Reciprocity, Bullying, and International Cooperation: Time-series Analysis of the Bosnia Conflict
Author(s): Joshua S. Goldstein and Jon C. Pevehouse
Source: *The American Political Science Review*, Vol. 91, No. 3 (Sep., 1997), pp. 515-529
Published by: American Political Science Association
Stable URL: http://www.jstor.org/stable/2952072
Accessed: 13/05/2009 15:31

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=apsa.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

American Political Science Association is collaborating with JSTOR to digitize, preserve and extend access to *The American Political Science Review.*

http://www.jstor.org

Reciprocity, Bullying, and International Cooperation: Time-series Analysis of the Bosnia Conflict

JOSHUA S. GOLDSTEIN *American University*
JON C. PEVEHOUSE *Ohio State University*

Although the role of reciprocity in international cooperation is central to neoliberal institutionalism, empirical understanding of the concept remains weak. We analyze strategic response patterns—the use of reciprocity or inverse response (bullying)—in the Bosnia conflict from 1992 to 1995. We construct weekly time series of conflict and cooperation among the parties to the Bosnia war, using machine-coded events data. Time-series statistical analysis identifies several important patterns of strategic response, both reciprocal and inverse. These include bilateral responses, which are central to the concepts of reciprocity and evolution of cooperation, and triangular responses, which are central to the debates on containment versus accommodation in regional conflicts. Specifically, Serb forces displayed inverse triangular response, cooperating toward Bosnia after being punished by NATO. Outside powers displayed triangular reciprocity, increasing hostility toward Serb forces after Serbian attacks on the Bosnian government.

The role of reciprocity in international cooperation is central to neoliberal institutionalism. Theorists in this tradition argue that reciprocity can promote the evolution of cooperation in situations of mixed interest, such as prisoner's dilemma (Axelrod 1984). Because it does not depend on central authority, reciprocity as a strategy for cooperation seems well suited to the "anarchic" realm of international relations, where (as even neoliberals grant) states operate autonomously in pursuit of self-interest (Axelrod and Keohane 1986, Keohane 1986a).

Yet, our empirical understanding of this central theoretical concept remains weak. Evidence regarding the existence and nature of international reciprocity is still largely limited to a few cases of great-power relations (mostly involving the United States and the Soviet Union) and is contentious.[1] Furthermore, theoretical insights from formal models have proven difficult to apply in the real-world international conflicts that occupy both empirical researchers and policymakers (see Goldstein and Freeman 1990, 1991). David Baldwin (1993, 22), in his edited volume on the neorealist-neoliberal debate, concludes:

> The most important research need is better understanding of the conditions that promote or inhibit international cooperation. The debate between neorealism and neoliberalism has generated at least six hypotheses worthy of more research and testing. The first concerns the strategy of reciprocity. Both the theoretical and practical conditions under which such strategies promote cooperation deserve attention.

Systematic testing of the role of reciprocity in international cooperation requires attention to three interconnected elements largely ignored by neoliberal theorists to date. First, in the post–Cold-War era, international cooperation hinges on the great-power management of regional conflicts more often than on the evolution of cooperation between two relatively equal parties, such as the Cold-War superpowers.[2] Yet, although managing regional conflicts has come to preoccupy policymakers, it has received less attention from scholars of international cooperation.

Second, in the context of managing regional conflicts, the concept of reciprocity must include triangular responses: Outside actions toward a regional power aim to influence that power's behavior toward other regional targets. For example, although Saddam Hussein's behavior toward the Western powers in 1990 presented some problems (notably taking hostages), the main task for great powers was changing his behavior toward Kuwait and Saudi Arabia. These triangular aspects of reciprocity have not been studied adequately in the theoretical and empirical literature, which has focused on bilateral reciprocity.[3]

Joshua S. Goldstein is Professor of International Politics and Foreign Policy, American University, Washington, DC 20016-8071. Jon C. Pevehouse is a Ph.D. candidate in political science at Ohio State University, Columbus, OH 43210.

This research was supported in part by the National Science Foundation (SBR-9617157), as well as American University's School of International Service and Ohio State University's Department of Political Science. For research assistance we thank Elizabeth Kittrell and Nicole Bair. For useful suggestions, we thank Phil Schrodt, John Freeman, Gary King, Bruce Russett, Barry O'Neill, David Lumsdaine, Brad Westerfield, Allan Stam, David Bearce, and Elizabeth Erickson.

[1] Even in the highly studied area of superpower arms races, empirical research has produced no consensus, whether from statistical analyses or case studies (George, Farley, and Dallin 1988; Patchen 1991). The dearth of systematic empirical work on reciprocity persists, moreover, seventeen years after Zinnes (1980) noted in her presidential address to the International Studies Association the contradictory results of empirical studies of reciprocity as one of three major "puzzles" facing the field of international relations.

[2] U.S.–Soviet relations provided the exemplary case of the problem of inducing mutual cooperation in a mixed relationship with a rival or potential rival of roughly equal power (e.g., Axelrod 1984, vii). As U.S. policy toward China is now debated in a similar framework (see Goldstein 1995), this problem clearly did not end with the Cold War, but it is no longer so central.

[3] Triangular responses among great powers also can occur (Goldstein and Freeman 1990, 1991), but triangularity is even more important in regional conflicts because the relationships of local parties are overlaid with the possible actions of outside powers. Triangular relationships also are possible among local parties in a regional conflict—such as Serbian, Croatian, and government forces

Third, in these triangular response relationships, the concept of reciprocity must be paired with its opposite, inverse response—which Hirshleifer and Coll (1988) in a game-theoretic context call a "bully" strategy. Inverse response, or taking advantage of cooperative actions but cooperating in response to hostile ones, is thought to characterize aggressive states, such as Nazi Germany in the late 1930s or Iraq in the 1990s. Yet, the concept has received scant attention to date in either the formal or empirical research on international cooperation.

Policymakers pay great attention to how particular states respond to actions, reciprocally or inversely, since the two types of response call for opposite strategies in managing relations with those states. Toward states that reciprocate, appropriate policies are "soft-line," using cooperative initiatives to promote future mutual cooperation. These soft-line policies underlie strategies of accommodation (or engagement). By contrast, appropriate policies toward states that respond inversely are "hard-line," using hostile initiatives to induce cooperation from an aggressive, opportunistic state, that is, a strategy of containment.[4]

A state may not follow the same response pattern triangularly as bilaterally. Thus, in analyzing the potentials for eliciting international cooperation in regional conflicts, both theoretical and policy concerns call for attention to the full range of response types—triangular along with bilateral responses, and inverse along with reciprocal responses. Several specific patterns of triangular response are of special interest.

Triangular response by a regional actor occurs when the actor changes its behavior toward another regional actor in response to the recent behaviors of an outside power. This response can be either reciprocal or inverse. If the regional actor responds with triangular reciprocity, it will increase cooperation toward a neighbor in response to cooperation from an outside power (and decrease cooperation in response to hostility). In this case, the outside power would best follow a strategy of accommodation and employ cooperative initiatives to induce local cooperation.

By contrast, if the regional actor uses triangular bully (inverse) responses, it will cooperate more toward its neighbor in response to the hostile actions of an outside power (and become more hostile toward its neighbor in response to the cooperation of an outside power). In this case, outside powers would best follow a containment strategy, using hostile initiatives to induce local cooperation.

For their part, outside powers may respond to local actions by using triangular reciprocity. This means that the outside power varies its behavior toward a regional actor in direct proportion to the regional actor's recent behavior toward another regional actor. The outside power would thus punish a regional actor's local misdeeds and reward its local good deeds. Such a response is consistent with the management of the conflict by either containment (if combined with hostile initiatives in expectation of an inverse triangular response by the regional actor) or accommodation (if combined with cooperative initiatives in expectation of a reciprocal triangular response).[5]

Intervention to manage a regional conflict may engender a collective goods problem among the outside powers who pay for the intervention. Some neoliberal theorists have suggested that the presence of a hegemonic state in the international system can help solve such problems. If so, then a hegemonic country—possibly including the United States in the 1990s—might play a distinctive role in the management of regional conflict.

HYPOTHESES AND MODEL

Our research aimed to assess empirically whether these various patterns of strategic response actually exist in the interactions that make up a real-world regional conflict—the war in Bosnia. In order to infer patterns of response by actors to the behaviors of other actors, we used time-series statistical analysis. We converted the actions of each actor toward each other actor into long weekly time series representing net cooperation (weighted cooperative actions minus weighted hostile actions), then analyzed how each actor's behavior correlates with the other actors' recent past behaviors, taking into account the first actor's own recent behavior. For example, do recent past actions by the international community help explain present actions by the Serbian forces toward the Bosnian government? By looking at thousands of large and small actions as they played out over time, we can draw general inferences about response patterns such as reciprocity and bullying (inverse response) in both their bilateral and triangular aspects.

Specifically, we wanted to know whether the bilateral reciprocity thought to characterize the "evolution of cooperation"—and/or the triangular responses which might characterize outside management of regional conflict—were actually present in the Bosnian case. We focused on the two most important dyads (with tests for omitted-variable bias): that of Serb forces and the international community (the United Nations and NATO members), and that of Serb forces and the Bosnian government.[6] We tested five hypotheses.

HYPOTHESIS 1 (BILATERAL RECIPROCITY): *The net cooperation of one actor toward another positively affects subsequent net cooperation by the second actor toward*

in Bosnia—but our focus is on great-power actions toward one or more local parties with the goal of inducing cooperation among them.

[4] On soft- and hard-line policies, see Snyder and Diesing (1977, 298–303) and Jervis (1988, 326); Axelrod's (1984) "nice" resembles soft line. The distinction parallels Jervis's (1976, 78–82) classic "spiral" and "deterrence" models. Unfortunately for policymakers, no consensus exists regarding the particular states or circumstances to which each model applies. Most scholars agree that neither strategy works all the time (see Keohane 1986b, 3).

[5] In principle, an outside power could use inverse triangular response in reaction to the local actions of a state. But such responses are impractical and theoretically uninteresting.

