An Event Data Analysis of Third-Party Mediation in the Middle East and Balkans

PHILIP A. SCHRODT
DEBORAH J. GERNER
Department of Political Science
University of Kansas

Event data on the Israel-Lebanon and Israel-Palestinian conflicts in the Levant (1979-1999) and the Serbia-Croatia and Serbia-Bosnia conflicts in the Balkans (1991-1999) are used to test two sets of process-related hypotheses embedded in the theoretical and qualitative literatures on mediation. Cross-correlation analysis is used to examine the time delay in the effects of mediation on the level of violence over time. Results show that these effects vary somewhat, depending on the conflict, and differ substantially, depending on who is conducting the mediation. Whether conflict reduction is most effectively achieved with sanctions or rewards is tested. Results show that a reduction in violence is generally associated with mediation combined with conflictual action directed toward both of the antagonists and combined with cooperative action directed to the weaker antagonist.

Keywords: mediation; sanctions; rewards; event data; cross-correlation; Middle East; Balkans

This study uses statistical time-series analyses of political event data to examine the dynamics of third-party international mediation on the Israel-Lebanon and Israel-Palestinian conflicts in the Levant and the Serbia-Bosnia and Serbia-Croatia conflicts in the Balkans. Event data—nominal or ordinal codes that record the interactions between international actors as reported in the open press—provide a rich set of indicators about the results of mediation, the political circumstances of the mediation (e.g., prior military success or failure by the protagonists), and the various strategies employed by the mediating parties.

Our analysis shifts from the generally structural focus of quantitative mediation literature to an emphasis on the dynamics of the mediation process. We first introduce results of cross-correlation analyses, which we undertake to identify the lag structure between mediation and levels of conflict-cooperation between antagonists. We then present a time series analysis that directly tests the "sticks-or-carrots" model: is media-
tion more effective when it is accompanied by negative sanctions or by positive incentives?

The variables we examine are generally referred to as process variables in the qualitative mediation literature, although we will also be looking at some dynamic variables that are usually put in the contextual category in that literature. For example, the relationship between the mediator and a disputant is generally considered a contextual variable, but it can change at critical moments, as with the December 1988 decision by the United States to deal directly with the Palestine Liberation Organization (Gerner and Wilbur 2000). Thus, this research fills a gap in the literature between the macrolevel variables emphasized in the existing quantitative studies (reviewed below) and the microlevel advice to individual negotiators that is found in the “wisdom literature” (e.g., Fisher and Ury 1978; Fisher et al. 1997) and in descriptive case studies.

We are primarily interested in how variables that change throughout the process of mediation affect the conflict. However, the use of separate time series analysis for different cases of conflict and mediation also allows us to compare this pattern across cases that are differentiated by major contextual differences. One obvious variation in our study is the prior relationship between the mediator and the fighting parties: the special relationship between the United States and Israel in the Levant is not paralleled in any of the cases in the Balkans. We expect such factors to affect the mediation pattern and the sticks-or-carrots effects for the various conflicts.

RESEARCH DESIGN

Most of the existing quantitative research on mediation has used a cross-sectional approach that considers the effects of a set of specific variables across a large number of cases, an approach that dates back to the 1960s. The initial work was done by Haas (1967), who focused specifically on the efforts of international organizations to control conflict through mediation and other active measures, such as collective security. This work was later extended by Nye (1968) and Butterworth (Haas, Butterworth, and Nye 1972; Butterworth and Scranton 1976); the Butterworth studies also included mediation efforts by individual nation-states and organizations not set up for collective security.

During the past 15 years, the most extensive quantitative analysis of mediation has been in the work of Bercovitch and his associates (Bercovitch, Anagnoson, and Wille 1991; Bercovitch and Wells 1993; Bercovitch 1996; Bercovitch and Houston 1996, 2000; Bercovitch and Schneider 2000). Bercovitch has assembled a data set on mediation efforts for 295 conflicts from 1945 to 1995, and used statistical methods to test a variety of hypotheses about mediation.

Our research design departs from these studies in two respects. First, we are looking at cases of mediation individually rather than combining them into a single data set that spans many disparate cases. This approach has been employed in several more recently published studies based on smaller data sets that focus on a limited number of crises. For example, Ayres (1997) looks at the quantitative dynamics of image change in three conflicts; Mooradian and Druckman (1999) do the same for six mediation
attempts in the Nagorno-Karabakh dispute. Carment and Rowlands (1998) develop a game theoretic model based on salience, intensity, capability, and expected gains for the belligerent, and assess the model on six crises from the 1990s.

