outside the GATT in the form of voluntary export restraints (VERs)—were conducted on a “principal supplier” basis, whereby “a country could be

53. See Verdier 2002.

54. See Pahre 2001 who uses the notion of “clustering” to describe the equivalent of multilateral bargains in the bilateral world that preceded the GATT.

requested to make a concession on a product only by the partner which provided the largest share of its imports.”<sup>55</sup>

The trade-off between externality and agent surplus gives one a sense of the relative role of multilateral and bilateral negotiations within the GATT. Consider a country like the United States, a group of trade partners, and a given basket of products. The United States would want to engage in an inflexible multilateral round of negotiations with suppliers of products for which it does not have a principal supplier, thereby limiting externalities, while it would want to have bilateral talks with suppliers of products for which it does, thereby limiting the agent surplus. Of course, the fact that the United States is also a supplier and that all countries are in this dual position may complicate the calculations, but it does not change the comparative statics of the institutional equilibrium, making room for bilateral negotiations when externalities, that is, transaction costs, are low in order to minimize the agent surplus.

It is interesting to note that within the GATT/WTO, all enforcement, whether it is mandated by panel rulings or pursued outside of the legal framework (for instance, through “Section 301”), is done bilaterally. There is no multilateral sanctioning mechanism within the trading regime.<sup>56</sup> While some have been eager to see something in enforcement that would make it bilateral rather than multilateral, note that one-on-one retaliation does not work in matters of proliferation, which calls instead for export control groups—the NSG in the nuclear regime, the Australia Group in the chemical and biological regimes, the equivalent body in the Missile Technology Control Regime, and the granddaddy of all, the Coordinating Committee on Multilateral Export Controls (COCOM).<sup>57</sup>

### *Aid*

A key assumption of the model is the market analogy: the multilateral instrument treats all members in the same way, while the bilateral instruments, in contrast, allow for market fragmentation. This assumption is useful in regimes such as nuclear proliferation, currency, trade with MFN treatment, and probably many other regimes, but not in the aid regime. Aid, whether multilateral or bilateral, can be made to reflect the recipient’s needs. As a result, the present model, applied as is, would predict that all aid should be multilateral because it has the advantage of minimizing transaction costs without the drawback of generating an agent surplus, since each recipient is (or could be) given to according to needs. Of course, this is not the case: no more than a third of all aid is estimated to be multilateral.<sup>58</sup> The

55. Winters 1990, 1290.

56. Unless, of course, a country violates its obligations and causes harm to all other countries—a very special case in which multilateralism results from the aggregation of bilateral behaviors.

57. Arguing in favor of the affinity between enforcement and bilateralism are Oye 1986 and conjecture M1 in Koremenos, Lipson, and Snidal 2001.

58. See Rowe 1978; Rodrick 1996; and Milner 2006.

mechanism responsible for the multilateral-bilateral structure in the aid regime calls for further analysis.

Recipient's need is the wrong dimension on which to classify countries. Recent studies<sup>59</sup> have made patent what one always suspected: the need for aid plays second fiddle to the recipient's strategic, political, or trading usefulness to the donor, with the result that bilateral aid, unlike multilateral aid, is almost always politicized. Moreover, the two dimensions of need and usefulness are negatively correlated, and this for two reasons: first, the poorest countries are also those that are the least likely to be useful to rich donors; second, the poorest countries offer a greater developmental bang for the buck—morally, if not economically—than wealthier ones. As a result, Milner notes, “multilateral aid tends to be given to poorer countries on average than does bilateral aid.”<sup>60</sup>

Analytically, the marginal gain of aid to the donor is one of two marginals. It is either the marginal developmental gain, which is obtained through multilateral aid and decreases with the recipient's wealth, or the marginal political gain, which is obtained through bilateral aid and increases with the recipient's wealth. The level of recipient's wealth at which these two marginals are equal defines a cutpoint below which the donor prefers multilateral aid to bilateral aid and beyond which the donor's preference switches to bilateral aid.