[6] We later disaggregate the international community into U.S. and European actors to explore a possible hegemonic U.S. role in conflict management.

the first. (There are four actor-target pairs to test in the international-Serbian and Serbian-Bosnian dyads.)

HYPOTHESIS 2 (BILATERAL BULLYING): *The net cooperation of one actor toward another negatively (inversely) affects subsequent net cooperation by the second actor toward the first. (The same four actor-target pairs can be tested.)*

HYPOTHESIS 3 (TRIANGULAR RECIPROCITY BY INTERNATIONAL COMMUNITY): *The net cooperation of the Serbian forces toward the Bosnian government positively affects subsequent net cooperation by the international community toward the Serbian forces.*

HYPOTHESIS 4 (TRIANGULAR RECIPROCITY BY SERBIAN FORCES): *The net cooperation of the international community toward Serbian forces positively affects subsequent net cooperation by Serbian forces toward the Bosnian government.*

HYPOTHESIS 5 (TRIANGULAR BULLYING BY SERBIAN FORCES): *The net cooperation of the international community toward Serbian forces negatively affects subsequent net cooperation by Serbian forces toward the Bosnian government.*[7]

To test these hypotheses we used a time-series model. Conceptually, our model derives from the vector autoregression (VAR) approach (Freeman, Williams, and Lin 1989). That is, we defined each actor's behavior toward each other actor as a (time-series) variable, then regressed each one's behavior (at time t) on its own recent past behavior and the recent past behaviors of all the other variables in the system. Because the right-hand-side terms are identical for each equation (there is one equation per variable), the estimation uses ordinary least squares (OLS). Since we ultimately found that only one lagged term needed to be included for each independent variable, we were able to interpret the sign and significance of coefficients directly, largely bypassing the more controversial aspects of VAR. Although VAR proved useful for model specification, our actual model was a simple set of OLS equations regressing each time series on lagged terms for all time series:

Model 1:

$$SB_t = \beta_{10} + \beta_{11}SB_{t-1} + \beta_{12}BS_{t-1} + \beta_{13}IS_{t-1} + \beta_{14}SI_{t-1} + \epsilon_1;$$

$$BS_t = \beta_{20} + \beta_{21}SB_{t-1} + \beta_{22}BS_{t-1} + \beta_{23}IS_{t-1} + \beta_{24}SI_{t-1} + \epsilon_2;$$

$$IS_t = \beta_{30} + \beta_{31}SB_{t-1} + \beta_{32}BS_{t-1} + \beta_{33}IS_{t-1} + \beta_{34}SI_{t-1} + \epsilon_3;$$

$$SI_t = \beta_{40} + \beta_{41}SB_{t-1} + \beta_{42}BS_{t-1} + \beta_{43}IS_{t-1} + \beta_{44}SI_{t-1} + \epsilon_4;$$

where SB is Serb forces' level of net cooperation toward the Bosnian government; BS is the Bosnian government's level of net cooperation toward Serb forces; IS is the international community's level of net cooperation toward Serb forces; and SI is the Serb forces' level of net cooperation toward the international community.

In this model, the test for hypothesis 1 (bilateral reciprocity) is for the significance and positive sign of β_{12}, β_{21}, β_{34}, and/or β_{43}. The test for hypothesis 2 (bilateral bullying) is for significance and negative sign for any of the same four coefficients. The test for hypothesis 3 (triangular reciprocity by the international community) is for the significance and positive sign of β_{31}. The test for hypotheses 4 or 5 (triangular reciprocity or triangular bullying by Serb forces) is for the significance of β_{13}, with either positive or negative sign, respectively. This last coefficient provides the test of greatest interest in terms of both theoretical importance (triangular response by regional actors) and policy salience (appropriateness of accommodation versus containment).

COMPETING VIEWS OF THE CONFLICT

The war in Bosnia is an excellent case to test for several reasons. First, it matters greatly to policymakers, as it has become for better or worse a paradigm for the problem of regional conflict management in the post–Cold-War era.[8] Second, the complexity of the conflict and the repeated outside attempts to manage it provide numerous possibilities for triangular and bilateral response behaviors. Third, no consensus exists about the nature of the parties' response patterns or the appropriateness of various possible great-power intervention strategies.

The controversies about great-power policies toward the Bosnia conflict hinge on the question of reciprocal or inverse triangular response by Serb forces. Two schools of thought quickly emerged in Western capitals about the nature of the war and each called for different policy approaches. One school saw the war as a case of aggression against a UN member, which had a tradition of multiethnic tolerance, by ultranationalist forces using genocide as an instrument of territorial conquest. There were bad guys (the Serbian leadership) and good guys (the Bosnian government and society).[9] Because the assault on Bosnia struck at fundamental principles—genocide is evil and territorial integrity should be preserved—the war was a challenge to world order that required (both politically and morally) an effective response by the international

[7] In any particular test, either hypothesis 1 or 2 (or neither), but not both, can be true. The same applies to hypotheses 4 and 5.

[8] Bosnia received the largest UN peacekeeping force in history and the first use of force by the NATO alliance. As Woodward (1995, 2) notes, "by 1994 this conflict of little significance had emerged as the most challenging threat to existing norms and institutions that Western leaders faced."

[9] See, for example, Donia and Fine 1994, Malcolm 1994. Some variants focused on "the Bosnian Muslims." Croatia and Croatian forces play an ambiguous role, as aggressor or victim, depending on time and context.

community.[10] This aggressor-victim school of thought was supported by Bill Clinton when a candidate (and at times as president), Senator Robert Dole, a bipartisan majority in the U.S. Congress, the majority of the UN General Assembly, the Islamic Conference, and (not surprisingly) the Bosnian government.

The other school of thought saw Bosnia as a civil war among essentially equivalent actors, notwithstanding some power differences among them. There were no aggressors and victims but warring factions (or parties), with plenty of blame to go around.[11] Supporters of this view were more likely to accept ethnic nationalists on all sides as the legitimate political leaders,[12] to see autonomy for each ethnic group (even through territorial partition) as more important than preserving a multiethnic society, and to view conflict among the groups as the inevitable result of "ancient hatreds."[13] This school of thought was supported by presidents Bush and (sometimes) Clinton, the great powers on the UN Security Council (especially Russia), UN Secretary-General Boutros-Ghali, the European Union, and (not surprisingly) Serbia.

Although the two schools of thought differed in their views of the importance of Western intervention in Bosnia, their most fundamental difference concerned the nature of Serbian responses to international actions. As a result, they produced two mutually exclusive policy recommendations.[14] If the Serb forces were aggressive bullies, as members of the aggressor-victim school believed, then they would only cooperate toward Bosnia if confronted with force (hypothesis 5). One policy variant would have accomplished this by lifting the international arms embargo on Bosnia. A second variant would have used air strikes against Serb forces. In both cases, increased hostility toward the Serb forces was held to be the best option for inducing Serbian cooperation toward the Bosnian government. This is a containment strategy, responding to a regional bully in a triangular setting. (Some observers also expected Serb forces to cooperate *toward the international community* in response to international hostility, which would be a bilateral inverse response, as in hypothesis 2.)

According to the warring-factions school of thought, by contrast, the Serb forces were not bullies and would more likely respond to accommodation than containment. Lifting the arms embargo or using air strikes would increase Serbian hostility toward the Bosnian government (hypothesis 4) and probably toward the international community as well (hypothesis 1), leading to an upward spiral of violence in the region. The best policy was one of neutrality, mediation, and monitoring while using UN peacekeepers to try to deliver humanitarian aid to civilians. This policy generally prevailed among the great powers until September 1995.[15]

Both President Clinton and the UN Security Council showed ambivalence by advocating containment in words but accommodation in deeds. Hoffmann (1994) argues that in Bosnia, as in Ethiopia and Munich in the 1930s, "the international community made the mistake of simultaneously pursuing two incompatible policies—collective security against aggression, and a negotiated compromise between parties that were treated as morally equivalent."[16] Because the warring-factions and aggressor-victim schools rest on different assumptions about the specific response patterns of Serb forces (hypothesis 4 versus 5), our tests for triangular reciprocity and bullying can inform the debate.

METHODS

We used events data to examine the interactions of the actors in the Bosnia conflict. Events data are particularly well suited to the analysis of reciprocity and bullying since they are a systematic collection of interactions among the actors coded in a particular domain (Azar, Brody, and McClelland 1972). Many objections to past events data research center on coding inconsistencies and biases (Andriole and Hopple 1984, Laurance 1990). The costs of coding have also slowed the construction and extension of events data sets, at times leaving the data many years behind the evolving practice of international relations. These problems of event coding have been mitigated, however, through the use

[10] To turn back Serbian aggression would cost something, but it could be done and was cheaper than setting a bad precedent that would encourage aggression and genocide elsewhere. For example, the *Washington Post* editorialized in May 1995: "Bosnia is not of direct strategic significance to the United States. But the final abandonment of Bosnia would rip at the threads of international order and harden a cruel post–Cold-War calculus based on the general perception of what aggressors can get away with."

[11] See Boyd (1995). For example, Jimmy Carter argued against taking sides in regional conflicts such as Bosnia: "In most cases, both sides are guilty of atrocities" (in Brinkley 1995, 96).

[12] We use the term "ethnic" for Bosnia's "national" groups since U.S. scholars refer to "ethnic conflict" in such cases. In Bosnia, however, the group differences hinge on religious background, not ancestry or language, which are both shared.

[13] See Kaplan 1993, Mearsheimer and Van Evera 1996. For example, President Clinton said after the February 1994 "marketplace massacre" in Sarajevo that "until those people get tired of killing each other, bad things will continue to happen." Woodward (1995, 285) notes: "Using arguments . . . that the hostilities were the result of ancient ethnic and religious hatreds . . . the West was again able to justify not deploying troops."

[14] In theory the two dimensions of interventionism-isolationism and of accommodation-containment need not be conflated. In practice only two poles of policy analysis emerged, since advocates of containment were also much more inclined toward intervention in some form. The two poles did not align on traditional political fault lines. Conservatives concerned about response to aggression and liberals concerned about human rights and genocide joined forces to support the aggressor-victim school and its containment policy. Isolationists, worried about costly overseas interventions, joined with multilateralists, worried about splitting the Atlantic alliance, to support the warring-factions school and its policy of accommodation.

