Negotiating Strategic Arms Control, 1969-1979: Modeling the Bargaining Process

Richard J. Stoll
William McAndrew

Abstract: Using data collected by Jensen (1984), we investigate whether cooperative and inverse reciprocity can successfully predict U.S. and Soviet actions during the 23 rounds of strategic arms negotiation during the 1969-1979 period. Each type of reciprocity gives rise to three models of bargaining behavior, which we implement as computer routines. In our analysis, we find that the cooperative reciprocity models fit more often than inverse reciprocity models

Beginning in 1969, the United States engaged in an almost continuous series of strategic arms control talks with the Soviet Union, climaxing with the signing of the SALT II Treaty in June 1979. The record of these negotiations gives us a valuable field laboratory for testing a variety of models of the negotiation process. This is an especially interesting series of interactions to study, because it involves the two most powerful nations in the world on issues of extreme importance.

Thanks to the recent work of Jensen (1984), we have available a coding of all major concessions and retractions made by the two sides during these negotiations. These data will be used to test two sets of assumptions about the bargaining strategies of the United States and the Soviet Union during the strategic arms control talks. One set of assumptions is built around the notion of cooperative reciprocity, and the other set is built around the notion of inverse reciprocity (Bartos, 1974). First, we outline both bargaining frameworks, and then we formalize each into a set of three different models. After reviewing Jensen's data-making procedures, we test all six models to determine whether the United States and the Soviet Union pursued and reacted to "soft-line" or "hard-line" bargaining strategies during their strategic arms negotiations.

Bargaining: Frameworks and Models

The study of bargaining spans several disciplines and encompasses many conceptions of the negotiation process; we will deal with only two in this article. First, we consider the main features of these two frameworks and then we provide three different models for each framework. Each model is a specific representation of a framework that can be tested against Jensen's data.

Cooperative Reciprocity

Jensen (1984: 535-536) captures the essence of the cooperative-reciprocity framework and its centrality to the bargaining literature as follows:

Perhaps no proposition relating to bargaining behavior has been better documented in both experimental studies and real life situations than the one suggesting that concessions tend to be reciprocated.

Often only implied but not stated is the other side of the reciprocity coin: Retractions tend to be reciprocated as well

One well-known example of such a reciprocating strategy is "tit for tat" in an iterated Prisoner's Dilemma game. In this game, each player has two strategies: cooperate or defect. Using a tit-for-tat strategy, the player begins by choosing the cooperative strategy on the first round. In subsequent rounds, the player selects the strategy played by the opponent in the previous round. Recently, Axelrod (1984) studied a wide variety of situations resembling iterated Prisoner's Dilemma games. He showed that actors pursuing their own self-interest without the presence of a central authority to police their actions will very often adopt a tit-for-tat reciprocating strategy. The reason for this is that the strategy is very robust and will produce a good outcome under a large number of circumstances.

Axelrod's study, and the other works alluded to in Jensen (1984), would lead us to believe that cooperative reciprocation will be widely used in a variety of negotiations that result in an agreement, including the strategic arms talks between the superpowers.

Before considering inverse reciprocity, we note two features of cooperative reciprocity, as embodied in the tit-for-tat strategy, that Axelrod feels are crucial to its success: It is "nice" and it is "forgiving." It is nice because an actor using tit for tat will never be the first to select the defection strategy. It is forgiving because in the face of a defection by the opponent, an actor using tit for tat punishes (by choosing the defecting strategy) only for a single round. Taken together, these two characteristics imply that a cooperative reciprocating strategy is a relatively "soft-line" bargaining strategy. An actor who plays this strategy does not seek to batter opponents into submission, but rather to lead them to a successful resolution by encouraging cooperative responses, while still punishing defecting responses.

Inverse Reciprocity

If cooperative reciprocation is nice, then inverse reciprocation is unpleasant. Instead of leading with the carrot and using the stick only if the opponent is recalcitrant, inverse reciprocation is exploitative (Bartos, 1974: 38). Concessions by the opponent are met with retractions, and retractions by the opponent are met with concessions.

