Contracting around International Uncertainty

BARBARA KOREMENOS *University of Michigan*

International cooperation is plagued by uncertainty. Although states negotiate the best agreements possible using available information, unpredictable things happen after agreements are signed that are beyond states' control. States may not even commit themselves to an agreement if they anticipate that circumstances will alter their expected benefits. Duration provisions can insure states in this context. Specifically, the use of finite duration depends positively on the degree of uncertainty and states' relative risk aversion and negatively on the cost. These formally derived hypotheses strongly survive a test with data on a random sample of agreements across all four of the major issue areas in international relations. Not only do the results, highlighting evidence on multiple kinds of flexibility provisions, strongly suggest that the design of international agreements is systematic and sophisticated; but also they call attention to common ground among various subfields of political science and law.

In an international environment characterized by uncertainty that is at times so pronounced that it could easily make states fearful of making meaningful commitments, how is cooperation possible? States often have large amounts of information at their disposal when initiating cooperative activity and use it wisely to set the terms of cooperation and to manage its evolution. Inevitably, though, events occur that could not have been foreseen and cannot be controlled. Why do states commit themselves to cooperation in the first place if they expect that circumstances might reduce their anticipated benefits? Not only do states cooperate; they also codify their cooperation through countless international agreements that form a substantial body of international law. How is this possible?

One answer is that perhaps cooperative efforts are undertaken only in the areas and among states for which uncertainty is less pronounced. This may very well be the case given that, according to the conventional wisdom in international relations, a main objective of international cooperation is to make states' commitments to each other more credible through hands-tying. If we acknowledge both the conventional wisdom and the pervasive uncertainty in the international environment, we might conclude that, because states do not like to tie their hands under conditions of high uncertainty, cooperation ensues only when uncertainty is low.

Does this imply that countless other cooperative possibilities go unrealized because states cannot "insure" themselves against the unanticipated negative consequences that uncertainty might bring? Or should we entertain the possibility that states can somehow protect themselves in situations of high uncertainty?

At first glance, this second possibility seems farfetched. After all, international cooperation takes place in anarchy, so the idea that international cooperation could be designed to provide some kind of insurance for states is hard to imagine. On the other hand, the need for such a scheme is especially pressing in international relations given the pronounced uncertainty that operates in that sphere. I show that such a scheme is possible, notwithstanding the complications imposed by anarchy.

Taking as my point of departure economic contracting theory, I develop a model that features flexibility rather than hands-tying to make such protection possible. The specific flexibility provision that serves as an "international insurance scheme" is a limited duration agreement that can be renegotiated.[1] I subject the model to empirical testing with data on a random sample of agreements drawn from the United Nations Treaty Series (UNTS). This is the first study that uses a random sample across all four of the major issue areas in international relations to examine matters of international agreement design.[2]

This study also forces us to consider whether international relations scholars are missing opportunities by not examining the extent to which there is common ground with scholars in some of the other subfields of political science. More specifically, unlike in international relations, in much of American and comparative politics the study of institutions and institutional design has gained prominence, with much of that work drawing on economic theory. If international law follows a logic similar to one developed for economic contracts, might there not be much more in common among the various subfields of political science than traditionally

Barbara Koremenos is Associate Professor of Political Science, University of Michigan, 5700 Haven Hall, 505 South State St., Ann Arbor, MI 48109-1045 (koremeno@polisci.ucla.edu).

This article draws on research carried out under a National Science Foundation CAREER Award: "Designing International Agreements: Theoretical Development, Data Collection, and Empirical Analysis" (SES-0094376). Several colleagues gave me helpful comments on various drafts of this work. In particular, I thank Beth Simmons, Matt Baum, Jerry Cohen, Ron Mitchell, Rip Smith, Jim Morrow, Bob Powell, David Lake, Duncan Snidal, and Paul Diehl. I thank seminar participants at Chicago, Columbia, and UCSD for their feedback, and Jana von Stein for valuable research assistance. The comments of editor, Lee Sigelman, and the three anonymous referees greatly improved this article. Finally, I thank Jeffrey Smith for many stimulating discussions as well as detailed comments on numerous drafts of this paper and George Tsebelis for helping me see the value of my work in a broader context.

[1] In models of the domestic labor context, such flexibility can insure actors against certain kinds of price shocks.

[2] Although this is the first study to use a random sample to examine issues of agreement design, there are a number of analyses featuring careful large-n empirical work in the field of international cooperation, including Simmons 2000, Guzman and Simmons 2002, and Martin 1992, 2000.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

acknowledged? Given that over 50 thousand international agreements are ready for us to analyze, the results of this analysis can help guide both future research within international relations and collaborative efforts across subfields and even disciplines.

EXPLAINING THE STRUCTURE OF INTERNATIONAL AGREEMENTS

This article adds to a small but growing theoretical literature on how states choose the flexibility provisions of international agreements. A few scholars consider an environment in which a state's gain from an agreement is affected by infrequent, transitory negative shocks. In the area of monetary policy under the gold standard, Bordo and Kydland (1995) show how states used "escape clauses" consisting of temporary suspensions of convertibility to deal with banking crises. Downs and Rocke (1995) and Rosendorff and Milner (2001) consider similar escape-clause models in a trade context, where the shocks represent the sudden mobilization of domestic interest groups in response to particular aspects of trade agreements.

A few economists, for example, Gray (1978), Dye (1985), and Harris and Holmstrom (1987), consider an environment in which some underlying random variable affects the payoffs to the parties from a contract under the assumption that the parties cannot make the contract contingent on this variable. These models focus on how the parties choose the optimal duration of their contracts in the presence of uncertainty.

In earlier work, I have considered an environment in which a cooperative endeavor is characterized by a one-time distributional uncertainty and states have to learn about it over time (Koremenos 2001). In that context, states may limit the duration of their agreement and renegotiate it once after enough information has been revealed about the initial uncertainty. My case studies included the Nuclear Non-Proliferation Treaty, for which the relatively fixed distribution of gains between the nuclear weapon states and the non nuclear weapons states was initially unknown. While drawing on the economic models, that analysis adds some refinements (including the ability of states to renege on their contracts) to better capture the essential features of the international environment.

The model presented here builds on my prior analysis, with one key difference: the uncertainty is persistent; hence, there is no learning. Thus, the model here generalizes the earlier results by considering a different kind of distributional uncertainty and, in response, a *series* of renegotiations.

The Theory

To characterize *persistent* uncertainty surrounding the distribution of future gains from an agreement, I assume that the parties select an initial distribution of gains based on their relative bargaining power, but this gain evolves over time under the agreement due to external shocks. The precise nature of the shocks depends on the issue area but could include climatic fluctuations (e.g., commodity agreements), exogenous private sector shocks (e.g., exchange rates), or political shocks (e.g., refugee agreements). I assume that the parties always know the distribution of gains in the current period, but know only the probability distribution for the distributions of gains in all future periods. The distribution of gains in each period equals the distribution in the previous period plus a random shock. These random shocks are cumulative, so the distribution of gains, left to its own devices, follows a random walk.

Put simply, a shock to the distribution of gains occurs in each period. As a result, either the gains to one state increase relative to that of the other, or there is no change if the shock equals zero. These shocks have three important features. First, they are completely random. Each shock is independent of every other, both before and after. They are also independent of everything else that the states can observe, which implies that they cannot be predicted, either at the time an initial agreement is made or later on. Second, the precise value of any shock is discoverable only at a cost. In other words, shocks are noisily observable. Third, the shocks add up over time. That is, the gain for a state in any one period is the sum of its initial gain plus each of the subsequent changes. Because the random shocks are cumulative, it is possible for the states' relative shares to evolve in a way that differs quite substantially from the assumed equal division at the start of the agreement. How great the differences become depends both on how much time has passed and on the specific sequence of realized shocks. As time passes, the potential divergence from the initial equal division increases because the passage of time allows, but does not require, the random walk process to walk farther away from its starting point. Hence, once a policy course is chosen, without subsequent redirection, the system may continue down particular paths longer than the parties originally intended. I refer to this environment as one characterized by "persistent uncertainty."

The extent or importance of the *uncertainty* is summarized by the variance of the shocks to the distribution of gains, with a larger variance implying greater uncertainty. Also, the larger the variance, the farther the distribution of gains will depart, on average, from its initial value over a given time.

Of course, states do not embrace such uncertainty; rather, the idea that they may gain far less than they anticipated from an agreement could make some of them too nervous to commit to one. To capture this, I employ a standard assumption in international relations that states are *risk averse.* I do, however, allow them to vary with respect to their *relative* level of risk aversion (see O'Neill 2001).

Types of Agreements

Suppose states determine that there are gains from cooperation and decide to negotiate an agreement. They have two options in a given agreement context:

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

(1) an agreement of indefinite duration (an inflexible agreement) or (2) a series of agreements renegotiated at regular intervals in order to adjust the distribution of gains for the effects of the shocks that cumulate during each agreement (a flexible agreement).

It is important to note a few alternatives that are not available. I exclude a completely contingent agreement, consistent with my assumption that it is costly to uncover the exact value of the shock to the distribution of gains. Most international agreements are multidimensional and are characterized by at least some gains that are not easily measurable; hence, it is a complex process to assign values to all possible outcomes. Thus, once a shock has occurred, some kind of meeting among the member states is usually necessary to bring together and analyze all of the information each has individually gathered. Such a meeting enables comparisons across an agreement's multiple dimensions regarding the true winners and losers. Because of the complexity of some of the issue areas, third-party "neutral" expertise is at times solicited. For example, in the environmental issue area, experts are sometimes invited to testify at meetings of the member states.

I also exclude the delegation of flexibility to an international organization, an option found in Koremenos 2000, which models explicitly the creation of an international organization as a choice variable. Current data limitations do not allow for testing the fuller model. In random samples, there is a predominance of bilateral agreements, which rarely call for the creation of international organizations.[3]

Consider first an indefinite duration agreement. States pay initial negotiation costs, but no additional costs thereafter. Given that there is no provision for renegotiating the distribution of gains in response to shocks, the distribution may wander well away from its initial level.

There is no international authority to enforce agreements. Hence, should the distribution of gains move sufficiently far enough away from the initial level, one state may break the agreement. The state that elects to do this pays a cost. The costs of reneging consist of any sanctions imposed by the other parties or by other international actors. These costs may be substantial. I assume that once reneging costs are paid, the states begin again with a new agreement with an indefinite duration. Put differently, I assume that the reneging state incurs no cost in terms of reputation or refusal to bargain, thereby avoiding problems of renegotiation-proofness.