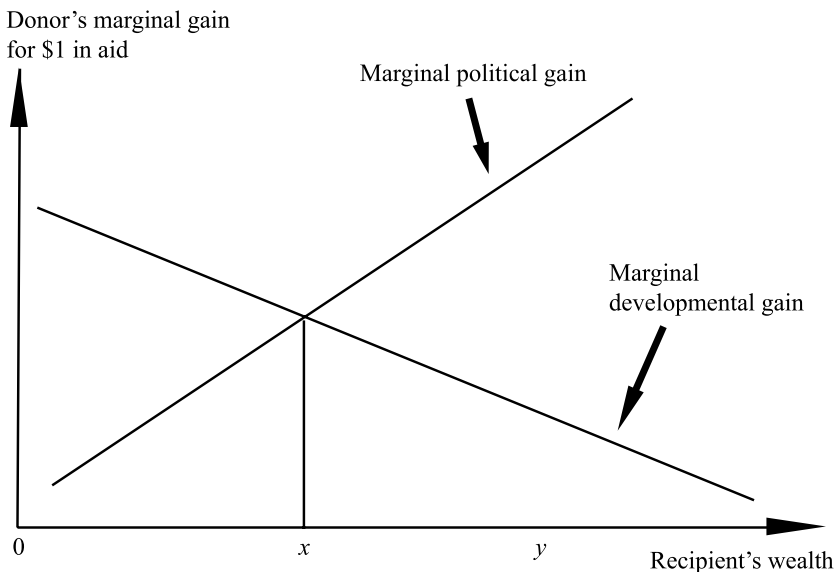


FIGURE 5. *Structure of the aid regime*

59. See Rodrick 1996; and Milner 2006.

60. Milner 2006, 114.

For a visual representation of this cutpoint, consider Figure 5, mapping the recipient's wealth on the horizontal axis and the donor's various marginal gains on the vertical axis. The descending curve maps the donor's marginal developmental gain of one dollar of aid as a function of the recipient's wealth, while the ascending curve maps the marginal political gain. The donor's marginal gain is maximized with multilateral aid up to the intersection point and with the addition of bilateral aid thereon. This intersection point corresponds to the  $x$  of the proliferation regime model, while the  $y$  falls to the right of  $x$  at a level of wealth past which it no longer is considered reasonable to ask for, or hand out, aid.

Therefore, only bilateral aid can be customized to the donor's political goals. The attribution of multilateral aid, in contrast, obeys fixed, impersonal criteria, over which individual governments have limited or no control. It is this second-order rigidity—the fact that all recipients have to be treated according to the same rules—that makes the member-surplus problem applicable to multilateral aid. Efficient for the most underdeveloped recipients, from which hardly any favor can be had, the multilateral mode of allocation becomes inefficient with countries that present donors with greater strategic or diplomatic opportunities. Rather than raising incentives for all recipients, donors prefer to customize incentives to targets of special interest to them. For instance, although both the United States and Europe value the protection of human rights and use aid to promote it, they have different geographic priorities, reflecting different colonial histories, inherited contacts, physical proximity, unresolved differences, and bargaining leverage. They use bilateral aid to foster human rights in their respective areas of influence.

In sum, the aid regime presents the same three-tiered structure as the nuclear proliferation regime. It can be modeled as a principal-agent contract and the logic for the cutpoint between multilateral and bilateral aid indirectly reflects the same member-surplus rationale, which I found to be at work in the NPT, currency, and trade regimes. Generally, to the extent that any multilateral regime treats members according to a set of fixed, impersonal criteria, it creates an inefficiency in the form of an agent surplus—a surplus distributed to members with compliance costs well below incentive levels set by impersonal rules.

### *Other Approaches to the Nuclear Proliferation Regime*

Existing studies do not seek to explain membership in the NPT regime but focus instead on the decision to pursue nuclear weapons programs or acquire nuclear weapons.<sup>61</sup> While nuclear proliferation has been given a good deal of scrutiny in the literature, I wish to restrict this survey to two approaches—quantitative and interpretive—that have recently received attention.

61. For a survey of more traditional approaches, see Sagan's (1996–97) useful typology.

On the quantitative side, three studies<sup>62</sup> agree on the importance of technological determinants, economic capacity, and the existence of a security guarantee in the decisions to launch a nuclear research program and acquire and deploy nuclear arms. Of the three, the two more recent studies assess the relevance of NPT membership to proliferation: while they both find a statistically significant negative relation between accession to the treaty and developing nuclear weapons, Jo and Gartzke also test for what they call the “systemic effect” of the regime, proxied by the number of total ratifications, and find no relation. They conclude that “these results cast doubt on the validity of constructivist arguments about the transformative effect of international agreements at the system level. . . .”<sup>63</sup>

The present analysis replicates most of these econometric results while using accession as a dependent variable, suggesting that the decisions to proliferate and join the NPT can be viewed as one and the same. The present study, however, cautions against Jo and Gartzke’s conclusion that the regime is irrelevant. Their non-finding only applies to the multilateral component of the NPT regime, which they test against the bilateral component (the existence of nuclear guarantees). Not surprisingly, they find the latter to be both significant and important in predicting the launching of a nuclear program, but too hastily credit the finding to *realpolitik* in contradistinction with the regime, overlooking the fact that nuclear guarantees are a key component of the regime as a whole.