The rationale behind the approach is twofold. First, a concession by the opponent is interpreted as weakness, so a tough stand will extract further concessions. On the other hand, if the opponent stands firm, this indicates that she or he has been pushed too far, and a concession is the appropriate response. The second part of the rationale simply turns this reasoning around to consider the impact of the actor's concessions on the opponent: An opponent will view a concession as a sign of weakness and act to exploit it; only a firm stand will convince an opponent to make concessions.

There is some evidence that a hard-line stand by the United States produced Soviet concessions in the SALT talks, as well as in earlier negotiations on the Threshold Test Ban Treaty (Jensen, 1984: 540-541); note, however, that the relationship found during SALT was much weaker than during the Test Ban negotiations.

Summarizing the Differences

It should be stressed that the cooperative and inverse reciprocity frameworks are not just different, they are in fact mirror images of one another. They predict opposite responses to actions of the opponent. Thus, it is possible that a single empirical test can serve to evaluate both frameworks. We will take advantage of this feature in the construction and evaluation of the models, a matter to which we now turn.¹

Models of Reciprocity

Although the differences between the frameworks are clear enough, converting each into an empirically testable model is no simple task. In fact, we have found it necessary to consider not a single dual model (i.e., a model that can simultaneously test both frameworks), but rather three such models.

The problem we faced in attempting to specify a precise description of the process of reciprocity was that a number of different possibilities appear to be reasonable interpretations of the concept. Consider the three following descriptions of a country behaving in a cooperatively reciprocal fashion:

- (1) Country A responds to a concession from country B with a concession, and responds to a retraction with a retraction.
- (2) Country A examines the recent actions of country B. If country B is increasing its concessions, country A increases its concessions. If country B is decreasing its concessions, country A decreases its concessions.
- (3) Country A compares its recent behavior to that of country B. If B's concessions have been greater than A's, A increases its concessions. If B's concessions have been less than A's, A decreases its concessions.

Each of the three descriptions falls within the cooperative reciprocity framework. The major difference between them lies in the type of calculation undertaken by A's decision makers. In the first description, A's decision makers simply assess whether B's move was a concession or a retraction, and respond in kind. In the second description, A's decision makers assess the change in B's recent behavior, and respond with a change in the same direction. In the final description, A's decision makers compare their recent behavior to that of B, assess the difference between B and A, and move in the direction of B's behavior.

Starting with the three cooperative descriptions, we can construct three parallel descriptions of inverse reciprocity by reversing the direction of A's response:

¹ For both frameworks it should be emphasized that although we are predicting response direction, either toward or away from one's opponent, we in no way are making a statement about the size of that response. Thus, a large concession in the cooperative reciprocity framework will not predict the size of the opponent's concession, but simply that a positive move of some size should then be made.

- (1) A responds to a concession from B with a retraction, and responds to a retraction with a concession.
- (2) A responds to a pattern of increasing concessions from B by decreasing its concessions. A responds to increasing retractions from B by increasing its concessions.
- (3) A compares its recent behavior to that of B. If B has shown the greater concessions, then A decreases its concessions. If A has shown the greater concessions, A increases its concessions.

These three sets of descriptions do not exhaust the possibilities for cooperative or inverse reciprocity,² but they do demonstrate two important points. First, a close examination of the term reciprocity reveals that it can have a variety of meanings, and there is little guidance in the standard descriptions of the term as to which is to be preferred. Second, reciprocity is usually taken to be a form of matching behavior, with A responding to B in a like manner. But for every description of cooperative reciprocity, we can construct a parallel description of inverse reciprocity, in which A responds to B's behavior in the opposite manner. With these points in mind, we turn to formalizing the three pairs of reciprocity models just discussed so that they can be subjected to systematic testing with Jensen's data on the strategic arms talks.