Instead of one indefinite duration agreement, states can choose to conclude a series of finite duration agreements of equal duration, with renegotiation taking place between each pair of agreements in the series. I consider only series of agreements that are of the same duration because of my assumption that the environment facing the parties does not change over time. As a result of this assumption, the states will make the same choice regarding agreement duration at the start of the first agreement and every subsequent agreement because the choice problem they face in each case does not change.

States choosing to conclude a series of finite duration agreements incur two kinds of cost. First, states incur the negotiation costs required to reach the initial agreement in the series. I assume that these are the same as in the case of a single agreement of indefinite duration, although they could be less if a shorter shadow of the future results in lower bargaining costs. Second, states incur *renegotiation costs* at the start of each new agreement in the series. Like negotiation costs, renegotiation costs vary depending on the states in the agreement and the environment. For example, I expect both costs to increase with the number of parties to the agreement.[4] In addition, with respect to renegotiation costs, the greater number of parties produces greater amounts of information that have to be analyzed to discover both the nature of the shocks and how the terms of the agreement must be adjusted to cancel out their cumulative effect.

The advantage that states derive from concluding a series of renegotiated agreements rather than one agreement of indefinite duration is flexibility: the division of gains can be *reset* to the initial level at regular intervals. States choose to reset the division of gains to its original level because I assume that bargaining power and other factors that influence the division do not change over time.

The planned readjustment to the division of gains that occurs under a series of renegotiated agreements greatly reduces the chance that either state will want to incur the costs of reneging or be forced to endure an unsatisfactory division of gains for long periods. Put differently, for risk-averse states, the opportunity periodically to reset the distribution of gains back to its initial level increases the *ex ante* value of a series of renegotiated agreements relative to a single indefinite agreement. It does so by reducing the variance of the discounted expected gain from the agreement. States that choose to renegotiate must decide how often to do so. Doing so involves trading off the costs of more frequent renegotiation against the costs of living with a distribution of gains that differs from that initially chosen in light of the parties' relative bargaining power.

Foreshadowing the formal results, states will choose to renegotiate more often when *renegotiation costs* fall. States will also renegotiate more often when the *variance of the shocks* increases because this increases the expected deviation between the realized and the desired distributions of gains and when their level of *risk aversion* increases because the variance is more costly to them.

Suppose that the variance of the shocks to the distribution of gains is very low so the benefits from flexibility are small and renegotiation costs are high. In this

[3] The UNTS data at the time my sample was drawn (1999) consisted of 32,939 bilateral agreements and only 2,330 multilateral ones.

[4] Conybeare (1985) makes this argument in relation to trade talks, noting that the multilateral Kennedy Round (1963–67) took considerably longer than the previous bilateral Dillon Round (1960–61).

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

case, states may choose an indefinite agreement. The 1967 Outer Space Treaty, which forbids the placement of weapons of mass destruction into orbit, onto celestial bodies, or in outer space, is not subject to great uncertainty of this kind. There are no supply-and-demand shocks, and no resource or territorial issues are at stake. Rather, the goal is to lock in the status quo and prevent the militarization of space. The states involved therefore chose not to incur renegotiation costs; the agreement is of indefinite duration.

By contrast, if the variance of shocks to the distribution of gains is large so that it quickly departs from its initial level, states may want to limit the duration of their agreement. The Group of 7 (G-7) macroeconomic cooperation is characterized by low renegotiation costs, given the small number of parties and the low-cost availability of economic information, and by a high variance of shocks to the distribution of gains, given a rapidly changing world economy. Thus the G-7 has traditionally chosen a series of short, renegotiated agreements. In the case of the Group of five finance ministers, costs are lower, the variance of the shocks is higher, and agreements are even shorter.[5]

Formal Model

I consider two states, State 1 and State 2. Let $Y_{1,t} = b_{1,t}$, $Y_{2,t} = b_{2,t}$, denote the outcomes for the two states in the absence of an agreement. The first subscript indicates the state, and the second indicates the time period in each case. These outcomes depend on the specific issue area under consideration, but may include things like GDP, trade levels, or exchange rate stability, appropriately measured. For simplicity, I assume that the base outcomes do not change over time.

In the presence of an agreement, the outcomes attained by the two parties to an agreement in each period t are given by $Y_{1,t} = b_{1,t} + m_t$, $Y_{2,t} = b_{1,t} + (g - m_t)$, where g is the total gain from the agreement (the size of the pie) and m_t is the portion received by State 1. For simplicity, I assume a fixed total gain g. Rather than having the total gain vary over time, I have only the division of gains between the parties vary over time. Although this sacrifices some realism, it simplifies the model and allows me to focus on the key issue of how states structure agreements in the face of distributional uncertainty.

To take account of the rapidly changing context of international agreements, I assume that in the absence of renegotiation, the distribution of gains evolves according to a random walk, with $m_t = m_{t-1} + e_t$, where m_0 is chosen under the agreement and where e_t has density function $h(e_t)$. I assume, without loss of generality, that the states are symmetric. This implies equal bargaining power and, under a Nash bargaining solution, equal initial shares, so that $m_0 = 0.5g$.[6]

I assume that $E(e_t) = 0$ and that e_t is independently and identically distributed across periods. The parties observe m_t in each period. Let $t = 0$ denote the first period of an agreement. By choosing m_0, the parties to the agreement are choosing the expected value, as of $t = 0$, of m_t for all future periods. Because the e_t are independent across periods and have a common variance, we have

$$m_t = m_0 + e_1 + e_2 + \cdots + e_t, \quad \text{and}$$
$$var(m_t) = var(g - m_t) = t * var(e_t).$$

That is, because the shocks are independent and cumulative, the set of possible distributions of gains from the agreement fans out over time as the agreement continues. Put differently, the probability that the realized value of the distribution of gains differs by any fixed amount from the initial choice of m_0 increases over time. In the general case considered here, the gain for one party (either m or $(g - m)$) from the agreement may become negative (though an agreement whose initial gain was negative for one or both parties would, obviously, not elicit much interest).

I denote the cost of negotiating a single, infinite duration agreement or the first in a series of renegotiated agreements by k_n. When an agreement gets renegotiated, states pay a renegotiation cost k_r.These costs are incurred by all of the parties to an agreement. I interpret both k_n and k_r much more broadly than is done in the economics literature. In addition to the costs of sitting down and talking, I treat both as including the costs of assembling or reassembling domestic political coalitions—often necessary for agreement ratification. Renegotiation costs also include the costs of discovering the value of e_t.

As noted earlier, when states renegotiate an agreement, they reset the distribution of gains under the agreement to that originally agreed upon. This follows from my assumption of stable relative bargaining power. This means that $m_t = m_0$ in periods immediately following a renegotiation.

To focus on the basic tradeoffs, consider a simple two-period case. The parties must choose between one 2-period agreement (an inflexible agreement) and two 1-period agreements (a flexible agreement in that it can be renegotiated). If they choose two 1-period agreements, each gains a reduction in the variance of its gain from $2 * var(e_t)$ to $var(e_t)$, but loses the additional renegotiation cost. The expected value (gross of negotiation and renegotiation costs) is the same in both cases; consequently risk-neutral parties would opt for the 2-period agreement, as variance does not matter to them.

[5] It could be argued the more an agreement matters, the more important is flexibility. In fact, why would states pay for flexibility in a shallow or trivial agreement? An additional variable measuring importance should perhaps be interacted with the uncertainty variable. Unfortunately, that measure does not exist and would be extremely difficult to create in any objective sense.

[6] This cooperative solution corresponds to the Rubinstein alternating-offers noncooperative solution when δ is close to 1. See Osborne and Rubenstein 1990.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

Formally, the parties compare their expected utilities with one 2-period agreement, given by

$$E(u(b_1 + m_0 + e_1 - k_n) + \delta u(b_1 + m_0 + e_1 + e_2)),$$
$$E(u(b_2 + (g - (m_0 + e_1)) - k_n) + \delta u(b_2 + (g - (m_0 + e_1 + e_2)))),^{7}$$

to their expected utilities from two 1-period agreements, given by

$$E(u(b_1 + m_0 + e_1 - k_n) + \delta u(b_1 + m_0 + e_2 - k_r)),$$
$$E(u(b_2 + (g - (m_0 + e_1)) - k_n) + \delta u(b_2 + (g - (m_0 + e_2)) - k_r)).$$

The parties choose the agreement type that provides the higher expected utility.

Comparative Statics

The following hypotheses implicitly hold everything else equal so they indicate the effect of changing one parameter of the model holding the others constant. The proofs are in Appendix A. These hypotheses focus on the effects of variation in states' characteristics and in the agreement context on the type of agreement selected. I also present hypotheses regarding the choice of agreement duration, should states choose to conclude a series of finite duration agreements. I have also simulated a version of the model that includes the possibility of reneging, using a variety of values for the key parameters: the level of risk aversion, renegotiation costs, and the variance of the shocks to the distribution of gains. The simulations support the comparative statics (see Appendix B).[8]

(CS-1) As **renegotiation costs** increase, the probability that the parties will choose finite, renegotiated agreements decreases.

(CS-2) If the parties conclude a series of renegotiated agreements, as **renegotiation costs** increase, parties will choose to make each agreement in the series longer.

An increase in renegotiation costs (CS-1) raises the costs of choosing a series of renegotiated agreements relative to an indefinite duration. With respect to (CS-2), the costs of renegotiation increase while the benefits of renegotiation remain unchanged; thus, the parties will renegotiate less frequently by making each agreement in the series longer.

(CS-3) As **uncertainty** increases, the probability that the parties will choose finite, renegotiable agreements to adjust for shocks increases.

(CS-4) If the parties conclude a series of renegotiated agreements, as **uncertainty** increases, the parties will choose to make each agreement in the series shorter.

An increase in uncertainty (CS-3) makes the parties value flexibility more. More specifically, it increases the variation in realized outcomes under an indefinite duration agreement, which makes it less attractive relative to the alternative. With respect to (CS-4), the value to the parties of renegotiating more often to undo the shocks increases; hence, they conclude shorter agreements.

(CS-5) As the **risk aversion** of the parties increases, the probability that they will choose finite, renegotiable agreements to adjust for shocks increases.

(CS-6) If the parties conclude a series of renegotiated agreements, as the **risk aversion** of the parties increases, the parties will choose to make each agreement in the series shorter.