[15] Even after the war ended, debate continued about the nature of Serbian responses—about whether strict enforcement of Dayton provisions against Serb forces (such as arresting indicted war criminals) would elicit Serbian compliance or provoke Serbian hostilities.

[16] The Security Council first defined the issue as aggression and territorial integrity. Resolution 752 (May 15, 1992) demands "that all forms of interference from outside Bosnia-Herzegovina . . . cease immediately," that states respect Bosnia's territorial integrity, and that irregular military forces in Bosnia be disbanded, withdrawn, or placed under Bosnian government control. Resolution 757 on May 30 followed suit. But by June the council had turned to humanitarian assistance, peacekeeping, and mediation.

TABLE 1. Event Count by Variable, January 1991–December 1995

Actor	Target	Variable Name	Number of Events
International (UN and NATO members)	Serbia and Serb forces	*IS*	2,969
Serbia and Serb forces	International (UN and NATO)	*SI*	1,699
Serbia and Serb forces	Bosnian government/army	*SB*	1,766
Bosnian government/army	Serbia and Serb forces	*BS*	1,091
International (UN and NATO)	Bosnian government/army	*IB*	2,228
Bosnian government/army	International (UN and NATO)	*BI*	993
Croatia and Croatian forces	International (UN and NATO)	*CI*	719
International (UN and NATO)	Croatia and Croatian forces	*IC*	1,114
Croatia and Croatian forces	Serbia and Serb forces	*CS*	958
Serbia and Serb forces	Croatia and Croatian forces	*SC*	820
Croatia and Croatian forces	Bosnian government/army	*CB*	634
Bosnian government/army	Croatia and Croatian forces	*BC*	627
Total			15,618

of machine-readable data sources and machine coding (Schrodt and Gerner 1994).

This study uses the Kansas Event Data System (KEDS) software to generate events data for the Bosnia conflict (see Gerner et al. 1994; Schrodt 1995; Schrodt, Davis, and Weddle 1994; Schrodt and Gerner 1994).[17] In this study, as in past work using KEDS, lead sentences of Reuters News Service articles were coded.[18] Leads—the first sentence of an article—usually contain the "who did what to whom" information: For example, "Bosnian Serb forces began heavily bombarding the centre of the Moslem stronghold city Tuzla on Monday." A Lexis/Nexis search identified any article in 1991–95 whose lead contained any of the following word roots: Yugoslav, Bosnia, Serb, Croat, Sarajev.[19] The 38,837 potentially codable leads for 1991–95 were downloaded, cleaned, and placed in KEDS-compatible format; most duplicate stories were removed.

KEDS uses a "sparse parsing" technique to extract the subject, verb, and object in a sentence. It uses two dictionaries (one of verbs, one of actors) to perform pattern matching. KEDS recognizes pronouns (which it dereferences), conjunctions, and passive voice constructions, which it converts to active voice (Schrodt 1995). KEDS assigns a type (actor, verb, pronoun, etc.) to each word, which it matches in a dictionary. KEDS then pattern-matches each verb using the verb dictionary. A subject (source) and object (target) are determined. The source is usually the first actor in the lead, while the target is usually the first actor after the verb (Gerner et al. 1994).

The KEDS dictionaries are customized for a particular region or issue, initially through machine-assisted coding. When KEDS codes an event incorrectly, one or both of the dictionaries are modified to correct the miscoding. For example, KEDS correctly coded the following lead as a double event of "agreement," with Bosnian-Muslims and Croatians as both the target and source: "Bosnia's Moslem and Croat leaders, once allies against rebel Serbs but now sworn foes, have agreed safe passage for convoys carrying relief through central Bosnia." But the following partial lead was initially (with undeveloped dictionaries) coded as a UN meeting with Bosnians: "United Nations forces are to mediate the release of two Bosnian Croat commanders." After adding the actor phrase "Bosnian Croat" to reference Croatian forces, the event was coded as a meeting between the UN and Croatian forces. The following lead was miscoded, and no dictionary adjustments we made could correct the coding: "German Chancellor Helmut Kohl on Wednesday described a Russian statement accusing NATO of genocide against Bosnian Serbs as unacceptable." We monitored and updated the dictionaries until they achieved roughly 85–90% coding accuracy, equivalent to human coders (Gerner et al. 1994, 99–102).

Once the dictionaries were sufficiently stable, all the leads were machine-coded with no human intervention. A number of events were excluded because they could not be attributed to a specific actor. Others were excluded because either the actor or target was not one of the actors we studied (e.g., Greece, Russia, Hungary) or because actor and target were the same. Just over 15,000 events containing relevant actors were coded for 1991–95, as shown in Table 1.[20]

We used the World Event Interaction Survey (WEIS) coding scheme, which has 63 event codes—

[17] The KEDS program and documentation are available on the Internet (http://www.ukans.edu/~keds).

[18] Reuters provides very dense coverage and tends to have less bias than other major wire services, such as UPI or Agence France Presse in terms of which events are and are not reported (Huxtable and Pevehouse 1996, Howell and Barnes 1993). The war in Bosnia is a particularly good source of events since it was so intensively and continuously reported by U.S. and international media; for example, it received more minutes of U.S. network news time in 1994 than any story except the O.J. Simpson trial (Tyndall Report 1995).

[19] Using roots ensured that words such as Yugoslavian, Bosnian, and Serbian would be found. The Lexis/Nexis search excluded any leads matching the criteria and containing the following roots: soccer, sport, Olympic, basketball. This reduced noise in the data, since Reuters combines news, sports, and human-interest stories in the same database.

[20] During crises, a particular event such as a military attack is often coded several times, but we do not consider this a serious problem since it simply means that added weight attaches to important events deemed worthy of repeated Reuters coverage.

such as diplomatic visits or economic sanctions (McClelland 1976). Each verb root and pattern is assigned a three-digit WEIS code, or a null code indicating that the event should not be recorded. We converted the ordinal WEIS codes to interval data using the Goldstein (1992) net-cooperation scale. The scale assigns a weighting (how friendly or hostile a type of action) to each WEIS event code, on a scale ranging from conflict (below zero) to cooperation (above zero). A given data point, for a discrete unit of time, represents a level of net cooperation (for a given actor toward a given target), which is the sum of all the cooperative events in that period (each times its weight) minus all the conflictual events (weighted).[21]

In converting coded events into such a time series, one must choose an appropriate level of temporal aggregation (Freeman 1990, Goldstein 1991). High levels of aggregation (such as quarterly or annual data) tend to swallow up important interaction effects. Since the events in Bosnia moved relatively fast, and to make sure the series captured the dynamics of the conflict, we looked at daily and weekly aggregations over the 1,379 days, or 197 weeks, from March 1992 to December 1995. In contrast to Goldstein and Freeman's (1990, 1991) monthly aggregated Cold-War data on strategic great-power interactions, the Bosnia war presents a string of rapid episodes with frequent twists and turns. The high density of the data can support a finely grained level of aggregation.

Daily data, however, had several disadvantages. First, the timing of events as reported by Reuters was not always accurate down to the day, and major events tended to be reported several times over several days in different leads. Second, the daily time series have many strings of zeroes (days when nothing was reported; nothing new happened), and we feared this data structure would introduce artificial correlation (possibly confusing, for example, autoregression with response to other actors). Finally, in specification tests to determine how many lagged terms to include (which are described shortly for the weekly data), we found that 18 daily lags of each variable should be included as explanatory variables in each equation, meaning that a model using daily data would have dozens or hundreds of right-hand coefficients. Since the responses were occurring over more than a week, we decided on weekly data aggregation (Sunday through Saturday).

The construction of our time series makes our analysis much more prone to Type II error (failing to detect a real relationship) than Type I error (detecting a relationship which does not really exist) for several reasons. First, like most events data, our series are "noisy," containing random error overlaid on the true data but few systematic biases. Such noise will produce inefficient estimates but not biased ones (Goldstein 1991). Second, the use of weekly rather than daily aggregation is similarly more likely to wash out significant coefficients for true relationships than to inflate the significance of phantom relationships. Weekly aggregation masks those very fast interactions that occur within a week; these interactions are reflected, therefore, only in the contemporaneous correlation of residuals from our model. Third, weekly data mean that each period analyzed had just a few dozen data points. All these considerations make the results as reported conservative—we are more confident that significant relationships really do exist than we are that nonsignificant relationships really do not exist. Thus, we occasionally discuss probability levels above the traditional significance standard of $p < .05$ in analyzing overall patterns of results across a variety of tests.

DATA VALIDITY AND TEMPORAL PHASES

In this section we will briefly review the history of the war with reference to graphs of our time series. Face validity tests, that is, examinations of these and other graphs to identify such known characteristics of the war as crises or cease-fires, convinced us that the event coding was working well. We also will describe and justify the potential break points, which will later be used to test coefficient stability through time. The placement of these possible break points, or delimiting phases of the war which may have particular response patterns, should reflect actual historical phases and changes so that the subsequent analysis captures changes in response patterns over time.

Figure 1 shows our four potential break points, that is, times at which we had reason, a priori, to think that strategies and patterns of response may have changed (we will test statistically whether they did).[22] At each point, great-power policymakers were reported to have agonized about whether and how to change policy toward Bosnia, typically in light of a dramatic failure of previous policy. The four points are the weeks ending 4/17/93, 2/12/94, 12/17/94, and 7/29/95 (hereafter, we drop the "week ending" designation, implied by a given date).