On the interpretive side, a new set of studies have portrayed nuclear proliferation as the new threat that has come to dominate the security discourse in the wake of the disappearance of the Soviet threat.<sup>64</sup> The “securitization” of proliferation and the demonization of proliferators as “rogue states,” these authors argue, further distanced the NPT regime from disarmament, which, along with nonproliferation, was among its initial goals. This trend has fueled the antagonism of developing countries led by India, who denounce the discrimination between the nuclear haves and the nuclear have-nots. According to Mutimer,<sup>65</sup> the current framing of proliferation as the main threat to world stability constitutes the main reason for India’s consistent refusal to sign the NPT and its more recent refusal of signing the Comprehensive Test Ban Treaty.

This line of argument in many ways dovetails with the present analysis, which, more than any other, demonstrates the importance of bilateral security guarantees to the working of the regime. Security guarantees would not be credible if the guarantors were considering serious disarming. The drop in the credibility of Soviet security assurances in the late 1980s and their official withdrawal in 1991, for instance, had negative consequences for the NPT regime, prompting the accelera-

62. See Singh and Way 2004; Singer and Tago 2005; and Jo and Gartzke 2007.

63. Jo and Gartzke 2007, 13.

64. See Klare 1995; Litwak 2000; and Krause and Latham 1999.

65. Mutimer 2000.

tion of nuclear programs in North Korea, Iraq, and Iran.<sup>66</sup> Given the negative effects that disarming would have on guarantees, the price that India asked in order for its government to sign the NPT—disarmament—exceeds what the superpowers are willing to pay.

The difference, if any, between the interpretive argument and the present analysis resides in the future consequences of the Indian exclusion. Interpretivists may see it as the discursive basis for the mobilizing of widespread opposition among developing countries. I see it as an uncharacteristically high cost of compliance, affecting only India and a handful of countries finding themselves in a situation similar to that of India.

## Conclusion

I used the NPT regime to illustrate the idea that dyadic diplomacy is not incompatible with the building of a multilateral regime, but an efficient component thereof. I also showed that exclusion is a constituent part of that regime.

Starting from a characterization of proliferation as the result of large number of prisoner's dilemmas played out between countries engaged in local dyadic rivalries, I used contract theory to characterize as equilibrium a strategy profile in which states with low compliance costs joined the regime on the terms of the multilateral treaty alone; states with intermediate compliance costs needed additional customized incentives, delivered through bilateral agreements; and states with high compliance costs were not only left out of the regime but also punished for nonparticipating.

I offered systematic historical evidence of the exclusion hypothesis, showing that non-nuclear weapons states' membership in the global regime was a negative function of their compliance costs and reservation value—the cost of building nuclear weapons. I offered anecdotal evidence that several other important regimes in the currency, trade, and aid areas, offered multilateral and bilateral dimensions in proportions compatible with the optimization calculation formalized in the model.

Dyadic diplomacy is a necessary component of the NPT regime because the NPT, as with all multilateral instruments, generates an "agent surplus," collected by signatories who would have signed for much less than is being offered. It is this feature that made it impossible for the nuclear countries to write a collective security guarantee to all non-nuclear countries into the multilateral instrument itself, but required, instead, the use of bilateral diplomacy. The inefficiency of multilateralism is a cause for bilateralism that differs from the current literature's emphasis on enforcement. I suggested that the agent surplus inefficiency also characterized the trade and currency regimes as well as, with some modifications, the aid regime.

66. Lessenberry 2005.



The exclusion of countries with high compliance costs is another requisite component of the NPT regime, because the nuclear cartel is unable or unwilling to meet their costs. Trying to bring India within the fold of non-nuclear countries at this juncture would require the nuclear five to cut their nuclear armament and thus weaken their capacity to extend security guarantees to countries with intermediate compliance costs, thereby reopening the question of these countries' membership in, and compliance with, the regime in the same way as the withdrawal of military security guarantees by Russia following the end of the Cold War led Iraq, Iran, and North Korea to reconsider their obligations under the terms of the treaty.