The decision to use a time-series case study approach rather than a cross-sectional design was motivated by three considerations. First, the thorough reporting on the conflicts in the Middle East and the Balkans provides us with data sets that contain tens of thousands of events, so a time-series analysis of each separate conflict is feasible; we do not need to combine cases to obtain adequate degrees of freedom. Second, pooling cases for cross-sectional time series analysis would create problems of heterogeneity, because it is difficult to simultaneously control for serially and cross-sectionally correlated errors. Finally, we are focusing on a limited number of regions and a relatively limited period of time to control, at least in part, for the effects of cultural and historical context.

The second major difference between our work and most of the existing studies is that we use data coded with the general-purpose World Events Interaction Survey (WEIS) coding scheme (McClelland 1976) and derive our measure of mediation through patterns of events rather than explicitly coding for mediation. The same data set is also used to assess the level of conflict and the levels of incentives and sanctions applied by the mediator. This goes back to McClelland’s (1970) original assumption that event data would break down complex political activities into the sequence of basic building blocks (e.g., comments, visits, grants, rewards, protests, demands, threats, and military engagements) from which more complex political activities can be constructed.

We define a mediation pattern as the occurrence of a cooperative behavior (WEIS cue categories 01 through 10) between a designated mediator (e.g., the United States, United Nations [UN], or European Union [EU]) and both parties in the dyad within a period of 7 days.1 (See Appendix A for a list of WEIS cue categories and the average Goldstein Scale scores within each category.2) Although this pattern is not a sufficient condition for mediation—a representative of a state might visit multiple parties to a conflict without trying to mediate—it is probably a necessary condition (any mediation will involve such cooperation, at least within the limitation of the news reports). This definition has been implemented as a relatively simple C program that operates on the original stream of event data, detects occurrences of the pattern, and then produces a new file that can be merged with other variables.

We understand mediation to be a tool of conflict management or resolution, involving third-party intervention; it is neither forceful nor binding, unless negotiating parties choose to make it so. Although the way we operationalize mediation relies on cooperative behavior between third parties and parties to a conflict, we do not intend to preclude from our definition those events that involve intervention of biased or impartial third parties. We do believe that even adversarial and manipulative mediation strategies (Bercovitch and Wells 1993) would typically be accompanied by

1. We did a few tests using an interval of 4 days; this made no discernible difference in the results.
2. For purposes of brevity, Appendix A gives only the average Goldstein value for each two-digit cue category. All calculations of Goldstein totals, however, used the Goldstein scores assigned to the original three-digit World Events Interaction Survey (WEIS) categories.
some form of consultation, if not a more cooperative interaction, between mediators and antagonists.

Figures 1 and 2 show both the Goldstein-scaled (Goldstein 1992) monthly aggregations of events of Israel to the Palestinians and Israel to Lebanon as well as the frequency of mediation events. This measure of mediation activity tracks the historical record fairly well. The Israel-Palestinian dyad receives mediation efforts almost continuously except during the period from 1983 to 1988, with conspicuous spikes corresponding to events such as the 1982 Lebanon invasion, U.S. resumption of formal negotiations with the Palestine Liberation Organization (PLO) in 1988, and various
agreements that were part of the Oslo process. In contrast, mediation in Lebanon tends—necessarily but not sufficiently—to coincide with periods of violence.

Figure 3 shows the total number of mediation events in the Balkans mediation by 6-month period for three different groups of mediators. The bars labeled “UN,” “USA,” and “EUR” count the number of mediation events that involved the UN, United States, and major European states (plus the EU), respectively; the dates indicate the month when the 6-month period begins. The general pattern shows more or less continuous—but relatively low-level—mediation efforts by the UN and the Europeans starting around mid-1992. The United States, in contrast, engages in only sporadic efforts until 1995, when it engages in a very substantial effort that dwarfs the earlier UN and European efforts and culminates in the Dayton Accords. Following that effort mediation, efforts largely disappear from the event data record. As with the Middle East case, the mediation effort, as measured by the proposed pattern, appears to accurately track the qualitative history of the conflict.