In (CS-5), as the parties become more risk averse, they increasingly value some form of flexibility in their agreement to reduce the variation of the realized outcomes. In the case of (CS-6), the intuition is that as the level of risk aversion increases, so do the costs of putting up with an agreement whose distribution of gains has moved far from that originally agreed on. As renegotiation costs are constant, states will choose to renegotiate more often so that the variance in the realized outcomes falls.

Subject to two caveats, the two-period model generalizes readily to an infinite horizon, particularly under the assumption of a time-homogeneous environment. States face the same two choices and again compare their discounted expected utility from each one, choosing the one with the higher value.[9]

The first caveat concerns the alternative of a series of finite duration agreements. In the two-period model, these agreements can only be of length one. In an infinite horizon framework, the states must choose their preferred duration from the set of all possible finite durations by comparing the discounted expected utilities associated with series of finite duration agreements of various lengths. Once they have identified the finite duration agreement length they like best, they can compare it to their discounted expected utility from the other three agreement type choices.

The second caveat concerns the possibility of reneging. Over an infinite time horizon, the distribution of gains under an indefinite agreement could move very far away from the parties' preferred division. If it moved far enough, one party might be losing more

[7] This formulation implicitly assumes neither party reneges in the two-period case. I assume the possible values of the agreement shock are such that, in only two periods, the distribution of gains never moves far enough away from the initial value to make paying the cost of reneging worthwhile. This is a reasonable assumption except for (hypothetical) states on the extreme margin that completely discount future benefits. Still, the complete case with reneging is described in Appendix B.

[8] These comparative statics also confirm the robustness of the logic articulated in Koremenos 2001 when the uncertainty variable is changed.

[9] Generalizing the model for an environment that changes over time would be conceptually simple but notationally burdensome and not add anything to the substantive results.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

from the agreement than the sum of the expected future gains and the costs of reneging. The same condition could hold in the midst of a relatively long finite duration agreement. Thus, reneging looms larger when the time horizon is extended.

The generalization to additional states is straightforward, subject to the following caveats. First, adding additional states will affect certain parameters of the model. As a result, holding the other parameters constant, agreement type choice, as well as the choice of whether to have an agreement at all, may change with an increase in the number of parties. For example, increasing the number of states will increase renegotiation costs. Moreover, generalizing the model both across time and to multiple parties calls attention to the relationship between potential changes in membership and the anticipated costs of renegotiation. If the agreement is such that membership is likely to increase dramatically, states may not choose renegotiation even though the initial membership is low. Given that membership rules themselves affect the number of future signatories to an agreement, this raises the issue of the interaction between two aspects of institutional design: flexibility and membership rules.[10]

EMPIRICAL RESULTS

Data

This article is the first to exploit a new data set, the "Continent of International Law" (COIL).[11] According to Article 102 of the Charter of the United Nations, "every treaty and every international agreement entered into by any Member of the United Nations after the present charter comes into force shall as soon as possible be registered with the Secretariat and published by it" (http://www.un.org/Overview/Charter/contents.html). Given the almost universal membership of the United Nations and its stature among international organizations, its list of international agreements is the most comprehensive to be found.[12] All international agreements registered or filed and recorded with the Secretariat since 1946 are published in the United Nations Treaty Series (UNTS). The Internet collection at the time the sample was drawn contained over 34,000 international agreements "which have been published in hard copy in over 1,450 volumes, which corresponds to all treaties and subsequent actions registered up to December 1986" (http://www.un.org/Depts/Treaty/).

The UNTS Internet site provides subject terms that can be used in searching for international agreements. The primary data for this paper are drawn from the monetary matters, finance, investment, agricultural commodities, environment, human rights, security, and disarmament subject headings. Within each issue area, I generated a random sample of the total set of agreements. Once an agreement was selected, it was subjected to screening criteria. I excluded agreements that only established procedures for other agreements or designated the host for international conferences, that did not include at least two states among the parties, and that did not prescribe, proscribe, or authorize behavior that was observable at least in principle. Agreements are also subjected to rules to avoid double counting.[13]

The characteristics of the selected agreements were recorded using a special coding form. Detailed information on flexibility provisions such as duration and renegotiation provisions, withdrawal and escape clauses, and the presence of quasi-legislative institutions designed to adapt to shocks as well as on a number of other design features like membership provisions and voting rules was coded.

Variables

The dependent variables, whether an agreement is finite or not and if so, how long it is, were measured directly by coding the agreements. Still, it is necessary to operationalize the following independent variables: renegotiation costs, uncertainty, and risk aversion. One reason there has been negligible testing of formal models of international cooperation is the difficulty of finding measures for commonly used variables, like uncertainty and risk attitude. Still, we must rise to the challenge or some of the claims made by critics of the formal approach will be substantiated (Brown 2000). Although I qualify my measures later, they represent earnest attempts to capture the theoretical concepts and thereby serve as useful starting points for much-needed testing in the subfield.

I use the number of original signatories to the agreement as a proxy for *renegotiation costs* because, according to bargaining theory, increasing the number of actors involved is likely to make the negotiation process more lengthy given the existence of multiple equilibria. Also, the amount of information that needs to be analyzed and agreed on before an existing agreement can be adjusted increases with the number of parties involved.

For *uncertainty*, the coding rules are based primarily on categories of agreements in each issue and subissue area. In the model, the uncertainty surrounds the variance of shocks to the distributions of gains from the agreement. Certain kinds of agreements are far more subject to these shocks than are others. For example, agreements whose distribution of gains is

[10] If I were to allow the bargaining power of the parties to be affected by the realized distribution of gains, this would reduce the expected utility of a renegotiated agreement for all values of the variance of the agreement shock. This would not change any of the comparative static results; rather it would only change the cut point at which states switch from one type of agreement to the other.

[11] COIL is supported by a 5-year National Science Foundation CAREER Award: "Designing International Agreements: Theoretical Development, Data Collection, and Empirical Analysis" (SES-0094376).

[12] The key exception to this completeness is the absence of informal agreements.

[13] See http://www.polisci.ucla.edu/faculty/koremenos for the set of agreements in the random sample.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

governed by the forces of supply and demand are coded as high uncertainty, whereas those primarily about coordinating policies to avoid suboptimal outcomes are coded as low uncertainty. Put differently, agreements for which changes in the environment can cause the distribution of gains to vary substantially over time, even while being welfare-enhancing in the aggregate, are high-uncertainty agreements, whereas those for which efficiency concerns dominate and period-to-period changes in the distribution of gains are not expected are low-uncertainty agreements. Two scholars weighed in on each decision, often supplemented by conversations with scholars specializing in the issue area, to avoid subjectivity as far as possible.

Beginning with economics, the subissue area of *monetary* agreements essentially contains exchange rate agreements. These are subject to supply-and-demand shocks that could dramatically alter the distribution of gains from period to period. They are all coded as having high uncertainty. All *trade* agreements are coded as high-uncertainty observations for the same reason. *Finance* agreements are of two types: those that resemble monetary or trade agreements—high uncertainty—and those that are about coordinating policies (e.g., Convention for the Avoidance of Double Taxation and the Prevention of Fiscal Evasion with Respect to Taxes on Income between Australia and Italy)—low uncertainty. Almost one third of finance agreements are high uncertainty. *Investment* agreements concern the promotion and protection of investments against nationalization and expropriation and thereby are subject to political shocks that could alter the distribution of gains; hence, they are high-uncertainty agreements.

In the *environmental* issue area, the following rules were followed: agreements addressing plant and bird protection or scientific cooperation on subjects like soil are coded as low-uncertainty agreements (e.g., the Exchange of Notes Constituting an Agreement on the Project Soil Management and Conservation in East Amazonia between Brazil and the Federal Republic of Germany). Such agreements are predominantly about coordinating policies. On the other hand, agreements about pollution abatement or fishing or other sea resources are coded as high uncertainty (e.g., the multilateral International Convention for the Conservation of Atlantic Tunas). Pollution control and sea resources implicate competitive industries. Shocks affecting the availability of or dependence on the resource (fish) as well as technologic shocks (e.g., in pollution control, expected positive developments may not be forthcoming) can alter the distribution of gains. Almost half of the environmental agreements are coded as high uncertainty.

For *human rights*, universal declarations (e.g., the multilateral Convention on the Prevention and Punishment of the Crime of Genocide) are coded as low uncertainty agreements. Such agreements serve to capture or establish ethically based international norms. On the other hand, agreements about refugees and detailed labor standards are subject to distributional shocks. For instance, political shocks may dramatically change the flow of refugees and thereby change the distribution of gains. Just over half of human rights agreements are coded as high uncertainty.

Security agreements fall into two basic categories: universal prohibitions (e.g., the multilateral Agreement Governing the Activities of States on the Moon and Other Celestial Bodies) and those related to mutual security. Following the logic of human rights, prohibitions are low uncertainty, whereas mutual security agreements are high uncertainty. Almost two thirds of security agreements are high uncertainty.

Risk aversion is a standard assumption in theoretical models of international relations, but little attention has been paid to developing measures of it and incorporating them into analyses of international cooperation. I use three different proxies, each of which captures some aspect of relative risk aversion.

I turn first to Bueno de Mesquita's (1985) risk attitude measure, a "*de facto* academic standard" (Bennett and Stam 2000a, 541) in the international conflict literature. The measure is based on the premise that the closer a state is to the alliance portfolio that maximizes its security, the more risk-averse it is.[14] To generate risk attitude scores, I used the *EUGene* risk attitude variable (Bennett and Stam 2000b).[15] Risk scores are region-based, ranging from −1 (very risk-averse) to +1 (very risk acceptant). For bilateral treaties, I use regional risk scores. For example, if states *i* and *j* sign a treaty, I calculate *i*'s risk attitude toward state *j*'s region and *j*'s risk attitude toward *i*'s region. For multilateral treaties, which include up to 119 signatories, I compute global risk scores (the mean of each signatory's regional risk scores) for each signatory. The risk attitude assigned to an agreement is a function of the risk attitude of the *most* risk-averse participant.[16] Finally, I convert the risk attitude score into a risk aversion measure by inverting the scale to run from −1 (least risk-averse) to +1 (most risk-averse).