It is hard to agree with some characterizations of the last seventy years as constituting a progressive shift either away from multilateralism to minilateralism or bilateralism,<sup>67</sup> or from bilateral and hegemonic regimes toward multilateral regimes.<sup>68</sup> Nor should the reassertion of bilateral diplomacy under the administration of U.S. President George W. Bush drive one to skepticism about the empirical relevance of multilateralism. The existence or resurgence of bilateral agreements in and of itself provides no evidence for one characterization over the other. Bilateralism is no proof of *realpolitik*, but could be evidence of its opposite—regime strength—in an area characterized by a high degree of subsidiarity. It must also be determined whether bilateral agreements complement or detract from the corresponding multilateral regime.

The article has consequences for the definition of an international regime. The commonality of norms is more important to the existence of a regime than the laterality of the instruments that carry out compliance. Notwithstanding Ruggie, the two dimensions should not be conflated but kept separate. Moreover, the degree of institutionalization of an issue, usually defined by how much of it falls under the supervision of a multilateral instrument or an international organization, is no indication of the strength of that regime. In some cases, as with the gold standard, the institutionalized part of the regime, like the legendary tip of the iceberg, hides a much larger body below its flotation line, whereas in other cases, such as Bretton Woods, foam floating on the surface of an ocean is the more apt comparison.

## Appendix: Proofs

### *Proposition (1)*

The principal's problem is to solve program  $P^1$  as specified in the text. Assume that  $\lambda, \theta, \gamma, \psi, \pi \geq 0$  are the Lagrangian parameters for each constraint in the order they are listed in  $P^1$ . The Kuhn-Tucker conditions are twofold: (1) a nonbinding constraint implies that its corresponding parameter is zero—call it KT1; (2) a positive parameter implies that

67. Yarbrough and Yarbrough 1992.

68. Ruggie 1992.

the constraint is binding—call it KT2. Forming the Lagrangian and deriving the first-order conditions with respect to  $a, d$ , and  $s$  yield

$$\lambda + \theta + \gamma = 2 \quad (\text{FOC}_a)$$

$$\lambda + \psi = 2\delta \quad (\text{FOC}_d)$$

$$\theta + \pi = 2\sigma \quad (\text{FOC}_s)$$

I do not solve the problem for all possible equilibria, but for one only, the mixed-incentive equilibrium.<sup>69</sup> A mixed-incentive equilibrium by definition is of the form  $a, d, s > 0$ , implying by KT1 that  $\gamma, \psi, \pi = 0$ . These zero values along with  $\text{FOC}_d$  and  $\text{FOC}_s$  imply that  $\lambda, \theta > 0$ , in turn yielding by KT2 the following system of three variables and two equations:  $\begin{cases} a = z - \frac{c-d}{1-q} \\ s = d + \frac{z-c}{1-q} q. \end{cases}$

Two variables are endogenously determined as a function of a given third. I choose to express both  $a$  and  $s$  as a function of a given value of  $d$  yielding the results listed in proposition (1).

To establish the parametric conditions under which the mixed-incentive equilibrium obtains, subtract  $\text{FOC}_d$  and  $\text{FOC}_s$  and then the result from  $\text{FOC}_a$ . Recalling that  $\gamma, \psi, \pi = 0$ , the calculation yields the unique circumstances of the mixed incentive equilibrium,  $1 = \delta + \sigma$ ; the marginal cost of the positive incentive is equal to the sum of the marginal costs of the negative incentives.

#### Proposition (2)

The principal now faces  $N$  dyads numbered according to the rivalry index  $z_i$  from 1 to  $N$ , such that  $z_1 = 1, z_2 = 2, \dots, z_N = N$ .  $z$  is distributed over the set  $(1, N]$  with  $z, N \in \mathbb{Z}^+$ , the set of positive integers.

Following lemma 1, assume that  $x$  and  $y$  partition the  $z$  segment into three sets, multilateral from 1 to  $x$ , multi-bilateral from  $x + 1$  to  $y$ , and exclusion from  $y + 1$  to  $N$ . A mixed instrument regime in equilibrium must satisfy the following set of inequalities:  $1 < x^* < y^* < N$ . This requires that  $x^*$  and  $y^*$  be internal maxima. Were the principal's objective function continuous, identifying such maxima would directly follow from showing concavity. Since the function is not continuous, additional cases must be considered.