Formalizing the Models

To formalize the models discussed above, we have created a series of computer routines. Each routine produces a prediction for a particular cooperative reciprocity model during a negotiating round in the strategic arms limitation talks. When the predictions are compared to the observed data, positive correlations indicate that the particular cooperative reciprocity model fits the data, whereas negative correlations indicate that the parallel inverse reciprocity model fits the data. Aside from the routines to make these predictions, the program contains routines to create the observed behavior from the raw data, and to calculate the fit of each model.³

In order to facilitate further discussion of the three sets of models, we will use a descriptive shorthand for each. We recognize that often such labels serve to confuse, but a complete description of each model every time we wish to reference it would just be too cumbersome. The first pair of models, in which A discriminates only between concessions and retractions, and predicts a concession or retraction response, will be called directional reciprocity models. The second pair of models, in which A tracks B's recent behavior and increases or decreases its concessions accordingly, will be called trend reciprocity models. The final pair of models, in which A compares its recent behavior to that of B, and increases or decreases its concessions accordingly, will be called comparative reciprocity models. Tables 1-3 summarize the predictions from each of the models.

To summarize, we identified and formalized two sets of models of reciprocity in the form of a series of computer routines. One set is based on the notion of cooperative reciprocity: to respond to an opponent with an equivalent action. The other set is based on the notion of inverse reciprocity: to respond to an opponent with an opposite action. These two sets are mirror images

² For example, Wilson (1971) considers several variants on the tit-for-tat strategy discussed earlier. But his strategies focus on the duration of particular responses, and are less appropriate for the short series of data to be examined in this article, so they will not be included.

³ The program was written in Turbo Pascal, and runs on an IBM personal computer with or without an 8087 mathematical coprocessor. Program listings are available from Richard J. Stoll.

of each other, producing opposite predictions to an opponent's behavior. The individual models differ in the complexity of calculation involved on the part of the actor. In the next section, we describe the data set created by Jensen that will be the focus of our empirical investigation.

Data on Bargaining During the Strategic Arms Talks

Jensen abstracted all available proposals and suggestions made by the United States and the Soviet Union to each other during the 1969-1979 time period. These proposals were culled from the large number of accounts (principally in the form of memoirs by participants) of the U.S.-Soviet discussions. Reliance on these materials for the list of proposals and the weighing of them (described below) may lead to certain biases in the data; for example, all the accounts Jensen used were written from the American point of view. But one is hard pressed to come up with a better method of generating the data.

Each proposal and suggestion was assigned a score to indicate how large a concession or retraction it represented. Concession scores range from 1 to 5, and retraction scores range from -1 to -5. "A move toward the position of the other side, as defined at the time of the move, is regarded as a concession, and a retraction involves movement away from the other side's position"(Jensen, 1984: 537). The procedures for abstracting proposals and assigning scores were developed in Jensen's (1963: 522-528) earlier work on U.S.-Soviet disarmament negotiations.

For the 1969-1979 period, Jensen (1984: 537) coded 133 concessions and 6 retractions for the Soviet Union, and 81 concessions and 7 retractions for the United States. He cautions that little should be inferred from the fact that the Soviet Union made more concessions than the United States, as this may have been a product of differences in their initial negotiating positions. Jensen also identified 23 rounds of negotiation during the total period, and created summary scores for both parties in each round by summing their individual concession and retraction scores within a round.