ANALYSIS

DATA

The data used in this study were coded into the WEIS scheme using the Kansas Event Data System (KEDS), a computer program that creates event data from machine-readable text. The events were coded from Reuters News Service lead sentences obtained from the NEXIS data service for the period from April 1979 through

3. Discussions of machine coding can be found in Gerner et al. (1994), Schrodt and Gerner (1994), Huxtable and Pevehouse (1996), and Bond et al. (1997).

We analyze Goldstein-scaled aggregate measures of conflict and event counts of the number of events that occur in a dyad. The scaled conflict variable, scaled conflict, uses the negative of the Goldstein Scale at monthly aggregations—the scaled values for the events are totaled for each dyad-month. For the event counts, we use the following categories based on the WEIS two-digit cue categories:

- Verbal cooperation: WEIS categories 02, 03, 04, 05, 08, 09, 10
- Material cooperation: WEIS categories 01, 06, 07
- Verbal conflict: WEIS categories 11, 12, 13, 14, 15, 16, 17
- Material conflict: WEIS categories 18, 19, 20, 21, 22

This reduces the number of distinct event categories that can be used as independent variables to a manageable number and eliminates the problem of three-digit WEIS categories that have very low frequencies. It is also likely to reduce the effects of coding error somewhat. Several of the verbal conflict codes in WEIS are ambiguous even for human coders, and the automated coding probably generates some misclassification within those categories.

We consider 10 different cases of third-party mediation:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Actor A</th>
<th>Actor B</th>
<th>Mediators</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISRLEB</td>
<td>Israel</td>
<td>Lebanon</td>
<td>United States, EU, UN</td>
</tr>
<tr>
<td>ISRPAL</td>
<td>Israel</td>
<td>Palestinians</td>
<td>United States, EU, UN</td>
</tr>
<tr>
<td>ALLBFR</td>
<td>Serbia</td>
<td>Bosnia</td>
<td>All mediators listed below</td>
</tr>
<tr>
<td>USABFR</td>
<td>Serbia</td>
<td>Bosnia</td>
<td>United States</td>
</tr>
<tr>
<td>EURBFR</td>
<td>Serbia</td>
<td>Bosnia</td>
<td>EU, France, Germany, Italy, United Kingdom</td>
</tr>
<tr>
<td>UNOBFR</td>
<td>Serbia</td>
<td>Bosnia</td>
<td>UN</td>
</tr>
<tr>
<td>ALLCRO</td>
<td>Serbia</td>
<td>Croatia</td>
<td>Same sets of mediators as Bosnia cases</td>
</tr>
<tr>
<td>USACRO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EURCRO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNOCRO</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**CROSS-CORRELATION**

Our first set of tests involves cross-correlation of the mediation measure with the total level of scaled conflict, defined as the negative of the sum of the Goldstein-scaled net-cooperation scores A-to-B and B-to-A (i.e., high values imply high levels of conflict). The cross-correlation approach is discussed in greater detail in Appendix B.

A cross-correlation analysis is useful because the mediation literature allows for the possibility that there will be some lag between use of mediation and the impact of

4. Note that this variable does not refer to the Goldstein-scaled, net-cooperation score, which is the typically used aggregate score; the variable used in this study gives positive weight to conflict and not cooperation.
that mediation on the level of conflict. Furthermore, this lag could be anywhere from a few weeks to a few months, and it might also differ between regions. For example, although some agreements—notably cease-fires—are supposed to be implemented immediately, others (e.g., disarmament, territorial disengagement, and deployment of peacekeeping forces) involve substantial negotiated delays of varying lengths. In addition, much of the qualitative literature asserts that de-escalation processes require time to take hold due to continued hostilities by groups opposed to the peace process, warfare by the population that hostilities have actually ended, gradual repatriation of refugees and reconstruction of infrastructure, and other time-consuming processes.

The problem of time lags is further complicated by the presence of a strong, but theoretically plausible, positive contemporaneous correlation between mediation and violence. In the post–World War II period, outbreaks of violence invoke almost immediate attempts at mediation. Schrodt’s (1990) sequence-recognition exercise using the Behavioral Correlates of War (BCOW) data set (Leng 1987) shows that mediation was the primary behavior distinguishing pre– and post–World War II crises. Many aspects of mediation have a strong stochastic element. Consequently, we would expect to see the effects of mediation spread across several time lags.