First, I calculate  $U_P$  in program  $P^N$  as specified in the text. The summation term is equal to  $2(y - x)(V + c + \bar{d} - \bar{d}\delta - \sigma(\bar{d} - c\frac{q}{1-q}) - T) + (y(y + 1) - x(x + 1))(-q\frac{\sigma}{1-q} - 1)$ . Substituting this expression into  $U_P$  along with the value for  $g(x)$  and rearranging yields

$$U_P = Ax + Bx^2 + Cy + By^2 + E \text{ with}$$

$$A = 1 + 2T + q\sigma \frac{1}{1-q},$$

69. The mixed-incentive equilibrium is not a mixed equilibrium but an equilibrium in pure strategy that mixes incentives.

$$B = -1 - q\sigma \frac{1}{1-q},$$

$$C = -1 - 2T + 2V + 2c + 2\tilde{d} + 2\tilde{d}\rho - 2\tilde{d}\sigma - 2\tilde{d}\delta - q\sigma \frac{1}{1-q} + 2cq\sigma \frac{1}{1-q}, \text{ and}$$

$$E = -2N\tilde{d}\rho - T. \text{ Lemma 2 follows.}$$

*Lemma 2.  $U_P$  is concave.*

**Proof.** Concavity requires that for any pair of distinct points  $(x_1, y_1)$  and  $(x_2, y_2)$  in the domain of  $\hat{U}_P$ , and for  $0 < \theta < 1$ , the following weak inequality holds:  $\theta U_P(x_1, y_1) + (1 - \theta) U_P(x_2, y_2) \leq U_P(\theta(x_1, y_1) + (1 - \theta)(x_2, y_2))$ . Developing yields  $\theta(Ax_1 + Bx_1^2 + Cy_1 + By_1^2 + E) + (1 - \theta)(Ax_2 + Bx_2^2 + Cy_2 + By_2^2 + E) \leq A(\theta x_1 + (1 - \theta)x_2) + B(\theta x_1 + (1 - \theta)x_2)^2 + C(\theta y_1 + (1 - \theta)y_2) + B(\theta y_1 + (1 - \theta)y_2)^2 + E$ . Rearranging and simplifying, one obtains  $\theta(1 - \theta)B((x_1 - x_2)^2 + (y_1 - y_2)^2) \leq 0$ , which is true since  $B$  is negative.

One can now prove proposition (2).

**Proof.** The principal's problem is to choose  $x$  for a given  $y$  so that  $x \in S_x|_y$  with  $S_x|_y = \{x \in \mathbb{Z}^+ | U_P(x|y) \geq U_P(x + 1|y) \text{ and } U_P(x|y) \geq U_P(x - 1|y)\}$  and to choose  $y$  for a given  $x$  so that  $y \in S_y|_x = \{y \in \mathbb{Z}^+ | U_P(y|x) \geq U_P(y + 1|x) \text{ and } U_P(y|x) \geq U_P(y - 1|x)\}$ , with  $1 < x^* < y^* < N$ . Consider  $x$  first. Define  $\hat{U}_P$  as the principal's objective function  $U_P$  defined on the continuous domain  $[1, N] \times [1, N]$ . Define  $\hat{x}^*$  as the maximum of  $\hat{U}_P$  for a given  $y$ . From the definition of  $U_P$  above,  $\hat{x}^* = \frac{1-q}{1-q+\sigma q} T + \frac{1}{2}$  while  $x^* \in S_x|_y$  implies  $\frac{1-q}{1-q+\sigma q} T \leq x^* \leq 1 + \frac{1-q}{1-q+\sigma q} T$ . Four types of solutions are possible:

- (i)  $\hat{x}^* \in \mathbb{Z}^+$  and thus  $\hat{x}^* \in S_x|_y$ . It is unique because  $\hat{x}^* + 1$  or  $\hat{x}^* - 1$  fall out of the allowed range and do not belong to  $S_x|_y$ .
- (ii)  $\hat{x}^* \notin \mathbb{Z}^+$  and  $U_P([x^*]|y) > U_P([x^*] + 1|y)$ , with  $[x^*]$  defined as the nearest integer below  $\hat{x}^*$ : it can be shown that  $[x^*] \in S_x|_y$  because  $\frac{\partial \hat{U}_P}{\partial x} > 0$  on  $(1, [x^*])$  implies that  $U_P([x^*]|y) > U_P([x^*] - 1|y)$ . Moreover, there is no other  $x$  that meets these conditions. Indeed, assume there exists  $x' < [x^*]$  with  $x' \in \mathbb{Z}$ ;  $\frac{\partial \hat{U}_P}{\partial x} > 0$  on  $(1, [x^*])$  implies that  $U_P(x'|y) < U_P(x' + 1|y)$ .
- (iii)  $\hat{x}^* \notin \mathbb{Z}^+$  and  $U_P([x^*]|y) < U_P([x^*] + 1|y)$ : a reasoning similar to that in (ii) yields  $[x^*] + 1$  as the sole maximizer.
- (iv)  $\hat{x}^* \notin \mathbb{Z}^+$  and  $U_P([x^*]|y) = U_P([x^*] + 1|y)$ : this case affords two maximizers,  $x_1^* = \frac{1-q}{1-q+\sigma q} T$  and  $x_2^* = 1 + \frac{1-q}{1-q+\sigma q} T$ , with  $[x^*] = \frac{1-q}{1-q+\sigma q} T$ .