Bueno de Mesquita's (1985) risk attitude variable is not above criticism. One problem is that it assumes that alliances work exclusively to make states more secure and therefore reflect risk aversion; it fails to capture the possibility that alliances increase autonomy

[14] More specifically, the variable is constructed in three steps: (1) define state *i*'s "security level" (the sum of all other states' expected utilities vs. *i*); (2) identify the hypothetical alliance portfolio that would minimize and the portfolio that would maximize *i*'s security; (3) how proximate *i*'s **actual** policies are to its **hypothetical** policies may be interpreted as an indication of *i*'s willingness to take risks. Bueno de Mesquita (1985, 157) assumes that "*i*'s risk acceptance increases as *i*'s security score approaches its level of greatest vulnerability, and that *i*'s risk aversion increases as its security approaches the level possessed by its safest policy portfolio.

[15] According to Bennett and Stam (2000a, 466), *EUGene* uses an improved and more accurate algorithm to generate risk attitude scores than does Bueno de Mesquita (1985). I employed *S* rather than Tau-b as a measure of alliance portfolio similarity for the risk attitude scores. Signorino and Ritter (1999) provide good reason to believe that Tau-b can seriously misrepresent the degree to which two states' alliance portfolios are similar; their measure *S* appears to measure better foreign policy similarity.

[16] This "weak-link assumption" is common in quantitative research on the causes of international conflict. See, for example, Dixon 1994 and Oneal and Russett 1997.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

and can therefore reflect risk acceptance.[17] The measure could be improved if data on national issue positions and the status quo on those positions were accessible (Morrow 1987, 436). Unfortunately, because these data are not available, we cannot tell whether the failure to distinguish between the security benefits and the autonomy benefits of alliances introduces error into the measure. Despite this shortcoming, Bueno de Mesquita's risk attitude is far from subjective and appears to be a fairly good measure of the concept I wish to operationalize—that states would be relatively more risk-averse in international cooperation with those states they fear the most from a security standpoint.

To create a second risk attitude variable, I use Gartzke and Jo's "Affinity of Nations Index" (2002), which measures preference similarity among states. Preference similarity does not measure directly the level of risk aversion. Still, it is reasonable to assume that states would be relatively more risk-averse with respect to unanticipated changes in the distribution of gains when dealing with partners with very different preferences; in contrast, if a partner with similar preferences gained more, it can be more safely assumed that the gains will be used for shared priorities. That is, it is likely that state A will be increasingly concerned about the distribution of gains with state B, the more divergent their preferences are.

The Affinity indicator reflects the similarity of state preferences based on their voting positions in the United Nations General Assembly (Gartzke and Jo 2002) and is calculated using Signorino and Ritter's (1999) "S" procedure. The possible values range from −1 (least similar interests) to 1 (most similar interests). For ease of interpretation, I invert this scale so that −1 indicates the least risk-averse state and +1 indicates the most risk-averse state. I use Gartzke and Jo's interpolated SUN2CATI variable, which indicates whether a state voted "yes" or "no." Because the Affinity data are dyadic, I simply take the Affinity value for each bilateral agreement. For the multilateral agreements, I first create a dyad for each pair of signatories. Hence, if there are three signatories, there are three dyads; if there are four signatories, there are six dyads, and so on. I then use the Affinity value of the dyad with the least similar interests, utilizing the "weakest link assumption."

The Affinity data's goal of measuring the similarity of state preferences is far from unproblematic, as Gartzke and Jo (2002, 1–2) acknowledge. The key difficulty is that preferences are not directly observable; hence, we must assume that a state's behavior reveals its preferences. The Affinity measure does have two considerable advantages over Bueno de Mesquita's (1985) risk attitude measure. First, it is based on an information source that is less distorted than alliance portfolios. Alliances are costly; hence, states that do not have strong motives or threats may not pursue one, even though their preference may be to have an alliance (Gartzke and Jo, 2). Second, there is much more variation in UN voting behavior than in alliance formation, particularly during the Cold War.

The third proxy for risk-aversion is a signatory's GDP growth. I argue that states are least risk-averse when they have either very low or very high growth levels and most risk-averse when growth levels lie in a middle range. The relationship between GDP growth and risk aversion is therefore assumed to follow an inverted-U shape. The motivation for the U-shaped function stems from the literature on electoral incentives and risk attitude and, in particular from those who argue that diversionary foreign policy will occur in times of trouble (Downs and Rocke 1995; Levy 1988; Smith 1996, 1998). The argument is that leaders with domestic problems who anticipate being removed often undertake adventurous foreign policies that they would not have attempted otherwise. It is also widely argued that economic variables, like growth rates, are good predictors of election results. Thus, leaders facing either very low or very high growth rates are likely to be relatively more risk-acceptant (they are either gambling for resurrection or extremely secure) than those with middle levels; the latter do not wish to rock the boat.[18]

I used the Penn World Table and calculated each signatory's change in real gross domestic income (adjusting for terms of trade changes) from the year prior to signing. I again used the weakest link assumption, taking the value of the signatory with the lowest growth rate. Next, I separated the growth rate variable into three equal categories, based on the variable's distribution. States with growth levels from the 0 to 33rd percentile of the distribution were considered low-growth states, states with growth levels from the 33rd to the 66th percentile of the distribution were considered mid-growth states, and finally states with growth levels from the 67th to the 100th percentile of the distribution were considered high-growth states. I then created a dummy variable to designate risk-averse states. This variable equals one for mid-growth states and zero otherwise.

Although my three risk indicators differ in numerous ways, an iterated principal component factor analysis reveals moderately strong factor loadings, ranging from about .27 to .44.[19] These factor loadings indicate that

[17] As Morrow (1987, 436) points out, according to Bueno de Mesquita's calculations, Hitler's Germany appears slightly risk-averse. The risk attitude "assigns to German security the autonomy benefits that Germany derived from its alliances with other revisionist powers."

[18] It is important to note that two thirds of the agreements in the sample are characterized by at least one party that has election data available for the years surrounding the signature date and that missing data does not imply the party was not having elections (e.g., Canada has no data for 1946). As a further check on the appropriateness of this measure for the majority of the agreement observations, I examined how many agreements are characterized by at least one democratic signatory (weakest link assumption). Depending on the cutoff used from the Polity data (5, 6, or 7), at least 142 out of 149 agreements have one democratic signatory. This in itself is an interesting observation worthy of future research: what kinds of states do and do not register their agreements with the United Nations?

[19] The Bueno de Mesquita (1985) and Gartzke and Jo (2002) measures are correlated at .121; the Bueno de Mesquita and growth

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

TABLE 1. Intended Duration of Agreements

	Finite or Indefinite? (*p*-Value from Test of Independence: 0.003)	
Issue Area	Percent Finite	Percent Indefinite
Economics	79.7	21.3
Environment	60.0	40.0
Human Rights	44.0	56.0
Security	48.0	52.0
All Agreements	66.4	33.6
	Mean Duration of Finite Agreements (*p*-Value from Test of Equality of Means: 0.107)	
Issue Area	Mean Duration (Standard Deviation of Duration)	
Economics	12.1 (11.6)	
Environment	6.6 (3.5)	
Human Rights	8.4 (3.2)	
Security	5.0 (0.0)	
All Agreements	10.0 (9.4)	

Source: Author's calculations using data on international agreements. The total sample contains 97 finite agreements and 49 indefinite agreements. Within the economics issue area, the sample includes 19 financial agreements, 17 investment agreements, and 9 monetary agreements.

the three indicators, although not substitutes, are significantly related to one another.

Duration Patterns in International Agreements

Perhaps the most striking feature of Table 1 (top panel), which presents the intended duration of my random sample of agreements, is that there is a lot of this kind of flexibility: about two thirds of the agreements have a finite duration. As the bottom panel of Table 1 illustrates, the average duration for a finite agreement is about 10 years.

In the top panel of Table 1, the p-value of 0.000 indicates that whether an agreement is finite or not depends strongly on issue area. For example, agreements in the economics issue area are almost twice as likely as human rights agreements to be of finite duration. That this kind of flexibility varies in important ways across the broad issue areas suggests that states design their agreements to match the broad features of different issue areas. But flexibility also varies in important ways within issue areas, which suggests that states also tailor their agreements to meet their individual preferences and to reflect the unique aspects of particular agreement contexts. The bottom panel of Table 1 displays the mean and standard deviation of the intended durations of the finite duration agreements.[20] Again, we see differences across issue area, although they are not as significant as those in the top panel. Of course, the sample size is also smaller, given that I am conditioning on an agreement being finite.

Subjecting the Model to Empirical Test

Using the absence or presence of a finite (flexible) agreement as the dependent variable, I conduct three probit analyses. In each specification, I include the variable indicating the number of participants (logged) and the uncertainty measure as well as the dummies designating the human rights, economics, and environmental issue areas, as discussed earlier. I test three separate models—one for each measure of risk aversion. Table 2 displays the results of these analyses, and Table 3 shows predicted probabilities.[21]

The statistical analyses provide strong support for the comparative statics. There is robust evidence that as renegotiation costs increase, states become less likely to choose finite agreements **(CS-1)**.[22] Indeed, in all three models, an increase in the number of participants (logged) significantly decreases the probability of a finite agreement ($p < .001$).[23] All else equal, agreements

measures, at −.072; and the Gartzke and Jo and growth measures, at −.144.

[20] I exclude those finite duration agreements whose duration is contingent on something outside of the agreement ($N = 36$). For example, the durations of some agreements are contingent on domestic laws or other international agreements.

[21] Scrutiny of the agreements in my sample suggests that certain states that conclude a large number of agreements, such as the United States, may follow a template when concluding similar agreements with different states. Thus, it could be argued that I should not be treating my agreements as entirely independent observations. To deal with this potential problem, I ran all three models using robust standard errors. The results do not differ significantly from those displayed in Table 2.

[22] I also operationalize the renegotiation cost variable in an additional way. Some agreements contain language requiring official ratification by the governments of the participating states; others do not. If we were to assume ratification adds to negotiation and renegotiation costs, agreements requiring ratification are more likely to be of indefinite duration and less likely to take the form of a series of renegotiated finite duration agreements **(CS-1)**. I find that not having the ratification requirement significantly increases the probability of a finite agreement. Although this test lends support to my hypothesis, the measure may be problematic. If decisions about which kinds of agreements do or do not need ratification are dictated by customary international law, we can treat these requirements as exogenous to any given agreement. But if states make ratification decisions at the same time they choose flexibility provisions, the test described here suffers endogeneity problems. A further complication arises from the fact that the specific requirements of ratification differ both across states and between the international law point of view and the domestic one. Still, the measure does capture some aspect of renegotiation costs with positive results.