In the cross-correlograms in Figures 4, 5, and 6, the values to the left of the lag of 0 (the center of the graph) are the correlations with mediation activity and conflict between the antagonists prior to the mediation; the values to the right of the 0 lag are the correlations with mediation activity and conflict following the mediation. If mediation is successful at reducing conflict, we would expect to see a negative correlation between mediation events at time $t$ and scaled conflict at time $t + k$ in these...
Figure 5: Cross-Correlations of Mediation and Serbia-Bosnia Conflict
NOTE: USABFR = conflict between Serbia and Bosnia, mediated by the United States; EURBFR = conflict between Serbia and Bosnia, mediated by the European Union, France, Germany, Italy, and the United Kingdom; UNOBFR = conflict between Serbia and Bosnia, mediated by the United Nations.

The parallel dotted lines on the correlograms show the critical values at the two-tailed 5% significance level; these were determined by Monte Carlo simulation (see Appendix B).

Figure 6: Cross-Correlations of Mediation and Serbia-Croatia Conflict
NOTE: USACRO = conflict between Serbia and Croatia, mediated by the United States; EURCRO = conflict between Serbia and Croatia, mediated by the European Union, France, Germany, Italy, and the United Kingdom; UNOCRO = conflict between Serbia and Croatia, mediated by the United Nations.

The parallel dotted lines on the correlograms show the critical values at the two-tailed 5% significance level; these were determined by Monte Carlo simulation (see Appendix B).
In the analysis of the Levant case, we looked at mediation efforts involving either the United States, UN, or European Community/EU. Most of this mediation activity, unsurprisingly, involves the United States: of the 95,464 events in the data set, 22,752 (23.8%) involve the United States as actor or target, 6,186 (6.5%) involve the UN, and only 579 (0.6%) involve the European Community or EU. Because we looked at interactions that involved any of these actors, a meeting between UN officials with Palestinians followed 5 days later by a meeting between U.S. officials and Israelis would count as a mediation effort. This is imprecise but is probably still a reasonable approximation. UN involvement is far more likely in the Israel-Lebanon case than in the Israel-Palestinian case.

Figure 4 shows the cross-correlation of the mediation indicator with Israeli-Palestinian conflict measure. The correlogram shows a very distinct pattern of positive correlations for lagged values of conflict and zero or negative correlations for conflict in the period following the mediation. In other words, mediation correlates with the level of conflict in the months before the mediation and correlates with decreased conflict following the mediation. The levels of correlation are relatively low, and the highest correlations are barely significant at the 5% level, but the overall pattern is quite regular. Although the individual cross-correlations in the period following mediation are not significant at the two-tailed 5% significance level, the overall pattern is significant. For example, the 5% critical value for the minimum absolute value of three consecutive cross-correlations is around ±0.05. This value was determined by Monte Carlo approximation (see Appendix B), and the correlations satisfy this for \( k > 7 \).

Figure 4 also shows the correlogram for mediation and the level of Israel-Lebanon conflict. This illustrates a very different pattern than that of the Israel-Palestinian case. The strongest correlations are contemporaneous—roughly 2 months before and after the mediation—and positive, indicating the mediation is most likely to occur when the level of conflict is high. However, the correlogram gives no evidence that the mediation is effective. The correlations between mediation and subsequent conflict remain near 0 or slightly positive. Conflict in the Israel-Lebanon case produces mediation efforts, but these have no effect on reducing that conflict.

Our analysis of the conflict in the former Yugoslavia looked at three different sets of mediators: the UN, the United States, and Europe (operationalized as the EU, France, Germany, Italy, and the United Kingdom). The measures for the Bosnian, Croatian, and Serbian conflict mediators differ slightly from the pattern described above. The United States and European Union are the most active mediators, while the UN is less active. The levels of conflict and activity are positively correlated for most of the conflicts, indicating that mediation is more likely to occur when the level of conflict is high. The correlations between mediation and subsequent conflict remain near 0 or slightly positive. Conflict in the Israel-Lebanon case produces mediation efforts, but these have no effect on reducing that conflict.