Similar reasoning for  $y^*|x$  yields  $\hat{y}^* = \frac{(v - T + (1 - \delta - \sigma + \rho)\tilde{d} + c(\frac{1+\sigma q}{1-q}))(1-q)}{1-q+\sigma q} - \frac{1}{2}$ ,  $\frac{(v - T + (1 - \delta - \sigma + \rho)\tilde{d} + c(\frac{1+\sigma q}{1-q}))(1-q)}{1-q+\sigma q} - 1 \leq y^* \leq \frac{(v - T + (1 - \delta - \sigma + \rho)\tilde{d} + c(1 + \frac{\sigma q}{1-q}))(1-q)}{1-q+\sigma q}$ , and four different cases.

Furthermore, assuming the particular case in which  $x^* = \frac{1-q}{1-q+\sigma q} T$  and  $y^* = \frac{(V-T+(1-\delta-\sigma+\rho)\tilde{d}+c\left(1+\frac{\sigma q}{1-q}\right))(1-q)}{1-q+\sigma q}$ , equilibrium values  $x^*$  and  $y^*$  must satisfy 3 constraints:

$$0 < x^* \Rightarrow T > 0,$$

$$x^* < y^* \Rightarrow T < \frac{1}{2} \left( V + c \left( \frac{1-q(1-\sigma)}{1-q} \right) + \tilde{d}(1-\sigma-\delta+\rho) \right), \text{ and}$$

$$y^* < N \Rightarrow T > V + (c-N) \left( \frac{1-q(1-\sigma)}{1-q} \right) + \tilde{d}(1-\sigma-\delta+\rho).$$

### Comparative Statics

Despite discontinuity, comparative statics is possible on  $U_P$  because it is monotonic in all its parameters.  $\frac{\Delta x^*}{\Delta V} = 0$ ;  $\frac{\Delta y^*}{\Delta V} > 0$ .  $\frac{\Delta x^*}{\Delta c} = 0$ ;  $\frac{\Delta y^*}{\Delta c} > 0$ .  $\frac{\Delta x^*}{\Delta \rho} = 0$ ;  $\frac{\Delta y^*}{\Delta \rho} > 0$ .  $\frac{\Delta x^*}{\Delta \delta} = 0$ ;  $\frac{\Delta y^*}{\Delta \delta} > 0$  if the sum of the marginals on the delivered incentives is greater than the sum of the marginals on the threats ( $1 + \rho \geq \delta + \sigma$ ),  $\frac{\Delta x^*}{\Delta \sigma} < 0$ ;  $\frac{\Delta y^*}{\Delta \sigma} < 0$ , with  $\frac{\Delta x^*}{\Delta \sigma} < \frac{\Delta y^*}{\Delta \sigma}$  if  $\tilde{d} < \frac{1+T-V}{1+q(\rho-\delta)} q$  and  $\frac{\Delta x^*}{\Delta \sigma} > \frac{\Delta y^*}{\Delta \sigma}$  otherwise.  $\frac{\Delta x^*}{\Delta \delta} = 0$ ;  $\frac{\Delta y^*}{\Delta \delta} < 0$ .  $\frac{\Delta x^*}{\Delta q} < 0$ ;  $\frac{\Delta y^*}{\Delta q} < 0$ , with  $\frac{\Delta x^*}{\Delta q} < \frac{\Delta y^*}{\Delta q}$  if  $\frac{\Delta y^*}{\Delta q} \cdot \frac{\Delta x^*}{\Delta T} > 0$ ;  $\frac{\Delta y^*}{\Delta T} < 0$ .

**TABLE A1.** *List of variables*

#### Exogenous variables

- V Principal's value for one individual agent's compliance.
- z Agent's compliance cost.
- c Agent's cost of pursuing nuclear weapons program.
- q Probability that cheating agent is not caught.
- N Highest cost of compliance; number of dyads.
- T Transaction costs.
- $\sigma$  Principal's marginal cost for threatening to sanction cheating agent.
- $\delta$  Principal's marginal cost for threatening to sanction rejecting agent.
- $\varepsilon$  Principal's marginal cost for sanctioning cheating.
- $\rho$  Principal's marginal cost for sanctioning rejection.