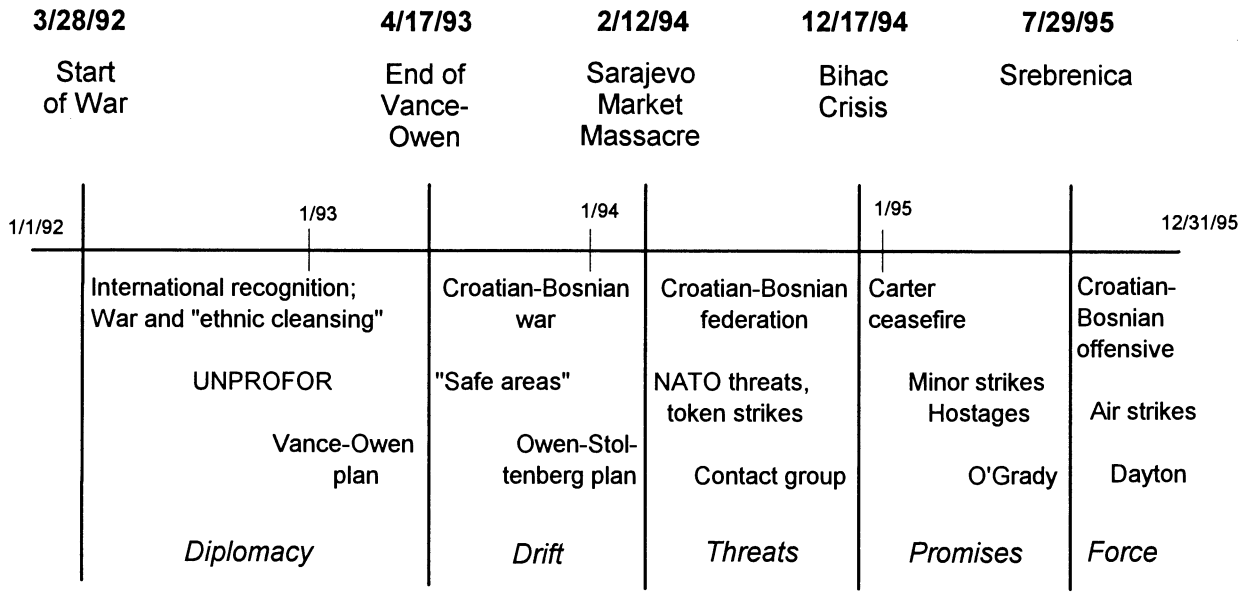
Figure 2 shows our weekly time series for the international-Serbian dyad; Figure 3 shows the Serbian-Bosnian dyad. The potential break points are also indicated. We will now describe each potential break point and the character of each phase between them, with reference to the time series in figures 2 and 3.

We begin our time series in the first week of March 1992 and, with allowance for several lags, begin the analysis with 3/28/92—just before the outbreak of war in earnest, which followed the international recognition of Bosnia on April 6. The early months of the war were very intense, with waves of territorial conquest and "ethnic cleansing" by Serb forces. In Figure 3 the time series drops accordingly (on the net-cooperation scale, intensified conflict drives the time series below zero). As the initial attacks let up, the conflict settled

[21] Some researchers argue that the conflict and cooperation dimensions should not be combined; see discussions in Goldstein 1992, 370–4, and Goldstein and Freeman 1990, 41.

[22] Because the typical behaviors of Serb forces and of the Bosnian government seem to remain fairly constant throughout, whereas great-power and Croatian behavior switch dramatically at several points, we focused on these latter behaviors in defining potential break points.

FIGURE 1. Potential Break Points in the Bosnia Conflict



into a somewhat less violent, ongoing war through spring 1993.

The international response during this first year was cautious, as reflected in the time series in Figure 2, which shows only modest levels of conflict in the international-Serbian dyad. President Bush was under pressure in an election year to focus on domestic issues, and European powers hoped that negotiations could bring about a settlement without a costly intervention. The UN Security Council authorized the UN Protection Force (UNPROFOR), which pursued humanitarian assistance and mediation. The Vance-Owen Plan proposed a complex system of ten cantons for Bosnia. We characterize this period overall as "diplomacy."[23]

Our first potential break point, in *mid-April 1993*, is marked by several important changes. First, Croatian forces and the Bosnian government began fighting each other, creating a more complex, three-sided war. Second, the Vance-Owen plan was rejected by Serb forces, dashing international hopes for a diplomatic breakthrough. Third, President Clinton—who had favored supporting the Bosnian government with arms and air strikes—had recently abandoned that policy when faced with European resistance. Finally, the UN declared six Bosnian government-held cities to be "safe areas" where civilians (including many refugees) must not be attacked. Five were enclaves surrounded by Serb forces (Sarajevo, Gorazde, Bihac, Srebrenica, and Zepa). They would become the flashpoints in the war, as Serb forces tried to starve or shell them into surrender, the Bosnian government tried to break out militarily, and the great powers tried to decide where to draw an ever-shifting line for acceptable behavior.

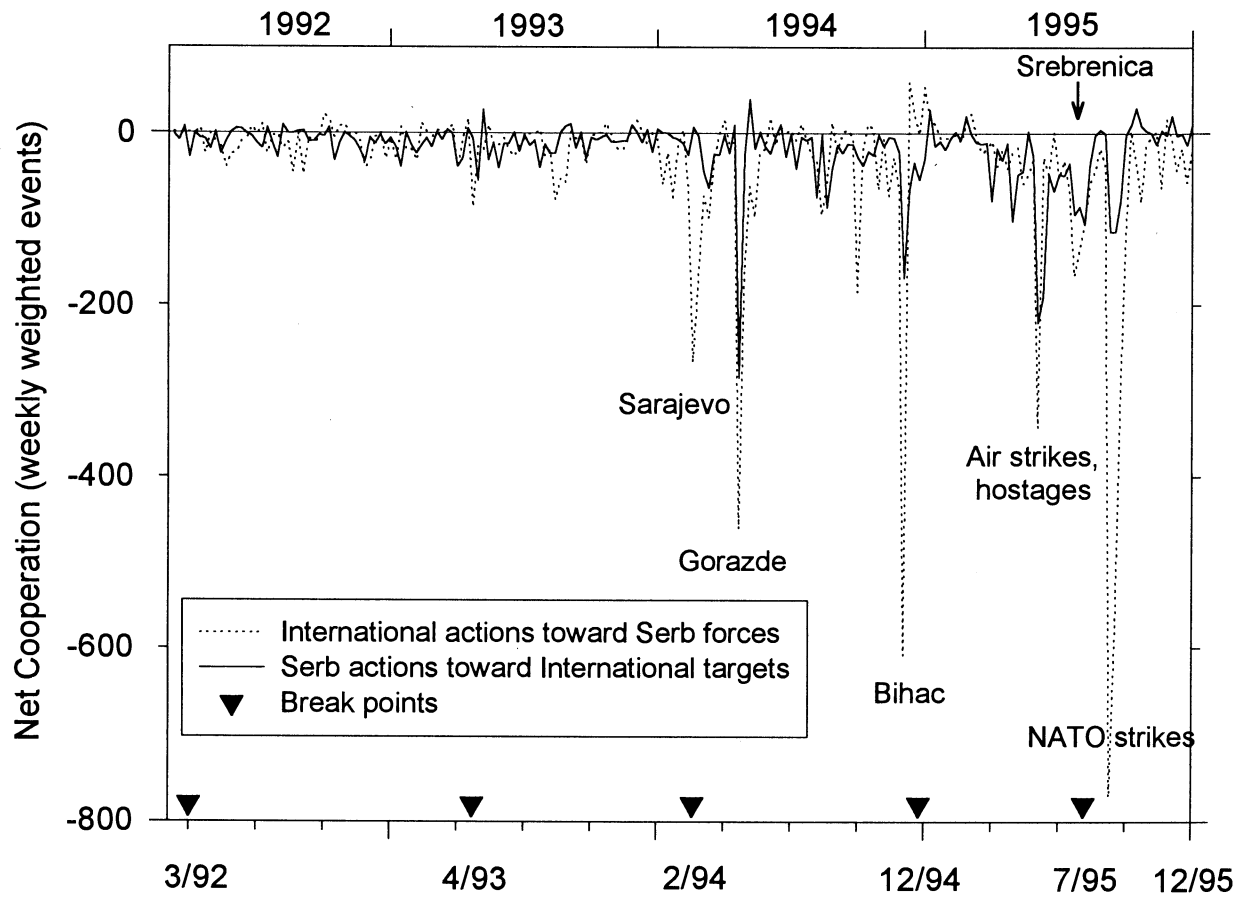
[23] These short-hand characterizations are for convenient reference and do not affect the actual analysis.

During this period, we characterize the great powers' management of the conflict as "drift" following the failures of the various Western ideas for ending the war. The time series in Figure 2 continue to show little action in the international-Serbian dyad until late in the period. Meanwhile, the Croatian-Bosnian war created the possibility for occasional cooperation—though there was still mostly conflict—in the Bosnian-Serbian dyad, as reflected in Figure 3. Eventually, European and UN negotiators crafted the Owen-Stoltenberg plan, which came close to the outright three-way partition of Bosnia that Serb forces demanded. Then, in early 1994, intense Serbian attacks on Sarajevo (shown as a negative spike in figures 2 and 3) caused the great powers once again to reevaluate their policies and begin to threaten air strikes.

The attacks on Sarajevo culminated in the shelling of a crowded marketplace in February 1994. We place our second potential break point here. Western policymakers seemed to draw a line at the fall of Sarajevo, and a NATO ultimatum was taken seriously by Serb forces, which backed off. NATO shot down four Serbian jets in February and, during the rest of 1994, carried out several token air strikes (with symbolic rather than military effect), which are visible in figures 2 and 3. The period after February 1994 also differed from the preceding period in that the Croatian-Bosnian war gave way to a federation (on paper, but a real ceasefire). Overall, we call the great-power strategy in this period one of "threats."