[23] It could perhaps be argued that, as the number of signatories increase, renegotiation becomes less important because the probability of enforcement decreases. On the contrary, according to rational design principles (see Koremenos, Lipson, and Snidal 2001: 789–90), states in these circumstances are more likely to delegate enforcement powers to some centralized authority. This conjecture receives some weak empirical support in Koremenos 2005, who finds that, as the number increases, states are more likely to delegate dispute resolution to third parties ($p = 0.15$ with an n of 83).

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

TABLE 2. Results of Probit Analyses of the Presence of a Finite Agreement

Independent Variable	Model 1 Coefficient (Standard Errors)	Model 2 Coefficient (Standard Errors)	Model 3 Coefficient (Standard Errors)
Number of Participants (Logged)	−.469 (.146)***	−.658 (.201)***	−.323 (.142)**
Uncertainty	1.268 (.279)***	1.223 (.294)***	1.266 (.281)***
Risk Aversion (Bueno de Mesquita 1985)	.648 (.275)**	—	—
Risk Aversion (Gartzke and Jo 2002)	—	.373 (.321)	—
Risk Aversion (GDP Growth)	—	—	.662 (.388)**
Human Rights Issue Area	.848 (.517)	.996 (.548)*	.638 (.512)
Environmental Issue Area	.975 (.450)**	.949 (.457)**	.813 (.436)*
Economics Issue Area	1.110 (.381)***	.865 (.378)**	.831 (.367)**
Constant	−.698 (.421)*	−.116 (.513)	−.664 (.422)
χ^2, Wald Test of Joint Significance of Number of Participants, Uncertainty, Risk aversion	12.57***	6.67***	9.08***
Percentage Correctly Predicted	.752	.748	.746
Number of Observations	145	135	138

Source: Author's calculations using data on international agreements. *$p < .10$; **$p < .05$; ***$p < .01$. The Bueno de Mesquita and Gartzke and Jo risk-aversion variables are continuous, and range from −1 (least risk-averse) to +1 (most risk-averse). The GDP growth risk variable is dichotomous and equals 1 for states that have midlevels of growth (and are therefore believed to be most risk-averse). The cutoff point for the calculation of percentage correctly predicted is .664, the sample mean of the dependent variable.

TABLE 3. Predicted Probability of the Presence of a Finite Agreement Based on Changes in Independent Variables of Interest

	Predicted Probability at Minimum Value of Independent Variable	Predicted Probability at Mean Value of Independent Variable	Predicted Probability at Maximum Value of Independent Variable
Number of Participants (Model 1)	.800	.706	.169
Uncertainty (Model 1)	.389	—	.832
Risk Aversion (Bueno de Mesquita 1985; Model 1)	.391	.703	.831
Risk Aversion (Gartzke and Jo 2002; Model 2)	.582	.687	.823
Risk Aversion (GDP Growth; Model 3)	.654	—	.844

Note: Predicted probabilities are based on results displayed in Table 2 and were calculated using *Clarify*. All other independent variables were held constant at their mean values.

with only two participants are between 47.5% and 73.7% more likely to be finite (depending on the operationalization of risk aversion employed) than are agreements with the highest number of participants in the sample (119).[24] The results also provide support for the argument that uncertainty increases states' propensity to choose a finite agreement (**CS-3**). In all three model specifications, the uncertainty effect is in the expected direction and highly statistically significant ($p < .001$). All else equal, agreements with high levels of uncertainty are about 44.0% more likely to be finite than are agreements with low levels of uncertainty.

All three operationalizations provide evidence that risk aversion increases a state's propensity to sign a finite agreement, yielding support for **CS-5**. The Bueno de Mesquita (1985) risk variable has a statistically significant effect at conventional levels ($p < .05$). According to this model, the most risk-averse participants are 44.0% more likely to create finite agreements than are the least risk-averse participants, all else equal. The growth risk variable also has a statistically significant effect at $p < .10$, indicating that risk-averse participants are 19.0% more likely to create finite agreements than their less risk-averse counterparts, all else equal. The sign of the Affinity risk variable is consistent with the theory, and its magnitude suggests important effects. However, perhaps due to the slightly smaller sample size, this operationalization of risk does not attain statistical significance ($p = 0.245$).[25] While none of the risk measures is ideal, each captures some aspect of what the concept means at the level of states operating in anarchy. Moreover, two of the measures are relational, that is, a state's level of risk aversion depends on its particular partner; the other (growth) is an individual attribute of the state that does not vary with partnership. Whereas international relations scholars are fond

[24] I used *Clarify* (Tomz, Wittenberg, and King 2001) to generate all predicted probabilities, varying the independent variable of interest and holding all other independent variables at their mean values.

[25] An argument can also be made that states become less risk-averse as their citizens become richer. This would imply that states with higher GDPs per capita should be more likely to sign an indefinite agreement. In an alternate specification, I used GDP per capita (taken from the Penn Table) as a proxy for risk-aversion. The results provided some evidence that states with higher GDPs per capita are more likely to sign indefinite agreements, although this result was not significant at standard levels ($p = .317$).

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

of relational measures, the other measure I have created is closer to the economist's conception.

As a final test of the appropriateness of the three variables, I perform three Wald tests of the null hypothesis that their *joint* effect equals zero.[26] The Wald tests yield strong support for my arguments: the joint effect of renegotiation costs, uncertainty, and risk is significantly different from zero at $p < .005$ when the Bueno de Mesquita (1985) or growth risk measures are used, and at $p < .01$ when the Affinity risk measure is employed.

Hence, the renegotiation costs, uncertainty, and risk aversion variables are performing as expected. With regard to the other variables in the analyses, the following findings are of note. In all three models, agreements in both the economics and environmental issue areas are significantly more likely to be of finite duration than are security agreements (p is between .004 and .062, depending on the variable and the model specification), but they do not differ significantly from those in the human rights area. Human rights agreements are more likely to be finite than security agreements, but the differences are not uniformly statistically significant at conventional levels across the three model specifications.

In an alternative specification, I added four controls to the models: a superpower variable, equal to 1 if the United States or the Soviet Union is a signatory and 0 otherwise; a democracy variable, equal to the mean Polity value of all signatories; a variable indicating whether the agreement contained an intergovernmental provision as a flexibility device; and the signatories' mean GDP per capita. The superpower, intergovernmental provision, and GDP per capita variables consistently bear no relationship to an agreement's intended duration. The democracy variable has a positive and significant effect ($p < .10$) in Model 3, providing some evidence that democracies tend to choose indefinite agreements.[27] In all specifications, the original results remain unchanged by the inclusion of these new variables. A Wald test of the joint significance of the independent variables of central interest—renegotiation costs, uncertainty, and risk aversion—remains significant ($p < .01$ in two models; $p < .10$ in one model), whereas a Wald test of the joint significance of the added variables never approaches significance.

My model also predicts that if states choose to conclude a series of renegotiated agreements, increases in renegotiation costs will lead them to choose agreements with longer intended durations (**CS-2**); conversely, increases in uncertainty and increases in relative risk aversion will lead states to choose agreements with a shorter duration (**CS-4** and **CS-6**, respectively). To test these predictions, simply performing a regression analysis in which the dependent variable is the intended duration is problematic because the intended duration of some agreements is indefinite. If I limit the sample to those agreements with a finite intended duration, I am selecting on the dependent variable and quite possibly biasing the results. A more appropriate way of testing is to set the intended duration of all indefinite treaties at some value greater than the longest finite treaty. I therefore set the intended duration of all indefinite treaties at four values: 42, 70, 100, and 200 years. Because the data are right-censored, a tobit regression model is used.

I conduct separate tobit analyses using the three risk-aversion proxies. The results provide strong evidence for **CS-2**. At all four values used, the intended duration of finite agreements increases as renegotiation costs increase (p is between .001 and .083, depending on the specification). The results also provide strong evidence in support of **CS-4**. At all four values used, as uncertainty increases, the intended duration of finite agreements decreases notably ($p < .0001$). Finally, at all four values used, the analyses provide some support for **CS-6**, suggesting that as risk aversion increases, the agreement's intended duration decreases. When the Bueno de Mesquita (1985) variable is used, the result is significant at between .037 and .077. When the *Affinity* measure is used, the result is significant at between .095 and .283. The coefficient on the growth risk variable is in the expected direction, but falls short of standard levels of statistical significance (p is between .460 and .806). Table 4 displays the results of the tobit analyses in which indefinite agreements are set at 100. Consistent with the previous analysis, agreements in the security issue area are significantly more likely to be longer than agreements in the other issue areas, suggesting that the security issue area is correlated with some variable(s) not included in my model that also affect agreement length.

Wald tests on Models 4 through 6 provide considerable evidence that renegotiation costs, uncertainty, and risk *jointly* have a significant impact on the intended duration of agreements. The three variables' combined effect is statistically different from zero at $p < .005$ with all three operationalizations of risk aversion. These results add further and significant support for **CS-2**, **CS-4**, and **CS-6**.

OTHER FORMS OF FLEXIBILITY

I have shown the choice of duration is affected by uncertainty, renegotiation costs, and relative risk aversion, but is there anything special about duration? We know that other forms of flexibility exist, and there has even been theoretical work on escape clauses.

[26] This is a test of the joint null that all three variables have zero coefficients, rather than a test of each of them one by one. If two (or more) of a model's theoretical variables are highly correlated, they might not show up statistically significant individually due to multicollinearity, but they would show up as statistically significant jointly. I use a two-tailed test, which is conservative given that my alternative hypotheses are in fact one-sided.

[27] Of course, many would argue that Polity is not a good measure of democracy. Critics smartly point out that not only is there no consensus about what should be included (e.g., there is very little overlap among Polity, Freedom House, and Dahl's dimensions), but also there is no reason to assume that changes in one dimension (chief executive constraints) are equivalent to changes in another (competitiveness of political participation). See Vreeland 2003 and those cited therein for a hard-hitting critique.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

TABLE 4. Results of TOBIT Analyses of the Intended Duration of Finite Agreements

Independent Variable	Model 4 Coefficient (Standard Errors)	Model 5 Coefficient (Standard Errors)	Model 6 Coefficient (Standard Errors)
Number of Participants (logged)	17.218 (6.852)**	32.549 (9.921)***	12.761 (7.255)*
Uncertainty	−67.585 (15.063)***	−68.525 (16.195)***	−72.893 (16.409)***
Risk Aversion (Bueno de Mesquita 1985)	−26.880 (12.814)**	—	—
Risk Aversion (Gartzke and Jo 2002)	—	−25.884 (15.582)*	—
Risk Aversion (GDP growth)	—	—	−12.720 (17.151)
Human Rights Issue Area	−81.032 (29.897)***	−94.319 (31.796)***	−79.713 (31.680)**
Environmental Issue Area	−102.753 (28.014)***	−109.587 (29.350)***	−103.949 (29.546)***
Economics Issue Area	−82.442 (25.875)***	−75.365 (26.728)***	−75.300 (27.498)***
Constant	−175.172 (29.425)***	144.685 (32.464)***	178.186 (31.178)***
χ^2, Wald Test of Joint Significance of Number of Participants, Uncertainty, Risk aversion	14.61***	10.29***	8.29***
Number of Observations	110 (61 Uncensored)	102 (57 Uncensored)	104 (50 Censored)

Source: Author's calculations using data on international agreements. *$p < .10$; **$p < .05$; ***$p < .01$. The Bueno de Mesquita and Gartzke and Jo risk-aversion variables are continuous, and range from −1 (least risk-averse) to +1 (most risk-averse). The GDP growth risk variable is dichotomous and equals 1 for states that have midlevels of growth (and are therefore believed to be most risk-averse). Indefinite agreements are set at an intended duration of 100 years.