#### Choice variables

- a Transfer to agent who is declared having complied.
- s Sanction cost for agent who is declared having cheated.
- d Sanction cost for agent who rejects regime.
- x Compliance cost of the agent who is indifferent between receiving bilateral incentives or not in addition to multilateral incentives.
- y Compliance cost of the agent who is indifferent between inclusion and exclusion.

## References

- Albright, David. 1994. South Africa and the Affordable Bomb. *Bulletin of the Atomic Scientists* 50 (4):37–47.
- Albright, David, and Corey Gay. 1998. Taiwan: Nuclear Nightmare Averted. *Bulletin of the Atomic Scientists* 54 (1):54–60.
- Bennett, D. Scott. 1998. Integrating and Testing Models of Rivalry Termination. *American Journal of Political Science* 42 (4):1200–32.

- Bennett, D. Scott, and Allan Stam. 2000. EUGene: A Conceptual Manual. *International Interactions* 26 (2):179–204.
- Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge, Mass.: Cambridge University Press.
- Burley, Anne-Marie, and Walter Mattli. 1993. Europe before the Court: A Political Theory of Legal Integration. *International Organization* 47 (1):41–76.
- Clausen, Peter A. 1993. *Nonproliferation and the National Interest*. New York: HarperCollins.
- Cole, Paul M. 1997. Atomic Bombast: Nuclear Weapon Decision-making in Sweden, 1946–72. *Washington Quarterly* 20 (2):233–51.
- Dominguez, Kathryn M. 1993. The Role of International Organizations in the Bretton Woods System. In *A Retrospective on the Bretton Woods System: Lessons for International Monetary Reform*, edited by Michael D. Bordo and Barry Eichengreen, 357–404. Chicago: University of Chicago 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.
- Drezner, Daniel W. 1999. *The Sanctions Paradox. Economic Statecraft and International Relations*. Cambridge, Mass.: Cambridge University Press.
- Fearon, James D. 1998. Bargaining, Enforcement, and International Cooperation. *International Organization* 52 (2):269–305.
- Forland, Astrid. 1997. Norway's Nuclear Odyssey: From Optimistic Proponent to Nonproliferator. *The Nonproliferation Review* 4 (2):1–16.
- Gallarotti, Giulio M. 1994. The Scramble for Gold: Monetary Regime Transformation in the 1870s. In *Monetary Regimes in Transition*, edited by Michael D. Bordo and Forrest Capie, 15–67. Cambridge, Mass.: Cambridge University Press.
- Gleditsch, Kristian Skrede. 2002. Expanded Trade and GDP Data. *Journal of Conflict Resolution* 46 (5):712–24.
- Grambsch, Patricia M., and Terry M. Therneau. 1994. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika* 81 (3):515–26.
- Jagers, Keith, and Ted R. Gurr. 1995. Tracking Democracy's Third Wave with the Polity III Data. *Journal of Peace Research* 32 (4):469–82.
- Jo, Dong-Joon, and Erik Gartzke. 2007. Determinants of Nuclear Weapons Proliferation. *Journal of Conflict Resolution* 51 (1):167–94.
- Kaiser, Karl. 1980. Nuclear Energy and Nonproliferation in the 1980s. In *Nuclear Policy in Europe: France, Germany and the International Debate*, edited by Erwin Häckel, Karl Kaiser, and Pierre Lellouche, 1–30. Forschungsinstitut der Deutschen Gesellschaft für Auswärtige Politik e.V., Bonn, Germany.
- Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton, N.J.: Princeton University Press.
- Klare, Michael. 1995. *Rogue States and Nuclear Outlaws: America's Search for a Foreign Policy*. New York: Hill and Wang.
- Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. The Rational Design of International Institutions. *International Organization* 55 (4):761–800.
- Krasner, Stephen D. 1991. Global Communications and National Power: Life on the Pareto Frontier. *World Politics* 43 (3):336–366.
- Krause, Keith, and Andrew Latham. 1999. Constructing Non-Proliferation and Arms Control: The Norms of Western Practice. In *Culture and Security*, edited by Keith Krause, pp. 23–54. London: Frank Cass.
- Kreps, David, and Robert Wilson. 1981. Reputation and Imperfect Information. *Journal of Economic Theory* 27 (2):253–79.
- Laffont, Jean-Jacques, and David Martimort. 2002. *The Theory of Incentives: The Principal-Agent Model*. Princeton, N.J.: Princeton University Press.
- Leeds, Brett Ashley. 2005. Alliance Treaty Obligations and Provisions (ATOP) Codebook. Department of Political Science. Rice University, Houston, Texas. Available at (<http://atop.rice.edu/download/ATOPcdbk.pdf>). Accessed 2 October 2007.