In early *December 1994*, the U.S. administration deliberately changed course away from NATO threats and pinprick air strikes, which seemed not to be working; these had not halted a Serbian attack on Bihac, and Serb forces had taken hundreds of UNPROFOR troops hostage (see Kelly 1994). President

FIGURE 2. Net Cooperation in International-Serbian Dyad



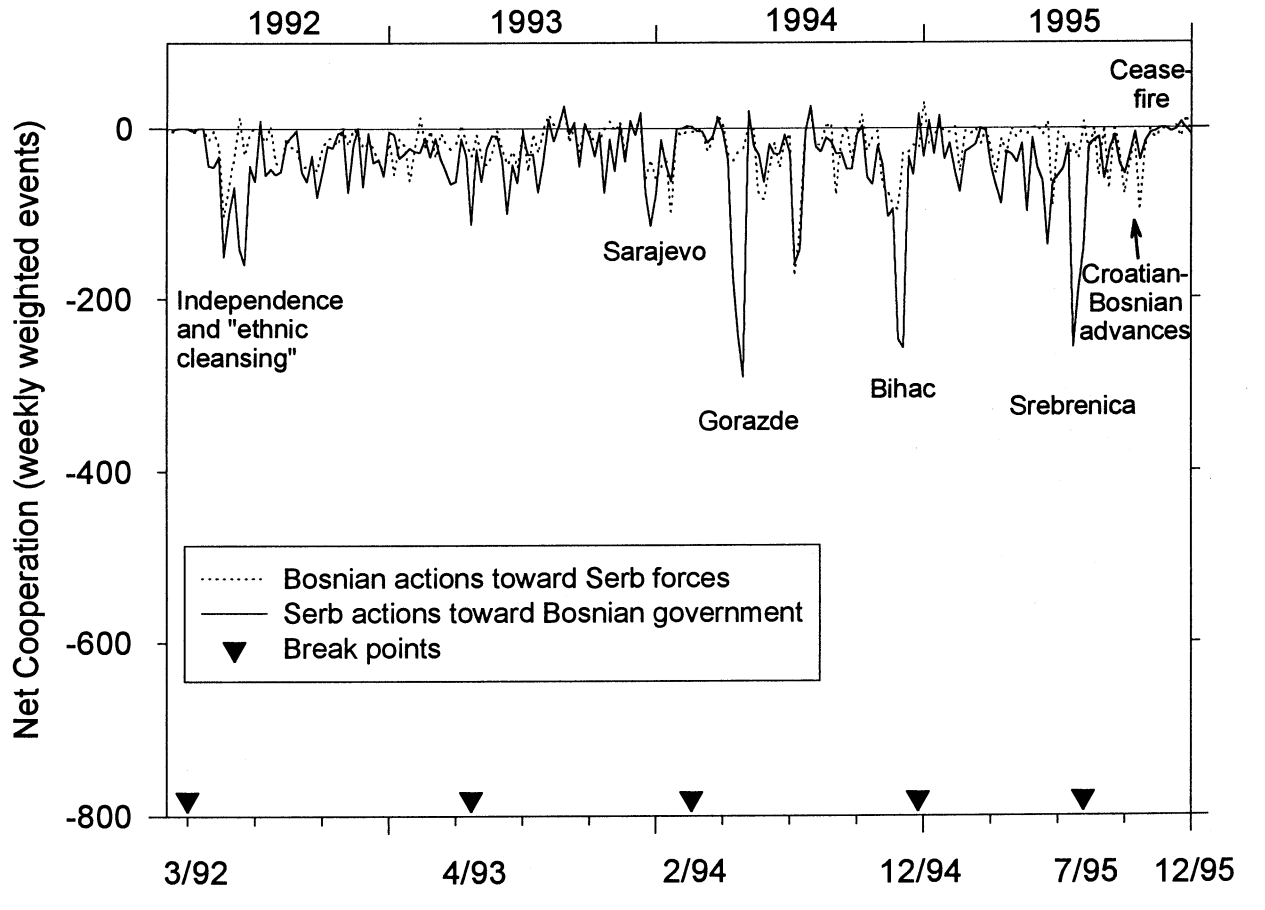
Clinton shifted toward the European position, which favored accommodation to induce Serbian cooperation.[24] Critics called the new policy "appeasement" and invoked the lessons of the 1930s (Hoffmann 1994, Herblock 1994). Jimmy Carter met with leading Bosnian Serb nationalists—formerly shunned by U.S. leaders and soon to be indicted for war crimes—to arrange a winter cease-fire. The abrupt shift in policy at the height of the Bihac crisis is evident in Figure 2; the level of net cooperation by the international community toward the Serb forces jumped from the most negative level thus far to the most positive.

The winter cease-fire gave way in spring 1995 to renewed fighting, which culminated in a short-lived effort by NATO to revive the air strike option (bombing an ammunition dump near Sarajevo). But when Serb forces responded by again taking hundreds of European UNPROFOR troops hostage, the international community backed down from further use of force and did not even retaliate when Serb forces shot down a U.S. plane. Thus, although relations deteriorated in spring 1995, we call the period one of great-power "promises" overall.

This period ended in *July 1995* when Serb forces overran the "safe area" of Srebrenica and executed thousands of the inhabitants, a signal to the great powers that current policies again were not working. Shortly thereafter, Croatian forces routed rebel Serbs from the Krajina region of Croatia, and a Croat-Bosnian offensive began to roll back Serb forces from territory in northwestern Bosnia. The U.S. Congress voted by a theoretically veto-proof majority to lift the arms embargo on Bosnia. Then, three U.S. officials—friends of President Clinton—died trying to reach besieged Sarajevo. The withdrawal of UNPROFOR in failure loomed, and Clinton had earlier promised U.S. troops to assist in that task. All these events contributed to a major shift in great-power policy, which culminated in sustained NATO air strikes in September 1995 (the largest negative spike in Figure 2). These were followed by Croatian-Bosnian territorial gains, renewed negotiations, a cease-fire (early October), and the Dayton Agreement (November). We characterize the international policy in the post-Srebrenica period overall as one of "force." We close our time series with *December 1995*, as the UNPROFOR mission ended and the NATO-led Implementation Force (IFOR) mission began.

[24] Administration officials signaled that the Bosnian government had lost the war and implied that a settlement could allow the Bosnian Serb forces to join their territories with Serbia.

FIGURE 3. Net Cooperation in Serbian-Bosnian Dyad



RESULTS

We spent most of our statistical effort on ruling out specification error in our model, since time-series statistics are especially vulnerable to such error. The first problem is the possibility of nonstationary time series. Tests showed stationarity and an absence of unit roots in each series (results are available from the authors on request; see Dickey and Fuller 1981, Hamilton 1994).

A second problem is omitted-variable bias. Plausibly, the omission of Croatian forces from our model would introduce such a bias. To test whether this was so, we estimated a model similar to Model 1, but with twelve equations and twelve independent variables in each, including all behaviors toward and by the Croatian forces.[25] We included two lagged terms (i.e., two weeks of past behavior) to reduce autocorrelation of errors, and we used an *F*-test to evaluate the joint significance of the two lagged terms for each independent variable. We estimated the model for March 1992 through December 1994, excluding the fast-moving major events of 1995 (the magnitude and possible instability of which, we feared, might wash out the effects of Croatian-related variables in earlier years).

Table 2 summarizes the results for the equations in which the four core variables of interest are dependent variables and are listed across the top. The four core independent variables are listed below, followed by the eight other variables in the model. The table shows that none of the noncore independent variables significantly affects the core variables. Therefore, we concluded that the Croatian forces could be dropped, and our model simplified to the four core variables, without introducing bias.

A third problem common to time-series analyses is omitted-lag bias. Including too few lagged terms may create serial correlation of error terms, with incorrect inferences about which variable is responsible for which effect (given contemporaneous correlations among the series, which is often the case). Including too many lagged terms will tend to reduce statistical power. Thus, we tested for the appropriate number to

[25] Clearly, the model becomes rather complex in this four-actor version. And we could have added more complexity with the inclusion of Russia or the disaggregation of Bosnian Serb forces from Serbia proper. Russia and to a lesser extent others, such as Greece, Slovenia, and the Organization of the Islamic Conference, played independent roles in the Bosnia conflict (not part of our Euro-American "international community"). Yet, these roles were secondary to, and more sporadic than, the UN/NATO roles in conflict management.

TABLE 2. Tests for Exogeneity (Omitted Variables)

Independent Variables		Probability level ($p < \ldots$) of F-statistic Dependent Variable			
		IS_t	SI_t	SB_t	BS_t
Core variables					
International toward Serb	$IS_{t-1,2}$	.39	.11	.23	.70
Serb toward International	$SI_{t-1,2}$	.27	.22	.002**	.60
Serb toward Bosnian	$SB_{t-1,2}$	.004**	.02*	.001***	.64
Bosnian toward Serb	$BS_{t-1,2}$	.79	.77	.39	.04*
Constant		.09	.001***	.001***	.001***
Other variables					
International toward Croat	$IC_{t-1,2}$	.17	.59	.30	.07
International toward Bosnian	$IB_{t-1,2}$	.39	.65	.79	.66
Serb toward Croat	$SC_{t-1,2}$	.16	.13	.08	.68
Croat toward International	$CI_{t-1,2}$	.42	.99	.75	.57
Croat toward Serb	$CS_{t-1,2}$	.51	.24	.16	.91
Croat toward Bosnian	$CB_{t-1,2}$	.57	.34	.11	.55
Bosnian toward International	$BI_{t-1,2}$	.66	.25	.62	.88
Bosnian toward Croat	$BC_{t-1,2}$	.74	.63	.15	.41
Durbin-Watson statistic		2.00	2.02	1.89	2.00

Note: Weekly data, 3/28/92–12/10/94. Number of weeks (N) = 142. Units are along net-cooperation scale. Estimation is by ordinary least squares. Probability levels shown are for joint probability of two lags of independent variable.
*$p < .05$, **$p < .01$, ***$p < .001$.

include, up to eight weekly lags.[26] No number of included lags performed significantly better than a model with just one lag. Thus, we could use Model 1, which explains each variable's behavior in a given week as a function of its own and the other variables' behavior the prior week. Needing only one lag, we can interpret coefficients straightforwardly without VAR methods.[27]

Fourth, analyses that pick a time frame to analyze arbitrarily (or based on data availability) may be misspecified if coefficients change during the course of that period. This is of special concern in a regional conflict in which patterns of behavior may plausibly change as actors learn and the war evolves. We tested for temporal stability of coefficients during March 1992 to December 1995, across the four potential break points defined above. These tests showed instability of coefficients across each of our four potential break points.[28] That is, the patterns of response in the Bosnia conflict differed significantly from one period to the next.