The Empirical Analysis: Results and Discussion

We now proceed to test the various reciprocity models against the negotiating-round data. Following the analysis of Jensen, we break the entire set of rounds into three negotiations: SALT I (November 1969- May 1972), Vladivostok (June 1972-November 1974), and SALT II (January 1977-June 1979). We also test all the models against the entire set of rounds. To measure the fit of each model, we calculate tau B between the predicted and observed behavior (refer to Tables I -3 for the predictions from each model). We treat U.S. and Soviet behavior separately, so we can determine if different models account for each superpower's actions. Table 4 displays this information.⁴

Considering first the entire set of 23 rounds, the comparative reciprocity model provides the best fit for both the United States and the Soviet Union, although in neither case is the fit outstanding. In fact, the fit for both is a bit less than the simple tau B between the round scores of

⁴ Note that dividing the time period into these three negotiation periods leaves out five rounds that occurred between the signing of the Vladivostok accords and the first negotiating round under President Carter. These "missing" rounds, however, are included in the analysis of all negotiating rounds.

the two superpowers (.36).⁵ The comparative reciprocity model is also the best-fitting model in three of the six other runs, with the trend reciprocity model accounting for the rest. In each case, the best fits for the individual negotiations are higher than the fits for the entire set of negotiating rounds.

What of the two frameworks? Unfortunately, the results are not as clear-cut as we would like. In four of the six analyses on separate negotiations (as well as for the entire set of rounds for both parties), a cooperative reciprocation model provides the best fit. However, the Soviet Union in the Vladivostok negotiations and the United States in the SALT II negotiations both display hard-line bargaining stances (specifically, the inverse trend reciprocity model).

The good fit of the inverse model for the Soviet Union during Vladivostok can be attributed to the single Soviet proposal of the negotiating round from August to December 1973. It was referred to in the American government as the "21 demands" and had a number of very objectionable features. It called for the withdrawal of U.S. nuclear forces abroad, or at least Soviet "compensation" for these forces; the United States had steadfastly insisted during all negotiations that these weapons systems did not properly belong in the strategic arms talks. In addition, the Soviet proposal called for a ban on the development of any new missiles for a 10-year period. This would have prohibited the deployment of any new U.S. systems (such as MX and Trident), while still allowing the Soviets to improve their strategic forces by introducing "modernized" versions of their systems. Jensen codes this retraction with a score of -4; only one other Soviet proposal in the entire data set receives a score this low, and none receives a score of -5.

This strongly negative round sits in the middle of a basically progressive series of Soviet concessions. Soviet scores for the five rounds of the Vladivostok negotiation were 6, 7, -4, 22, and 18. We certainly do not want to dismiss the fit of the hard-line model as merely due to one piece of "bad" data. The proposal was an extremely negative one: Henry Kissinger (1982: 1015) called it a "piece of effrontery." It is clearly incompatible with all of the cooperative reciprocity models. But looking at the overall pattern of Soviet behavior during the five rounds, there is also evidence of a pattern of escalating concessions on their part.

Accounting for the fit of the inverse model for the United States during the Carter administration's SALT II negotiations is not as straightforward a task. U.S. scores during the six negotiating rounds were 6, 16, 8, 0, 8, and 10. The fit of the inverse model is good because of the decline in scores after the high concession of the second round, which took place from June to December 1977. The high level of concessions in the second round was part of an attempt by the United States and the Soviet Union to conclude SALT II before the expiration of SALT I (October 3, 1977). Similar flurries of concessions occurred by both sides in the final round of SALT I and Vladivostok (as well as in the final round of SALT II).

⁵ Note that the directional models never have the best fit for a negotiation, and in fact have a tau B of zero in four of the eight runs. One reason for these exceptionally poor fits lies in the simplicity of the models. Because most rounds have positive scores, the models produce the same prediction (a concession for the cooperative and a retraction for the inverse reciprocation models) through all the rounds of a negotiation. The lack of variability in predictions leads to the poor fits.

Given that the treaty was not completed at the end of the round, it is not surprising that the concession score for the United States is lower in the subsequent round (although the summary score for the Soviet Union rises slightly, from 12 in the second round to 13 in the third round). Further, over the next several rounds, the Soviet Union offered a number of "concessions" that were considered very unfair by the United States. For example, the Soviets offered to ban all new MIRVd ICBMs in return for an exemption for a single-warhead missile. Because the United States had no single-warhead missile under development (but did have the 10-warhead MX on the drawing board), this was seen as a very one-sided offer. A subsequent Soviet offer to ban all new missiles was also regarded as inequitable, because the Soviets would be giving up a single-warhead missile in return for the United States giving up the MX, as well as suspicions that the Soviets would find ways to "modernize" their missile force significantly without creating any new type of missile.