TABLE 5. Results of Probit Analyses of the Presence of an Escape Clause

Independent Variable	Model 1 Coefficient (Standard Errors)	Model 2 Coefficient (Standard Errors)	Model 3 Coefficient (Standard Errors)
Number of Participants (Logged)	.342 (.160)**	.470 (.238)*	.353 (.168)**
Uncertainty	.264 (.425)	.261 (.451)	.182 (.442)
Risk Aversion (Bueno de Mesquita 1985 measure)	−.006 (.450)	—	—
Risk Aversion (Gartzke and Jo 2002 measure)	—	−.197 (.452)	—
Risk Aversion (GDP growth measure)	—	—	−.070 (.592)
Security Issue Area	(Omitted Category)	(Omitted Category)	(Omitted Category)
Human Rights Issue Area	.874 (.563)	1.048 (.597)*	.917 (.578)
Environmental Issue Area	.046 (.709)	.157 (.753)	.074 (.723)
Economics Issue Area	.266 (.599)	.471 (.666)	.095 (.634)
Constant	−2.578 (.680)***	−2.923 (.903)***	−2.542 (.710)***
χ^2, Wald Test of Joint Significance of Number of Participants, Uncertainty, Risk aversion	.90	.94	.34
Number of Observations	145	135	138

Source: Author's calculations using data on international agreements. *$p < .10$; **$p < .05$; ***$p < .01$. The Bueno de Mesquita and Gartzke and Jo risk aversion variables are continuous, and range from −1 (least risk-averse) to +1 (most risk-averse). The GDP growth risk variable is dichotomous, and equals 1 for states that have midlevels of growth (and are therefore believed to be most risk-averse).

To test whether other forms of flexibility can substitute for duration, I perform three probit analyses using the presence of an escape clause as the dependent variable and three using the presence of a withdrawal clause as the dependent variable. The independent variables are identical to those displayed in Table 2. As Tables 5 and 6 demonstrate, most of the independent variables do a very poor job of predicting the outcomes of interest. Uncertainty and risk aversion bear no significant relationship to the probability that an agreement will contain an escape or a withdrawal clause. The effect of the number of participants in each model is not in the expected direction, suggesting that agreements with more participants are more, not less, likely to include escape and withdrawal clauses. Finally, Wald tests that the joint effect of the key variables—number of participants, uncertainty, and risk aversion—equals zero yield insignificant results across every model.[28] I also analyze a dependent variable coded 1 if an agreement possesses at least one of the aforementioned forms of flexibility (finite duration, escape clause, or withdrawal provision) and 0 otherwise. Not one of the independent variables has a statistically significant impact on flexibility in any of the probits. Together, these results suggest that the determinants of escape and withdrawal provisions are different than those of duration; hence, these other forms of flexibility are solving different problems.

[28] Interestingly, compared to security agreements, human rights agreements are significantly more likely to contain withdrawal clauses across all three models and escape clauses across one model.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

TABLE 6. Results of Probit Analyses of the Presence of a Withdrawal Provision

Independent Variable	Model 1 Coefficient (Standard Errors)	Model 2 Coefficient (Standard Errors)	Model 3 Coefficient (Standard Errors)
Number of Participants (logged)	.286 (.129)**	.195 (.163)	.253 (.132)*
Uncertainty	.347 (.258)	.271 (.261)	.274 (.260)
Risk Aversion (Bueno de Mesquita 1985 measure)	−.317 (.227)	—	—
Risk Aversion (Gartzke and Jo 2002 measure)	—	.091 (.281)	—
Risk Aversion (GDP growth measure)	—	—	.018 (.286)
Security Issue Area	(Omitted Category)	(Omitted Category)	(Omitted Category)
Human Rights Issue Area	1.090 (.439)**	.993 (.445)**	1.066 (.453)**
Environmental Issue Area	.617 (.377)	.545 (.379)	.653 (.382)*
Economics Issue Area	.466 (.314)	.418 (.323)	.536 (.318)*
Constant	−.706 (.398)*	−.559 (.472)	−.781 (.404)*
χ^2, Wald Test of Joint Significance of Number of Participants, Uncertainty, Risk Aversion	.74	2.42	1.54
Number of Observations	145	135	138

Source: Author's calculations using data on international agreements. $^{*}p < .10$; $^{**}p < .05$; $^{***}p < .01$. The Bueno de Mesquita and Gartzke and Jo risk aversion variables are continuous, and range from −1 (least risk-averse) to +1 (most risk-averse). The GDP growth risk variable is dichotomous and equals 1 for states that have midlevels of growth (and are therefore believed to be most risk-averse).

What is it that makes a finite duration with the possibility of renegotiation uniquely important in this context? This particular institutional design allows *adjustment* in the face of international uncertainty *without dismantling cooperation*. Escape clauses do not allow adjustment; rather, they allow states to temporarily escape cooperation and return to an unadjusted agreement. Escape clauses are, however, appropriate responses to *domestic* uncertainty. States may agree to particular terms of cooperation but then suffer domestic shocks that make these terms politically difficult. What they require is a temporary relief from their obligations. Even the typical wording of escape clauses suggests this purpose: "extraordinary circumstances that jeopardize extreme national interests." Human rights agreements contain significantly more escape clauses than agreements in other issue areas, with the domestic shock usually being civil war.[29] Article 4 of the "International Covenant on Civil and Political Rights" states that "in time of public emergency which threatens the life of the nation . . . [states] may take measures derogating from their obligations under the [agreement] to the extent strictly required by the exigencies of the situation." If a state takes such measures, it must inform other state parties through the Secretary-General of the UN regarding "the provisions from which it has derogated and of the reasons by which it was actuated."

Withdrawal clauses are also very different from duration provisions because cooperative institutions cease to exist in the bilateral cases (by far, the majority) or the membership changes in a multilateral setting. The latter can be consequential: when North Korea withdrew from the Nuclear Nonproliferation Treaty, the agreement remained intact, yet the implications of the membership change were serious. Withdrawal clauses are responses to shocks that alter a state's basic interest in cooperation. Although such shocks rarely occur, the risk they impose is great. Thus withdrawal clauses are pervasive, but their use is infrequent and therefore dramatic.[30]

It is also illuminating that 62% of the agreements have withdrawal provisions and about 8% have escape clauses, but the correlation between these variables and duration never exceeds .18. Flexibility provisions are not simply chosen as a set; nor do particular pairs go together. The problems these provisions uniquely solve occur in different combinations depending on the cooperative endeavor. The conclusion to be drawn is that the landscape of international agreements is far from crude.

Although escape and withdrawal clauses seem to solve different problems than finite duration, another design tool allows adjustment in the face of shocks: an agreement embodying a quasi-legislative institution with the power to modify the distribution of gains. I have argued that such a design may be optimal when uncertainty regarding future gains is pervasive, but renegotiation costs are high because of the number of parties involved. In this context, an institution with an amendment provision characterized by majority rule cuts down on adjustment costs relative to a full renegotiation (Koremenos 2000).[31] The International Monetary Fund provides a good example of an indefinite

[29] This fact poses a challenge to those who believe rationalist logic is inappropriate in this issue area.

[30] The relationship between "bedrock" preferences, which are fundamentally stable, and constraints, which arise from the fact that the state is a composite actor, also provides insight (Koremenos, Lipson, and Snidal 2001, 1073). I would argue that withdrawal clauses are used in the event of "bedrock" changes, whereas escape clauses are used in the event of unchanged bedrock preferences but different domestic constraints.

[31] Not all amendment provisions are substitutes for finite duration. Often amendments are binding only on those accepting them. Environmental agreements have the most amendment provisions at 33%.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

duration agreement that establishes an institution that does many things including, importantly, adjusting the distribution of gains. As I mentioned earlier, this option is left out of the full analysis because of current data limitations, but future work will address this question of substitutability.[32]

OTHER THEORIES

Because this analysis fits squarely within the rationalist institutionalist paradigm in international relations, I now consider what the two competing theoretical approaches, constructivism and realism, would predict, although neither has generated testable implications regarding international institutional design. With respect to constructivists, Finnemore's (1996) work draws on sociological institutionalism, which documents isomorphic outcomes across a variety of substantive areas in different parts of the world. Similar outcomes are driven by a common global culture. If we were to take this argument to the level of agreement design, we would expect emulation regarding agreement form. Additionally, Finnemore and Barnett (1999) argue that agency culture drives institutional form, so the World Bank looks like it does because economists run it. Of course, most agreements are bilateral and establish no agency. Still, it could be argued that international lawyers have a culture; because they write the agreements, these lawyers are likely to follow a template. The design implication of using templates is observationally equivalent to the one predicted by isomorphism: emulation.

To capture the idea of emulation manifested through either more or less frequent use of finite duration provisions over time, I ran a test with time as the only independent variable. The test yielded no significant results. Additionally, the descriptive statistics presented previously also suggest that world culture does not produce isomorphism with respect to flexibility provisions; rather, what emerges is striking variation.

Another line of constructivist scholars emphasizes the role of norms and activist networks in shaping international outcomes. However, the primary outcomes being explained are whether states form cooperative institutions or not and, if they do, whose values are reflected in those institutions. These scholars are not explaining agreement design. In fact, Keck and Sikkink (1998) argue that even normative actors need to be strategic and therefore should pay attention to the details of institutional design. Hence, it could be argued that my analysis complements their analysis because we are predicting outcomes at very different levels.