We found instability *within* the already short final period (July to December 1995), which included escalation, NATO air strikes, and then a cease-fire. Thus, we could not analyze the final period.[29] Furthermore, in the first period (3/92–4/93) no response coefficients in our model were significant, either because of data problems like sparser reporting in that period, or because actors were unresponsive to one another in the first year of the war. Thus, we will drop the first period from further discussion, leaving the middle three periods to analyze.

Model 1 was estimated separately for each of the three periods from 4/93 to 7/95. In only the last of these did the Bosnian government's behavior have any significant role either as cause or effect, so for the other two periods we used a simpler three-variable model (*SB*, *IS*, and *SI*).[30] Table 3 reports the response coefficients and their probability levels (as well as the constants and "self-driven" coefficients, which are theoretically uninteresting). The rows with coefficients that test for triangular responses—hypotheses 3 and 4/5—are marked by arrows at the left.

DISCUSSION

The results lend support to several of our hypotheses, although statistical significance is generally weak. First, with regard to hypothesis 1—bilateral reciprocity in the international-Serbian and Serbian-Bosnian dyads—we found little evidence of such responses in the first two periods.[31] In the promises period (12/94–7/95), however, there is evidence of bilateral reciprocity

[26] Lag tests check for any significant difference in the explanatory power of two models—one with a smaller number of lagged terms included and one with a larger number. Weekly lag tests were performed for 3/92–12/94. The lag tests use a modified likelihood ratio test (see Sims 1980, 17–8). Results are available from the authors on request.

[27] Including only one lag may leave the model vulnerable to serial correlation of errors; Durbin-Watson statistics on several equations fell below 1.9 (as low as 1.4 in the worst case). Therefore, we checked key results for robustness against a model with two weekly lags included; Durbin-Watson statistics were improved, and results were substantively similar (though with less statistical power).

[28] Two models are compared, using the same ratio test as in the lag tests. One model includes two or more subperiods; the other model blanks out the data during one period by using one dummy variable for each time point. This tests whether the dynamics of a subperiod differ significantly from the dynamics of the longer period taken as a whole.

[29] Attempts to estimate the model during the period resulted in poor Durbin-Watson statistics and other indicators that the model was misspecified.

[30] This simpler model thus has three equations (for *SB*, *IS*, and *SI*, respectively), with the same three variables (lagged) on the right-hand side of each equation.

[31] Very fast bilateral reciprocity may have existed (actions and reactions within a single week), as suggested by the high, positive

TABLE 3. Estimated Models for Three Periods

Dependent Variable	Independent Variable	4/93–2/94 (Drift) Coefficient	$p <$	2/94–12/94 (Threats) Coefficient	$p <$	12/94–7/95 (Promises) Coefficient	$p <$
Explaining International toward Serb							
IS_t	IS_{t-1}	+ 0.1	.41	− 0.0	.97	+ 0.2	.39
	SI_{t-1}	+ 0.1	.64	− 0.6	.22	− 0.0	.89
→	SB_{t-1}	+ 0.0	.86	+ 0.8	.006**	+ 0.1	.68
	BS_{t-1}					+ 0.1	.82
	Constant	−13.7	.04*	−35.3	.12	−22.0	.29
Explaining Serb toward International							
SI_t	IS_{t-1}	+ 0.0	.91	+ 0.0	.93	+ 0.5	.005**
	SI_{t-1}	− 0.2	.13	− 0.2	.32	+ 0.1	.72
	SB_{t-1}	+ 0.1	.18	+ 0.2	.05*	− 0.2	.33
	BS_{t-1}					+ 0.1	.78
	Constant	− 9.3	.02*	−21.1	.04*	−29.5	.03*
Explaining Serb toward Bosnian							
SB_t →	IS_{t-1}	− 0.4	.11	− 0.2	.07	+ 0.2	.33
	SI_{t-1}	+ 0.1	.71	+ 0.4	.12	+ 0.2	.51
	SB_{t-1}	+ 0.2	.17	+ 0.6	.001***	+ 0.2	.25
	BS_{t-1}					+ 0.5	.24
	Constant	−27.4	.003**	−27.3	.04*	−19.2	.20
Explaining Bosnian toward Serb							
BS_t	IS_{t-1}					− 0.0	.72
	SI_{t-1}					− 0.0	.67
	SB_{t-1}					+ 0.2	.004**
	BS_{t-1}					− 0.3	.14
	Constant					− 8.6	.13
Contemporaneous Correlation of Residuals							
IS_t with SI_t		.01		.62		.74	
IS_t with SB_t		.06		.43		.39	
SI_t with SB_t		−.05		.43		.11	
		[*BS* with *IS*, *SI*, and *SB*, 12/94–7/95, respectively, −.05, .06, .15]					
Durbin-Watson statistics							
Equation for IS_t		1.56		1.88		2.13	
Equation for SI_t		1.81		1.97		2.08	
Equation for SB_t		1.97		1.64		2.10	
Equation for BS_t						2.04	
Number of weeks (*N*)		43		44		32	

Note: Bosnian actions toward Serb forces (*BS*) are included only for 12/94–7/95. Arrows indicate triangular responses of Serb and international forces. Estimation is by OLS. Probability levels ($p < \ldots$) are for the *t*-statistic on each coefficient.
*$p < .05$, **$p < .01$, ***$p < .001$.

in both dyads. Serbian behavior toward the international community responds directly to international behavior toward the Serbian forces (.005 in the right-hand column of Table 3), while the Bosnian government reciprocates the actions of Serb forces toward it (.004 in that column). Thus, bilateral reciprocity seems to appear late but strong.

With regard to hypothesis 2—bilateral bullying—we found no evidence in either dyad in any period. This suggests that bullying (inverse response) may be an inherently triangular concept connected with outside powers' attempts to regulate the behavior of regional actors. Formal models which include bilateral inverse response (such as Hirshleifer and Coll 1988) might profitably be extended to explore triangular versions.

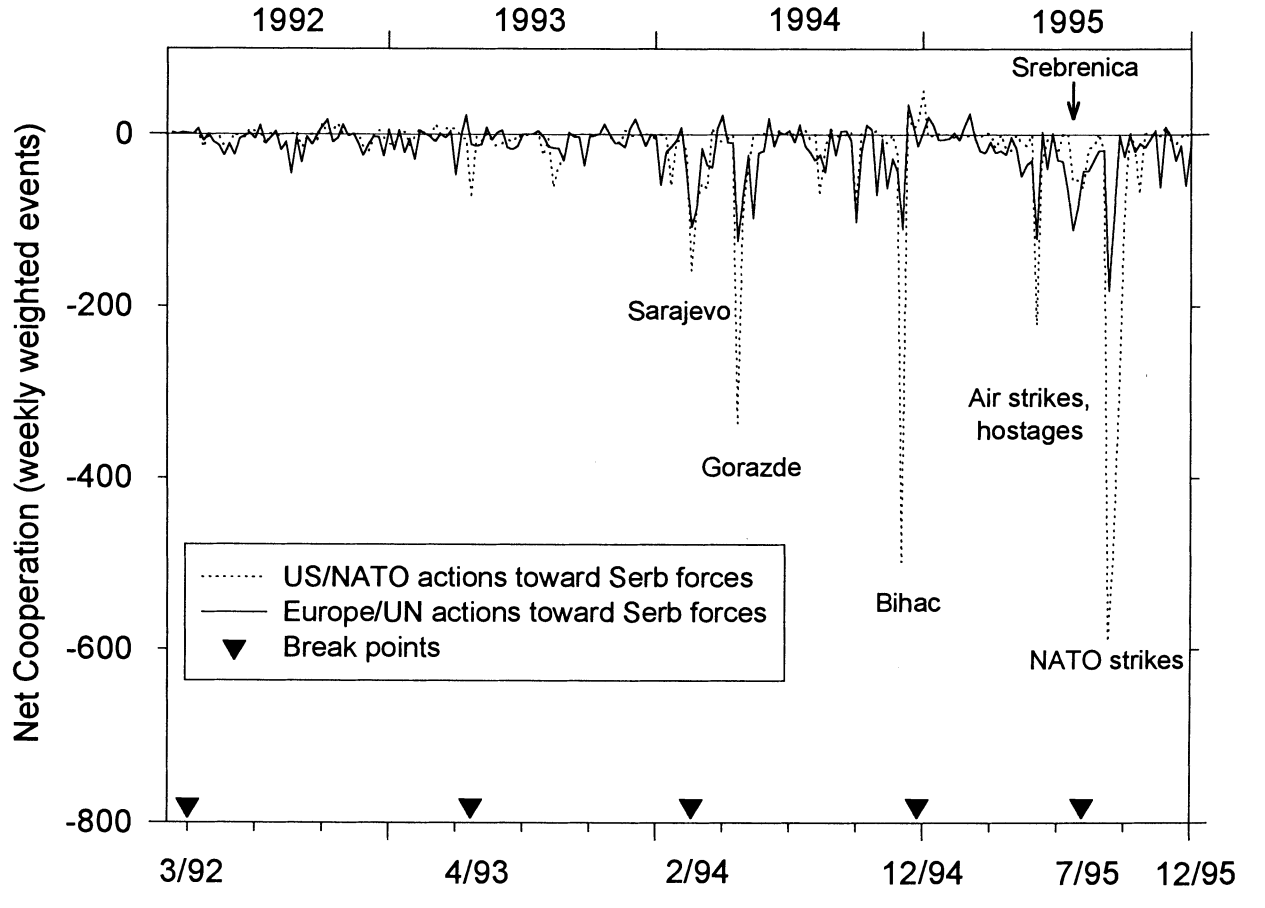
Hypothesis 3—triangular response by the international community to Serb actions toward Bosnia—corresponds with the third row of Table 3 (marked by an arrow). In the threats period, after February 1994, the significant positive coefficient ($p < .006$) indicates that Serbian attacks on (or restraint toward) Bosnia were followed by like actions by the international community toward Serb forces. Interestingly, recent Serbian behavior toward Bosnia was a far better predictor of international behavior toward Serb forces than was either the international community's own recent behavior or the behavior it just received from Serb forces. Before February 1994, hypothesis 3 is not significant ($p < .86$). And as the international community backed off in the promises period after December 1994, the international responses to Serb actions toward Bosnia seemed to weaken again ($p < .68$).