The pattern of negotiations in the first half of 1978 consisted of American rejections of these Soviet offers, as well as continued deadlock on several issues until the final rounds (for example, limits on the number of cruise missiles carried by bombers, the duration of the protocol, and the issue of telemetry encryption). Clearly, the pattern during the middle rounds is not one of cooperative reciprocity, although the model does correctly predict the U.S. action in the final round.

Conclusion

The 23 rounds of negotiation on strategic arms control between the United States and the Soviet Union during the 1969-1979 period are important not only for their substantive content, but for the valuable data they provide for the testing of a variety of models of the bargaining process.

In this article, we tested the fit of two very different (in fact opposite) kinds of reciprocity to the actions of the United States and the Soviet Union during the negotiations. Despite the small number of rounds in the three major negotiations we examined, we find that U.S.-Soviet interactions were characterized more often by cooperative reciprocity (various forms of tit for tat) than by inverse reciprocity (various forms of exploitation). However, for one negotiation for each side (Vladivostok for the Soviet Union, SALT II for the United States), an inverse reciprocity model provided the best fit to the data.

These findings of a pattern of reciprocity in U.S.-Soviet negotiations are consistent with previous quantitative studies by others (Jensen, 1963, 1968; Hopmann and Smith, 1977). This effort differs from these earlier analyses in the manner in which predictions are generated, and the number of explicit models that were tested. The convergence of findings using different methods across a variety of negotiations indicates that reciprocity is the dominant process in superpower negotiations.

Although reciprocity has turned up in a number of negotiation studies, much work remains to be done. Some effort needs to be made to extend research to levels below the negotiating round to provide a more detailed picture of the interactions of the parties. Additionally, the

reasons for the shift in the best-fitting model across the various negotiations for each side needs more attention than we devoted in this article.

Although acknowledging the need for more research, we feel we have demonstrated two important points: First, reciprocity characterizes the bargaining process between the United States and the Soviet Union on strategic arms, and usually it is cooperative reciprocity. Second, we need to examine more closely key terms such as reciprocity, to determine whether they have more than a single implication, and, if they do, we should make an effort to test all of the implications, not just one of them.

REFERENCES

- AXELROD, R. B. (1984) *The Evolution of Cooperation*. New York: Basic Books.
- BARTOS, O. (1974) *Process and Outcome of Negotiations*. New York: Columbia Univ. Press.
- HOPMANN, P. T. and T. C. SMITH (1977) "An application of a Richardson process model: Soviet-American interactions in test ban negotiations 1962-1963." *J. of Conflict Resolution* 21: 701-726.
- JENSEN, L. (1963) "Soviet-American bargaining behavior in the post-war disarmament negotiations." *Int. Studies Q.* 12: 152-160.
- — — (1968) "Approach-avoidance bargaining in the test ban negotiations." *Int. Studies Q.* 12: 152-160.
- — — (1984) "Negotiating strategic arms control, 1969-1979." *J. of Conflict Resolution* 28: 535-559.
- KISSINGER, H. (1979) *White House Years*. Boston: Little, Brown.
- — — (1982) *Years of Upheaval*. Boston: Little, Brown.
- NEWHOUSE, J. (1973) *Cold Dawn*. New York: Holt, Rinehart & Winston.
- SMITH, G. (1980) *Doubletalk: the Story of SALT I*. Garden City, NY: Doubleday.
- TALBOTT, S. (1979) *Endgame: the Inside Story of SALT II*. New York: Harper & Row.
- WILSON, W. (1971) "Reciprocation and other techniques for inducing cooperation in the Prisoner's Dilemma game." *J. of Conflict Resolution* 15: 167-196.