- Lessenberry, Brian K. 2005. The Rise of the Rogues: Reassessing Twenty-First Century Nuclear Proliferation. Paper delivered at the 2005 Annual Meeting of the American Political Science Association, September, Washington, D.C.
- Levite, Ariel E. 2002-3. Never Say Never Again: Nuclear Reversal Revisited. *International Security* 27 (3):59–88.
- Litwak, Robert S. 2000. *Rogue States and US Foreign Policy: Containment after the Cold War*. Washington, D.C.: Woodrow Wilson Center.
- Lohmann, Susanne. 1997. Linkage Politics. *Journal of Conflict Resolution* 41 (1):38–67.
- Martin, Lisa L. 1992. Interests, Power, and Multilateralism. *International Organization* 46 (4):765–92.
- Mearsheimer, John J. 1994-95. The False Promise of International Institutions. *International Security* 19 (3):5–49.
- Milner, Helen V. 2006. Why Multilateralism? Foreign Aid and Domestic Principal-Agent Problems. In *Delegation and Agency in International Organizations*, edited by Darren G. Hawkins, David A. Lake, Daniel L. Nielson, and Michael J. Tierney, pp. 107–139. Cambridge: Cambridge University Press.
- Mutimer, David. 2000. *The Weapons State: Proliferation and the Framing of Security*. London: Lynne Rienner.
- Oye, Kenneth A. 1986. *Cooperation Under Anarchy*. Princeton, N.J.: Princeton University Press.
- Pahre, Robert. 1994. Multilateral Cooperation in an Iterated Prisoner's Dilemma. *Journal of Conflict Resolution* 38 (2):326–52.
- . 2001. Most-favored-Nation Clauses and Clustered Negotiations. *International Organization* 55 (4):859–90.
- Ribicoff, Abraham A. 1976. A Market-Sharing Approach to the World Nuclear Sales Problem. *Foreign Affairs* 54 (4):763–87.
- Rodrick, Daniel. 1996. Why Is There Multilateral Lending? In *Annual World Bank Conference on Development Economics, 1995*, edited by Michael Bruno and Boris Pleskovic, 167–903. Washington, D.C.: International Monetary Fund.
- Rowe, Edward Thomas. 1978. National Attributes Associated with Multilateral and US Bilateral Aid to Latin America: 1960–1971. *International Organization* 32 (2):463–75.
- Ruggie, John Gerard. 1992. Multilateralism: The Anatomy of an Institution. *International Organization* 46 (3):561–98.
- Sagan, Scott D. 1996–1997. Why Do States Build Nuclear Weapons? Three Models in Search of a Bomb. *International Security* 21 (3):54–86.
- Singer, J. David, and Atsushi Tago. 2005. The Proliferation of Nuclear Weapons: Correlates and Consequences. Paper presented at the Annual Meeting of the American Political Science Association, September. Washington, D.C.
- Singh, Sonali, and Christopher R. Way. 2004. The Correlates of Nuclear Proliferation: A Quantitative Test. *Journal of Conflict Resolution* 48 (6):859–85.
- Sidnal, Duncan. 1985. The Limits of Hegemonic Stability Theory. *International Organization* 39 (4):579–614.
- . 1991. Relative Gains and the Pattern of International Cooperation. *American Political Science Review* 85 (3):701–26.
- Strulak, Tadeusz. 1993. The Nuclear Suppliers Group. *The Nonproliferation Review* 1 (1):2–10.
- Verdier, Daniel. 2002. *Moving Money: Banking and Finance in the Industrialized World*. Cambridge, Mass.: Cambridge University Press.
- Williamson, Oliver E. 1975. *Markets and Hierarchies: Analysis and Anti-trust Implications*. New York: Free Press.
- Winters, Alan L. 1990. The Road to Uruguay. *Economic Journal* 100 (403):1288–303.
- Yager, Joseph A. 1985. Nuclear Supplies and the Policies of South Korea and Taiwan toward Nuclear Weapons. In *The Nuclear Suppliers and Nonproliferation*, edited by Rodney W. Jones, Cesare Merlini, Joseph F. Pilat, and William C. Potter, 187–195. Lexington, Mass.: Lexington Books.
- Yarborough, Beth V., and Robert M. Yarborough. 1992. *Cooperation and Governance in International Trade: The Strategic Organization Approach*. Princeton, N.J.: Princeton University Press.