Hypothesis 4—triangular Serbian reciprocity—received no support in these results.

Hypothesis 5—triangular Serbian bullying of Bosnia—corresponds with the second row marked by an arrow in Table 3. For the 4/93–2/94 period (drift), it is the only response anywhere close to statistically

contemporaneous correlation of residuals (.62) for *SI* and *IS* in the 2/94–12/94 threats period.

FIGURE 4. Net Cooperation toward Serb Forces by United States/NATO and by Europe/UN



significant ($p < .11$). It is nearly significant ($p < .07$) for 2/92–12/94 (threats), when international efforts to manage the conflict were most active. In both these periods the negative sign of the coefficient indicates the responses were inverse, not reciprocal. For 12/94–7/95 (promises), the response is not significant; indeed, all the triangular relationships seem to disappear in that period. But that picture turns out to be inaccurate.

FURTHER RESULTS AND DISCUSSION

We disaggregated the international actors, distinguishing European and UN targets and actions (which we call *E*, for Europe), from American and NATO ones (which we call *A*).[32] We split them this way because actions by and toward the UN in Bosnia generally involved UNPROFOR troops from the European countries; actions by and toward NATO typically involved U.S. airplanes and U.S. political initiatives. Our model then had six variables instead of four (*IS* and *SI* are each split). Figure 4 shows the time series for European and American actions toward the Serb forces. Generally, the two series moved in tandem, though NATO took far more conflictual actions in crises (using military force).[33] This led us to expect no difference in response patterns when we disaggregated the international community. Furthermore, disaggregation did not create substantial new results in the period before February 1994.

In the important threats period of 1994, disaggregation produced only one interesting result: The responses found between *IS* and *SB* involved U.S./NATO actions and responses, while responses found between *SI* and *SB* involved European/UN ones. That is, the international actions *toward* Serb forces that mattered were those of the United States and NATO, while the Serb forces' own actions toward the international community that mattered were those toward Europe and UNPROFOR. These results elaborate those presented

[32] In separate preliminary analyses, we tried disaggregating Serbian variables into those of or toward Serbia proper (Belgrade) and those of or toward the Bosnian Serb forces (Pale). For 3/92–10/94, international behaviors toward Belgrade and Pale, respectively, were positively cross-correlated at .63, but Belgrade and Pale's behaviors toward the international community were at only .19. Reestimating our model with the two variables replacing Serb forces (both as actors and targets), we did not find substantively meaningful differences in response patterns.

[33] In the periods before February 1994, unlagged cross-correlations (between *AS* and *ES*, and between *SA* and *SE*, respectively) are small, below .22. But in subsequent periods they jump to around .6 to .7 positive correlation.

TABLE 4. Estimated Model for 12/94–7/95 with "International" Disaggregated

Independent Variables	Dependent Variable: America/NATO toward Serb AS_t Coefficient	$p <$	Europe/UN toward Serb ES_t Coefficient	$p <$	Serb toward America/NATO SA_t Coefficient	$p <$	Serb toward Europe/UN SE_t Coefficient	$p <$	Serb toward Bosnian SB_t Coefficient	$p <$	Bosnian toward Serb BS_t Coefficient	$p <$
AS_{t-1}	− 0.4	.26	− 0.7	.002**	+ 0.2	.11	− 0.2	.50	− 0.9	.006**	− 0.1	.71
ES_{t-1}	+ 0.9	.03*	+ 1.1	.001***	+ 0.4	.02*	+ 0.8	.02*	+ 1.7	.001***	+ 0.0	.97
SA_{t-1}	+ 0.8	.09	+ 0.7	.04*	+ 0.5	.008**	+ 0.6	.08	+ 1.5	.002**	− 0.1	.66
SE_{t-1}	− 0.5	.14	− 0.1	.74	− 0.2	.12	− 0.3	.28	− 0.3	.31	+ 0.0	.94
SB_{t-1}	− 0.2	.25	− 0.2	.18	− 0.2	.006**	− 0.2	.19	− 0.2	.30	+ 0.2	.02*
BS_{t-1}	+ 0.2	.53	+ 0.4	.10	+ 0.1	.46	+ 0.3	.25	+ 1.0	.02*	− 0.3	.15
Constant	−11.9	.33	− 6.3	.43	− 9.6	.05*	−22.8	.02*	−16.7	.17	− 7.7	.21
Durbin-Watson statistic	2.14		1.83		1.92		1.96		1.93		2.03	

Note: Number of weeks (N) = 32. Estimation is by ordinary least squares. Boxed coefficients are for Serb triangular response to international actions. Probability levels ($p < \ldots$) are for the t-statistic on each coefficient.
*$p < .05$, **$p < .01$, ***$p < .001$.

earlier (with the international community aggregated) but do not substantially change them. They do suggest, consistent with hegemonic stability theory, that Serb forces were more responsive to American actions than European ones.

The most interesting effects of disaggregation came in the promises period, 12/94–7/95. Table 4 summarizes the results. Although it had appeared (refer to Table 3) that Serb forces no longer displayed a triangular bullying response to international actions in this period, the actual responses were highly significant, but with opposite signs for the United States and Europe (see the boxed coefficients in Table 4 with $p < .006$ and .001, respectively). The inverse response pattern (triangular bullying) found in the two previous periods continues to be found and continues to apply specifically to U.S./NATO actions (as it did in the previous period). But this response is now supplemented by a triangular *reciprocal* response to European/UN actions—the only time Serb forces showed such a response and hence the only bit of evidence in this study supporting hypothesis 4.

In this disaggregated model for 12/94–7/95, bilateral reciprocity appears relatively strong and widespread in both dyads.[34] Thus, this one model for one period seems to contain bilateral reciprocity, triangular bullying (toward Bosnia) in the Serbian response to America, and triangular reciprocity (toward Bosnia) in the Serbian response to Europe—but no triangular response by the international community. The results for this period with the international community disaggregated, however, should be interpreted cautiously, since there are fewer degrees of freedom than for the results presented earlier.[35]

Across all three periods, the hypothesis most strongly supported by this study is hypothesis 5—Serbian triangular bullying. It was close to statistical significance in two periods and strongly significant in the third with regard to U.S./NATO actions when disaggregated from European/UN actions. Combining these three results gives an overall significance for this hypothesis at $p < .01$.[36] The results thus support the assumptions of the aggressor-victim school of thought that the international use of force could induce Serbian cooperation in this regional conflict. The containment interpretation of the events of fall 1995—that robust NATO air strikes finally caused the Serb forces to cooperate with the Bosnian government—resonates with our results.

By contrast, the warring-factions school of thought, with its preferred policy of using international cooperation to elicit Serbian cooperation toward Bosnia, receives little support. The results, for example, appear to contradict those who argued that air strikes would inflame Serbian hostility toward Bosnia.[37] The one supportive finding for this school is the apparent response of Serbian behavior (toward Bosnia) to *European* actions in the first half of 1995. Here, we find the triangular reciprocity posited by hypothesis 4, but it coexists with the triangular bullying of hypothesis 5 in response to U.S./NATO actions (and at a time when

[34] Across ten possible variations of bilateral reciprocity on the international-Serbian and Serbian-Bosnian dyads, five are statistically significant. *BS* (top right of Table 4) responds reciprocally to *SB* (fifth line down), $p < .02$, and vice versa (also $p < .02$). *AS* responds reciprocally to *SA* ($p < .09$) but not *SE* (negative sign, $p < .14$), while *SA* responds reciprocally to *AS* ($p < .11$) and *ES* ($p < .02$). *ES* does not respond to *SE* ($p < .74$) but does respond reciprocally to *SA* ($p < .04$), while *SE* reciprocates *ES* ($p < .02$) but not *AS* ($p < .50$).

[35] The model now includes six variables and only 32 weeks of data. To test robustness, however, we reran this model with two weekly lags instead of one, and the response of *SB* was still negative to *AS* and positive to *ES*.

[36] We use Fisher's (1950, 100) method of doubling the sum of $-\ln(p)$, where p is the probability level in each test, to estimate chi-square with degrees of freedom equal to twice the number of tests. Here, p levels of .11, .07, and .006 give chi-square (6 DF) = 19.97 ($p < .01$).

[37] Boyd (1995, 37–8), for example, wrote just before the NATO air strikes in September 1995 that they "can only reinforce the paranoia that drives [Serb forces] to continue the fight so relentlessly."

U.S./NATO actions had become more salient than European/UN ones).

The divergence in European and American roles in the conflict is interesting. Apparently, by early 1995 the United States used "sticks" and Europe "carrots" to get Serb forces to cooperate. The results give some support to the idea that U.S. participation (as a hegemon) is crucial to the effective management of a major regional conflict—even in Europe. The results also suggest that a regional belligerent may try to play great powers off against one another and/or that the great powers may use a "good cop, bad cop" strategy.

CONCLUSION

Returning to the theories of cooperation promoted by neoliberal institutionalism, this study raises three main challenges. First, research on bilateral relationships—whether by formal modeling, statistical analysis, or case studies—may miss triangular relationships that are central to the management of regional conflicts. Since such conflicts are increasingly important in the post-Cold-War era, theorists of cooperation should give far more attention to triangularity as a context for the evolution of cooperation.

Second, the study of reciprocity as a strategy for eliciting cooperation has, to date, largely ignored strategies of inverse response (bullying). We have shown that inverse response is not just an abstract possibility; it characterizes the actual responses of the key actor in one of the most important regional conflicts of recent years. Clearly, such "nice" strategies as Axelrod's (1984) tit for tat or Osgood's (1962) GRIT do not work well against bullies. In eliciting cooperation from a bully, the relatively cheap options of the accommodation approach, such as mediation and peacekeeping, are unlikely to succeed. Thus, inverse response deserves far more attention from theorists and empirical analysts than it has received to date. Models of inverse response should, of course, reflect the triangular context of bullying, not the abstract notion of bilateral bullying (for which we found no empirical evidence). Theories of the evolution of cooperation under anarchy ideally should allow for our empirical finding that bilateral reciprocity can exist simultaneously with triangular bullying.