Realists, who focus on power considerations, do not have a theory of institutions; they have not taken them seriously. (Glaser [1994] is one of a few realists writing about cooperation, but he does not believe that institutions, let alone design, matter.) Hence, variation in flexibility provisions would not surprise them. Importantly, however, they would expect no systematic variation given their argument that international agreements and institutions are epiphenomenal; rather, any variation would be noise. The data clearly refute this prediction. (And, as discussed previously, the superpower variable is insignificant.)[33]

CONCLUSIONS

The effects of uncertainty are dramatic in international relations given the fundamental insecurity posed by anarchy. How do states transcend this uncertainty and cooperate? The answer is straightforward: they negotiate agreements that include the proper amount of flexibility and thereby create for themselves a kind of international insurance.

Challenging a long and widely held view that anarchy makes international relations qualitatively distinct from other fields, I develop a model of duration provisions based on economic contracting theory that highlights renegotiation costs, uncertainty regarding the distribution of gains, and relative risk aversion. I then subject this model to a test featuring a random sample of international agreements.[34]

The descriptive statistics alone force us to reconsider one of the conventional wisdoms concerning international cooperation: states tie their hands in order to make their commitments credible. In fact, two thirds of the agreements in this sample are designed *not* to last forever. Does this statistic imply that states are not making many credible commitments? On the contrary, my model suggests that flexibility may *enhance* credibility in environments subject to shocks; reneging, which can be quite damaging, can often be avoided if renegotiation is an option instead.

It could be argued that because some of the agreements in the random sample are minor, the conclusions I draw may be meaningless. Yet it is doubtful these agreements are minor to the governments that signed them, and these sorts of agreements govern most of

[32] The degree of precision/ambiguity in an agreement's obligations is another form of flexibility that might be a substitute for duration, although it would not be a good response to the persistent uncertainty modeled in this paper. Rather, imprecision could be a response to a one-time initial uncertainty, like that modeled in Koremenos 2001. The current broader recoding of each of the agreements in the sample tries to capture such a variable with four categories of precision/ambiguity. In future work, the kinds of analyses performed with escape clauses and withdrawal clauses will be possible with the precision variable. See Abbott and Snidal 2000, for a discussion of precision and of soft law more generally.

[33] This discussion reflects a first (yet serious) attempt at teasing out some testable implications from these other paradigms. I welcome others with more expertise to elaborate alternative testable hypotheses.

[34] Duffield (2003, 426) criticizes certain rationalist approaches to international cooperation for their reliance on case studies. Quite rightly he argues that nothing can be "proved" with a case study. This empirical analysis therefore meets the Duffield challenge. Nonetheless, case studies are also essential. They reveal the historical importance of our theoretical propositions, they help us understand the mechanisms by which our causal variables operate, and they provide us with much needed guidance regarding how to operationalize our variables.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

the day-to-day cooperation between states. So if cooperation is to be explained, such agreements deserve attention as a category.[35] Certainly, future work must address what happens once the agreements are designed. Although a number of case studies illustrate that duration and renegotiation provisions are meaningful with respect to how cooperation evolves, an important next step would be to randomly select 15 to 20 of the agreements used in this analysis and research and evaluate their effectiveness.

The sharpest potential critique, however, is that it makes no sense to focus analysis on agreement design when the value of international agreements is itself contested. Still, one way of shedding light on this particular international relations debate is to examine whether the design of agreements is consistent with the presumption that the agreements themselves are important. In fact, instead of random variations among agreements or automatic replication of the same agreement provisions over and over, the detailed provisions of international agreements are chosen in ways to increase the incidence and robustness of cooperation. States would not spend so much effort getting institutions "right" if the institutions didn't matter.

International agreements are nuanced and sophisticated—just like the domestic agreements and institutions studied in American and comparative politics and in law and economics.[36] Although anarchy makes uncertainty dramatic, states contract around this uncertainty with various forms and combinations of flexibility, like duration, escape, and withdrawal provisions. Each is a unique response to a distinct kind of uncertainty. Nor is flexibility the only agreement feature that proves systematic. Additional analyses (e.g., Koremenos N.d.) suggest that provisions regarding monitoring and dispute resolution also are chosen according to principles of rational design.

International agreements are important because they regulate cooperation, and the fact that they obey lawlike regularities indicates that serious efforts are made for them to be able to regulate interactions in lasting and successful ways. So, because agreements matter, they are designed in rational ways, and the fact that people make efforts to design them in such ways corroborates their significance.

[35] Of course, our intuitions about the forms of international cooperation are often shaped by a few high-profile agreements. High-profile agreements get that way because they have big effects; thus, it is important to be able to make statements about their characteristics. Future work will build a sample of high-profile agreements, and we will be able to make a comparison between the two samples. Only then can we say with any confidence whether past reliance on case studies was somewhat justified.

[36] International relations is not the only field implicated in this analysis. A critical debate in law and economics centers on the importance of the "shadow of the law." Given that the design of international cooperation obeys lawlike regularities without any compelling supranational authority, perhaps the "law" is not quite as important as many thought *within* the state. Not only can international relations scholars learn from law scholars, but also we have important results to share with them.

APPENDIX A

Assuming that State A and B are symmetric, compare State A's expected utility with one two-period (inflexible) agreement, given by $EU_2 = E(u(b_1 + m_0 + e_1 - k_n) + \delta u(b_1 + m_0 + e_1 + e_2))$, to its expected utility from two 1-period (flexible) agreements, given by $EU_1 = E(u(b_1 + m_0 + e_1 - k_n) + \delta u(b_1 + m_0 + e_2 - k_r))$. The only differences between EU_1 and EU_2 occur in the second period. In that period, with EU_2 the states again live with the first period shock e_1 but do not pay renegotiation cost k_r. The reverse is true for EU_1, where the states avoid having to experience e_1 by paying the renegotiation cost k_r.

With respect to **CS1**, the key point is that EU_1 is decreasing in the renegotiation cost k_r whereas EU_2 is unaffected by k_r. Thus, suppose that initially $EU_2 > EU_1$; that is, an inflexible agreement dominates. Holding the distribution of e_1 fixed, we can decrease k_r until at some point a threshold is reached and $EU_1 > EU_2$. States will then change from an inflexible two-period agreement to a flexible series of one-period agreements. A similar story holds in the other direction, where initially two 1-period agreements are preferred. Increasing k_r will decrease EU_1 until at some point the two expected utilities are equal, and after which point the states jointly prefer an inflexible two-period agreement because the increased cost of renegotiation makes the one-period agreements no longer optimal.

With respect to **CS3**, holding the form of the utility function constant and assuming that it is strictly increasing and concave, a mean-preserving increase in the variance of the shocks to the distribution of gains (i.e., an increase in uncertainty), e, reduces both EU_1 and EU_2, because e_1 and e_2 enter both and because concavity of the utility function implies risk aversion (i.e., states prefer a lower variance to the shocks at a given mean, in this case zero). But because the shocks are independent, in the second period, the combined shock $(e_1 + e_2)$ experienced under the two-period agreement has a variance equal to $2\,\mathrm{var}(e)$, whereas the second period shock in the case of two 1-period agreements, just e_2, has only $\mathrm{var}(e)$. Thus, a given increase in $\mathrm{var}(e)$, holding the mean constant, increases the variance of the combined shock more for the two-period agreement, and thereby leads to a larger decrease in EU_2 than in EU_1. That is, increasing the variance of the shocks decreases EU_2 relative to EU_1. Thus, if we start from a situation where an inflexible two-period agreement is preferred by the states and then continually increase the $\mathrm{var}(e)$, holding constant k_r and the mean of the shocks, at some point EU_2 will fall sufficiently relative to EU_1 that the states will prefer a flexible regime of two 1-period agreements. At this point, it is worth it to the states to pay k_r in order to reduce the variance of the second-period outcomes. The same argument holds in the other direction. If initially the states prefer two 1-period agreements, we can decrease $\mathrm{var}(e)$, which will now increase EU_2 relative to EU_1, until $EU_2 > EU_1$. In the limiting case, where $\mathrm{var}(e) = 0$, states will prefer a two-period agreement as long as $k_r > 0$.

A similar argument holds with respect to **CS5**. Here we hold the variance of the shocks constant while varying the concavity of the utility function. The key point is that the variance of the outcomes is higher under the inflexible two-period agreement, given that states experience the combined shock $(e_1 + e_2)$ in the second period. As a result, increasing the concavity of the utility function (i.e., increasing the degree of risk aversion) reduces both expected utilities, EU_1 and EU_2, but decreases EU_1 relatively more. Suppose initially the two-period agreement is preferred by both states. If we increase the concavity of the utility function, say by increasing the degree of the exponent, a, when $u(x) = x^a$ as in

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

TABLE B1. Results from Simulation of Two-Period Model

Utility Function	δ	var(ε)	k_n	k_r	c	$E(U_R)$	$E(U_N)$	Choice
Base Case								
Square Root	0.9	Low	1.0	0.5	1.0	12.4927	12.5130	One 2-Period (Inflexible)
Decrease Renegotiation Costs								
Square Root	0.9	Low	1.0	0.1	1.0	12.5199	12.5185	Two 1-Period (Flexible)
Increase Variance of Shocks (Uncertainty)								
Square Root	0.9	High	1.0	0.5	1.0	12.4900	12.4557	Two 1-Period (Flexible)
Increase Risk Aversion								
Cubic	0.9	Low	1.0	0.5	1.0	6.6682	6.6658	Two 1-Period (Flexible)

the simulations, or by increasing the risk aversion parameter a in a constant absolute risk aversion (CARA) utility function where $u(x) = -\alpha e^{-ax} + \beta$ (Kreps 1990: 86), EU_2 declines relative to EU_1. At some point, $EU_1 > EU_2$, and both states now prefer to enter into two 1-period agreements. As previously, the same arguments work in reverse if we start from a situation where the states choose two 1-period agreements, and then decrease the level of concavity (and therefore risk aversion) in the utility function. Eventually the states will become risk-tolerant enough that they prefer an inflexible two-period agreement to two 1-period agreements, so long as $k_r > 0$. In the limiting case of a linear utility function (risk neutrality), the states will always prefer the longer agreement if $k_r > 0$.

APPENDIX B

This appendix describes the construction of the solutions to the discretized version of the two-period model used for the simulations. The value of base outcome, a, is set to 40.0. This value is completely arbitrary as long as it is positive; its main function in the simulations is to be large enough to keep the outcome from going negative even after a series of bad agreement shocks. The value of each state's share of the initial distribution of gains is set to $m = 4$, which implies a total gain from the agreement, given equal *ex ante* division due to Nash bargaining, of $g = 2m = 8$. The general class of utility functions considered in the simulation is polynomial. In particular, I consider $U(X) = X$, $U(X) = X^2$, $U(X) = X^3$, $U(X) = X^4$ or, in words, a linear utility function, a quadratic utility function a cubic utility function, and a quartic utility function. For the values of X considered here, which are all positive and greater than 1, an increase in the exponent of the utility function represents an increase in the level of risk aversion.

To keep the programming simple, the distribution of shocks is discrete, not continuous. The shock in the first period is denoted ε_1, and the shock in the second period is denoted ε_2. The potential values of the shocks to the distribution of gains under the agreement are denoted by ε_{1j} and ε_{2j}, where $j \in \{1, 2, 3\}$ denotes one of the three possible values of the shocks. In the simulations in Table B1, the values of the shocks are -2, 0 and 2 so that $\varepsilon_{11} = -2$, and so on. The variance of the distribution of shocks is increased/decreased through a mean-preserving increase/decrease in the spread of the distribution; in this case, there is a symmetric increase in the probability of the two extreme values, with a corresponding reduction in the probability of a zero shock.

I examined values for the costs of negotiation, k_n, and renegotiation, k_r, in the range $[0, 1]$. I considered values of the discount rate, δ, in the set $\{0.6, 0.7, 0.8, 0.9\}$, whereas the values of the costs of reneging, c, ranged over the interval $[0, 5]$. For values of c near the high end of this range (how high depends on the shape of the utility function), reneging does not occur given the values of the shocks utilized here.

For each of the variables corresponding to the comparative static results, I examined essentially the entire range of possible values (or distributions) within this framework; all were (not surprisingly given the proofs in Appendix A), consistent with the model. The particular values presented in Table B1 were chosen to represent the wide range of simulation results actually generated.

The expected utility associated with two 1-period agreements is given by

$$E(U_R) = \sum_{j=1}^{3} \Pr(\varepsilon_1 = \varepsilon_{1j}) U(a + m + \varepsilon_{1j}) + \delta \sum_{j=1}^{3} \Pr(\varepsilon_2 = \varepsilon_{2j}) U(a + m + \varepsilon_{2j} - k_r),$$

where $U(\cdot)$ is the utility function.

The equation for a single two-period agreement is more complicated due to the possibility of reneging. Consider the problem from the standpoint of state A. We can define an indicator variable, R_A, for whether or not State A will renege as follows:

$$R_A(\varepsilon_{1j}) = 1[E(U(a + m + \varepsilon_2 - k_r - c)) > E(U(a + m + \varepsilon_{1j} + \varepsilon_2))],$$

where the expectations are over the possible values of ε_2. The expectations are as of the end of the first period, so that ε_1 is known but ε_2 is not. In other words, State A compares the expected utility of reneging, given by the first term, with the expected utility of not reneging, given by the second term. The gain from reneging is getting rid of ε_1; the cost is the combined cost of reneging plus renegotiation, given by $(k_r + c)$. State B makes a similar calculation; we can define an indicator variable R_B for its decision as follows:

$$R_B(\varepsilon_{1j}) = 1[E(U(a + m - \varepsilon_2 - k_r - c)) > E(U(a + m - \varepsilon_{1j} - \varepsilon_2))].$$

With these indicator variables in hand, we can now give the formula for State A's expected utility from a two-period agreement. It is

$$E(U_N) = \sum_{j=1}^{3} \Pr(\varepsilon_1 = \varepsilon_{1j}) \{U(a + m + \varepsilon_{1j} - k_n) + \delta[R_A(\varepsilon_{1j}) T_1 + R_B(\varepsilon_{1j}) T_2 + (1 - R_A(\varepsilon_{1j}) - R_B(\varepsilon_{1j})) T_3]\},$$

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press

where

$$T_1 = \sum_{j=1}^{3} \Pr(\varepsilon_2 = \varepsilon_{2j}) U(a + m + \varepsilon_{2j} - c - k_r),$$

$$T_2 = \sum_{j=1}^{3} \Pr(\varepsilon_2 = \varepsilon_{2j}) U(a + m + \varepsilon_{2j} - k_r),$$

and

$$T_3 = \sum_{j=1}^{3} \Pr(\varepsilon_2 = \varepsilon_{2j}) U(a + m + \varepsilon_{1j} + \varepsilon_{2j}).$$

The formulae for State B are symmetric.

REFERENCES

Abbott, Kenneth W., and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54 (Summer): 421–56.

Barnett, Michael N., and Marthega Finnemore. 1999. "The Politics, Power, and Pathologies of International Organizations." *International Organization* 53 (Autumn): 699–723.

Bennett, D. Scott, and Allan Stam. 2000a. "*EUGene*: A Conceptual Manual." *International Interactions* 26, 179–204.

Bennett, D. Scott, and Allan C. Stam. 2000b. "A Universal Test of an Expected Utility Theory of War." *International Studies Quarterly* 44 (September): 451–80.

Bordo, Michael, and Finn Kydland. 1995. "The Gold Standard as a Rule: An Essay in Exploration." *Explorations in Economic History* 32 (October): 423–64.

Brown, Michael E., ed. 2000. *Rational Choice and Security Studies: Stephen Walt and His Critics.* Cambridge: MIT Press.

Bueno de Mesquita, Bruce. 1985. "The War Trap Revisited: A Revised Expected Utility Model." *American Political Science Review* 79 (March): 156–77.

Conybeare, John. 1985. "Trade Wars: A Comparative Study of Anglo-Hanse, Franco-Italian, and Hawley-Smoot Conflicts." *World Politics* 38 (October): 147–72.

Dixon, William J. 1994. "Democracy and the Peaceful Settlement of International Conflict." *The American Political Science Review* 88 (March): 14–32.

Downs, George, and David Rocke. 1995. *Optimal Imperfection?: Domestic Uncertainty and Institutions in International Relations.* Princeton, NJ: Princeton University Press.

Duffield, John. 2003. "The Limits of 'Rational Design."' *International Organization* 57 (Spring): 411–30.

Dye, Ronald A. 1985. "Optimal Length of Labor Contracts." *International Economic Review* 26 (February): 233–50.

Finnemore, Martha. 1996. *National Interests in International Society.* Ithaca, NY: Cornell University Press.

Gartzke, Erik, and Dong-Joon Jo. 2002. "The Affinity of Nations Index, 1946–1996." Version 3.0. http://www.columbia.edu/~eg589/datasets.html.), November 8, 2002.

Glaser, Charles. 1994. "Realists as Optimists: Cooperation as Self Help." *International Security* 19 (Winter): 50–90.

Gray, Jo Anna. 1978. "On Indexation and Contract Length." *Journal of Political Economy* 86 (February): 1–18.

Guzman, Andrew, and Beth A. Simmons. 2002. "Litigation and Settlement." *Journal of Legal Studies* 31: S205–S235.

Harris, Milton, and Bengt Holmstrom. 1987. "On the Duration of Agreements." *International Economic Review* 26 (June): 389–406.

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca, NY: Cornell University Press.

Koremenos, Barbara. 2000. "Bending but Not Breaking: Flexibility in International Financial and Monetary Agreements." Working Paper 1.73. University of California, Berkeley: Center for German and European Studies.

Koremenos, Barbara. 2001. "Loosening the Ties that Bind: A Learning Model of Agreement Flexibility." *International Organization* 55 (Spring): 289–325.

Koremenos, Barbara. 2005. "Bringing More 'Precision' to the Three Dimensions of Legalization." Paper presented at the European University Institute, Fiesole, Italy.

Koremenos, Barbara. N.d. "If Only Half of International Agreements Have Dispute Resolution Provisions, Which Half Needs Explaining?" *Journal of Legal Studies*. Forthcoming.

Koremenos, Barbara. http://www.polisci.ucla.edu/faculty/koremenos.

Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. "Rational Design: Looking Back to Move Forward." *International Organization* 55 (Autumn): 1051–82.

Kreps, David M. 1990. *A Course in Microeconomic Theory.* Princeton, NJ: Princeton University Press.

Levy, Jack S. 1988. "When do Deterrent Threats Work?" *British Journal of Political Science* 18 (October): 485–512.

Martin, Lisa L. 1992. *Coercive Cooperation: Explaining Multilateral Economic Sanctions.* Princeton, NJ: Princeton University Press.

Martin, Lisa L. 2000. *Democratic Commitments: Legislatures and International Cooperation.* Princeton, NJ: Princeton University Press.

Morrow, James D. 1987. "On the Theoretical Basis of a Measure of National Risk Attitudes." *International Studies Quarterly* 31 (December): 423–38.

Oneal, John R., and Bruce M. Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950–1985." *International Studies Quarterly* 41 (June): 267–93.

O'Neill, Barry. 2001. "Risk Aversion in International Relations Theory." *International Studies Quarterly* 45 (December): 617–40.

Osborne, Martin J., and Ariel Rubinstein. 1990. *Bargaining and Markets.* Boston: Academic Press.

Rosendorff, Peter, and Helen Milner. 2001. "The Optimal Design of International Trade Institutions: The Sources and Impact of Escape Clauses." In *The Rational Design of International Institutions*, ed. Barbara Koremenos, Charles Lipson, and Duncan Snidal. *International Organization*. Special Issue 55 (Autumn): 829–57.

Signorino, Curtis, and Jeffrey Ritter. 1999. "Tau-b or Not Tau-b: Measuring the Similarity of Foreign Policy Positions." *International Studies Quarterly* 43 (March): 115–44.

Simmons, Beth A. 2000. "International Law and State Behavior. Compliance with the Public International Law of Money." *American Political Science Review*, 94 (4): 819–35.

Smith, Alastair. 1996. "Diversionary Foreign Policy in Democratic Systems." *International Studies Quarterly* 40 (March): 133–53.

Smith, Alastair. 1998. "International Crises and Domestic Politics."*American Political Science Review* 92 (September): 623–38.

Tomz, Michael, Jason Wittenberg, and Gary King. 2001. "Clarify: Software for Interpreting and Presenting Statistical Results." http://gking.harvard.edu/clarify (April 23, 2003).

Vreeland, James Raymond. 2003. "A Continuous Schumpeterian Conception of Democracy." Working Paper, Yale University, New Haven, CT.

https://doi.org/10.1017/S0003055405051877 Published online by Cambridge University Press