Third, if responding to regional conflicts creates collective goods dilemmas among outside powers, these dilemmas surely intensify when accommodation works poorly and containment becomes necessary. Regional conflicts involving bullies are most likely to exact high costs, to become dangerous precedents if left unsolved, and to create potential divisions among great powers about who should respond and how. Our results suggest that the participation of the world's hegemonic military power—the United States—was central to the containment of Serbian behavior toward Bosnia. For theorists of cooperation, then, propositions about hegemonic leadership should be incorporated into the formal and empirical study of cooperation, alongside the elements of triangularity and inverse response.

Methodologically, this research contributes to the study of regional conflict in several ways. In applying KEDS to a new regional conflict (beyond the Middle East, which was used in developing KEDS), we have provided additional evidence that: (1) intensive news reporting, such as that by the Reuters wire service, can serve as raw material for constructing events data; (2) machine-coding, with its obvious advantages in terms of cost and bias reduction, is a practical method for generating data from such textual sources; and (3) time-series analysis based on lagged correlations of movements along a net-cooperation scale can illuminate the dynamics of regional conflict. These approaches should allow future researchers and policymakers to understand the dynamics of international conflicts better and sooner.

We plan to analyze a series of regional conflicts in the 1990s, which should increase the statistical power of our tests.[38] Analysis of multiple cases may also illuminate two tantalizing suggestions from the present findings. The first is the idea that bilateral reciprocity (tit for tat) becomes stronger as a conflict persists. In our study, the bilateral reciprocity hypothesis received support only in the 12/94–7/95 period (the last year of the war). One interpretation is that parties learn to reciprocate from the experience of repeated interaction. An alternative interpretation is that a relative equality of power (which had developed in Bosnia by 1995) may contribute to the emergence of bilateral reciprocity.[39] Such power-balanced contexts, after all, most resemble the formal models of "evolution of cooperation" as well as great-power relations (both of which have proven conducive to reciprocal strategies). We cannot say in light of a single case which interpretation is correct. The second tantalizing idea, similarly indeterminate in a single case study, is that the United States as hegemon plays a special role in regional conflict management—a role which at times may run counter to that of other great powers. Analysis of multiple cases may allow exploration of that possibility.

The war in Bosnia took a terrible toll, both on the people of that small country and on the international institutions whose intervention efforts were so problematical. Nothing can change that history. But by better understanding what happened, and specifically by bringing the tools of social science to bear (in addition to the more common methods of journalism and history), we hope to contribute to the better management of future regional conflicts. To the extent that the models of cooperation studied by scholars of international relations can correspond more closely with the realities faced by policymakers, those models will become more useful.

[38] NSF grant SBR-9617157 (1997–98) will fund analysis of about a dozen regional conflicts that occurred in 1990–96.

[39] By early 1995 the Bosnian government had built up its army sufficiently to give Serb forces a real fight, in contrast to the earlier extreme imbalance of power. At the same time, the international community's lack of will seemed to bring its power down toward the level of the Bosnian Serbs.

REFERENCES

Andriole, S. J., and G. W. Hopple. 1984. "The Rise and Fall of Event Data: From Basic Research to Applied Use in the U.S. Department of Defense." *International Interactions* 10(3/4):293–309.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, Robert, and Robert O. Keohane. 1986. "Achieving Cooperation under Anarchy: Strategies and Institutions." In *Cooperation Under Anarchy*, ed. Kenneth A. Oye. Princeton: Princeton University Press.

Azar, Edward E., R. A. Brody, and Charles A. McClelland, eds. 1972. *International Events Interaction Analysis: Some Research Considerations*. Beverly Hills, CA: Sage.

Baldwin, David A., ed. 1993. *Neorealism and Neoliberalism: The Contemporary Debate*. New York: Columbia University Press.

Boyd, Charles G. 1995. "Making Peace with the Guilty: The Truth about Bosnia." *Foreign Affairs* 74(September/October):22–38.

Brinkley, Douglas. 1995. "Jimmy Carter's Modest Quest for Global Peace." *Foreign Affairs* 74(November/December):90–100.

Dickey, D. A., and W. A. Fuller. 1981. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica* 49(June):1057–72.

Donia, Robert J., and John V. A. Fine, Jr. 1994. *Bosnia and Hercegovina: A Tradition Betrayed*. New York: Columbia University Press.

Fisher, R. A. 1950. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.

Freeman, John R. 1990. "Systematic Sampling, Temporal Aggregation and the Study of Political Relationships." In *Political Analysis*, ed. James A. Stimson. Ann Arbor: University of Michigan Press.

Freeman, John R., John T. Williams, and Tse-min Lin. 1989. "Vector Autoregression and the Study of Politics." *American Journal of Political Science* 33(November):842–77.

George, Alexander L., Philip J. Farley, and Alexander Dallin, eds. 1988. *U.S–Soviet Security Cooperation: Achievements, Failures, Lessons*. New York: Oxford University Press.

Gerner, Deborah J., Philip A. Schrodt, Ronald Francisco, and Judith L. Weddle. 1994. "The Analysis of Political Events Using Machine Coded Data." *International Studies Quarterly* 38(March):91–119.

Goldstein, Joshua S. 1991. "Reciprocity in Superpower Relations: An Empirical Analysis." *International Studies Quarterly* 35(June): 195–209.

Goldstein, Joshua S. 1992. "A Conflict-Cooperation Scale for WEIS International Events Data." *Journal of Conflict Resolution* 36 (June):369–85.

Goldstein, Joshua S. 1995. "Great Power Cooperation under Conditions of Limited Reciprocity: From Empirical to Formal Analysis." *International Studies Quarterly* 39(December):453–78.

Goldstein, Joshua S., and John R. Freeman. 1990. *Three-Way Street: Strategic Reciprocity in World Politics*. Chicago: University of Chicago Press.

Goldstein, Joshua S., and John R. Freeman. 1991. "US–Soviet–Chinese Relations: Routine, Reciprocity, or Rational Expectations?" *American Political Science Review* 85(March):17–35.

Hamilton, James D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

Herblock. 1994. Editorial cartoon. *Washington Post*, December 11.

Hirshleifer, Jack, and Juan C. M. Coll. 1988. "What Strategies Can Support the Evolutionary Emergence of Cooperation?" *Journal of Conflict Resolution* 32(June):367–98.

Hoffmann, Stanley. 1994. "What Will Satisfy Serbia's Nationalists?" *New York Times* December 4, Section E.

Howell, Llewlyn, and Gillian Barnes. 1993. "Event Data for Region Specific Interactions: A Research Note on Source Coverage." In *International Event Data Developments: DDIR Phase II*, ed. Richard L. Merritt, Robert G. Muncaster, and Dina A. Zinnes. Ann Arbor: University of Michigan Press.

Huxtable, Phillip A., and Jon C. Pevehouse. 1996. "Validity Problems in Events Data Collection: News Media Sources and Machine Coding Protocols." *International Studies Notes* 21(Spring):8–19.

Jervis, Robert. 1988. "Realism, Game Theory, and Cooperation." *World Politics* 40(April):317–49.

Kaplan, Robert D. 1993. *Balkan Ghosts: A Journey Through History*. New York: St. Martin's Press.

Kelly, Michael. 1994. "Surrender and Blame." *New Yorker*, December 19, pp. 44–51.

Keohane, Robert O. 1986a. "Reciprocity in International Relations." *International Organization* 40(Winter):1–27.

Keohane, Robert O. 1986b. "Neorealism and World Politics." In *Neorealism and Its Critics*, ed. Robert O. Keohane. New York: Columbia University Press.

Laurance, E. J. 1990. "Events Data and Policy Analysis: Improving the Potential for Applying Academic Research to Foreign and Defense Policy Problems." *Policy Sciences* 23(2):111–32.

Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton: Princeton University Press.

Malcolm, Noel. 1994. *Bosnia: A Short History*. New York: New York University Press.

McClelland, Charles. 1978. *World Event/Interaction Survey (WEIS) Project, 1966–1978* [computer file] (Study #5211). Conducted by Charles McClelland, University of Southern California. 3rd ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1978.

Mearsheimer, John J., and Stephen Van Evera. 1996. "Hateful Neighbors." *New York Times*, September 24, Section A.

Osgood, Charles E. 1962. *An Alternative to War or Surrender*. Urbana: University of Illinois Press.

Patchen, Martin. 1991. "Conflict and Cooperation in American-Soviet Relations: What Have We Learned from Quantitative Research?" *International Interactions* 17(2):127–43.

Schrodt, Philip A. 1995. "KEDS: Kansas Event Data System Manual." University of Kansas. Typescript.

Schrodt, Philip A., Shannon G. Davis, and Judith L. Weddle. 1994. "Political Science: KEDS—A Program for Machine Coding Events Data." *Social Science Computer Review* 12(Fall):561–88.

Schrodt, Philip A., and Deborah J. Gerner. 1994. "Validity Assessment of a Machine-Coded Event Data Set for the Middle East, 1982–1992." *American Journal of Political Science* 38(August):825–54.

Sims, Christopher A. 1980. "Macroeconomics and Reality." *Econometrica* 48(January):1–48.

Snyder, Glenn H., and Paul Diesing. 1977. *Conflict among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton: Princeton University Press.

Tyndall Report. 1995. *The Tyndall Report 1994 Year in Review*. New York: ADT Research.

Washington Post. 1995. "The Bosnia Humiliation." Editorial, May 30, Section A.

Woodward, Susan L. 1995. *Balkan Tragedy: Chaos and Dissolution after the Cold War*. Washington, DC: Brookings.

Zinnes, Dina A. 1980. "Three Puzzles in Search of a Researcher." *International Studies Quarterly* 24(September):315–42.