5. As a check that this pattern is actually measuring mediation and not just interactions, we also ran a cross-correlation between Goldstein-scaled behavior from the United States-Israel and Israel-Palestinian conflict. The Goldstein-scaled score differs from the mediation score because it only measures interactions between the United States and Israel, without adjusting for whether the United States is talking (or otherwise cooperating) with both sides, and also takes into account both positive and negative interactions (e.g., U.S. criticism of Israel). The resulting correlogram, which can be viewed on the KEDS Web site (http://www.ukans.edu/~keds/ISA01.supplement/ISA01.Supplement.html), is quite different from Figure 4. It shows the typical positive spike of contemporaneous correlation at -1, 0, and +1 months, but otherwise the correlation is flat and close to 0. We conclude from this that the mediation indicator is picking up something more than simple interaction.

6. A second positive spike occurs at +13 and +14 months. We have encountered this before in other cross-correlation studies of data on Lebanon, and it appears to be due to some sort of weak seasonality in the levels of conflict in this dyad.

7. We also analyzed mediation by Russia and the Ukraine. This series has substantially less variance than the European mediation series (17.8 vs. 97.6), but generally shows a pattern similar to that of Europe.
and Serbian actors combine the activities of the various ethnic factions with those of the governments. In other words, Serbia includes the actions of ethnic Serbs in Bosnia and Croatia as well as the actions of the Serbian government. As with all event data, the identification of the ethnicity of individuals or groups responsible for actions was dependent on how the event was reported in the news story.

The anecdotal accounts of the Balkans conflict suggest that the effectiveness of mediation efforts varied substantially, depending on who was doing the mediation (see Kaldor 1999, 31-68; Weiss 1999, 97-136). This assessment is strongly supported by our cross-correlation analysis. Figure 5 shows the cross-correlogram of the various mediators and the level of Serbia-Bosnia conflict. The three mediators show quite different patterns. The correlations for the United Nations are significantly positive both prior to and following the mediation. In other words, the UN mediation increased during periods of increased conflict in the dyad, but, in contrast to the pattern seen for mediation in the Levant, the level of conflict actually became greater following the mediation. U.S. mediation efforts, in contrast, reduced the level of conflict: there is a positive correlation with conflict prior to the mediation, but a significant negative correlation in the period following the mediation.8 Finally, European mediation efforts have no effect. The European mediation variable shows the usual positive correlation in the lagged period, but most of the correlations with conflict levels following the mediation are close to 0. Unlike the UN efforts, European mediation does no harm, but it does no good, either.

The pattern for the Serbia-Croatia case (see Figure 6) is generally similar, but with a couple of differences. First, the period of significant lagged correlations (i.e., mediation responding to increased conflict) is about half the length of the comparable period for Bosnia; this may be due in part to the more concentrated character of fighting in the Serbia-Croatia conflict, which generally occurred in a few months in 1991 and 1995. Second, the period of positive correlations following UN mediation is shorter, and, unlike the Bosnia case, there are no positive correlations following European mediation. The most significant negative correlations are found with the “ALL” measure of mediation, which could either indicate successful coordination of mediation efforts or simply be a statistical artifact.

A comparison of all the correlograms reveals an interesting trend about the initiation of mediation activities: in the Serbia-Bosnia and the Serbia-Croatia cases, correlations peak 2 to 5 months before mediators intervene (i.e., they occur to the left of the middle of the correlograms), indicating that mediators respond to rising levels of conflict about 2 to 5 months after they occur. In the Middle East, and particularly in the Israel-Palestinian conflict, we do not see major changes in the correlation between mediation and conflict levels across the 40-month period, except for mediation efforts in the Israel-Lebanon case that occur almost contemporaneously with high levels of conflict. This is probably due to the nearly constant involvement of mediators. In short, in comparing the cases, we not only see a variation in the time it takes mediation

8. Although the individual correlations for the United States are barely significant at the .05 level, the critical value for the minimum of three consecutive correlations is ±.07 (see Appendix B), and the U.S. pattern clearly satisfies this criterion.
to show results, but also an important difference in the time period it takes mediators to mobilize and respond to different conflicts.

**TIME SERIES ANALYSIS OF THE STICKS-OR-CARROTS MODEL**

Our second set of tests looks at the sticks-or-carrots issue: is mediation more likely to be effective when it is accompanied by material cooperation or conflict? Material cooperation involves provision of tangible incentives or side-payments, which “by [enlarging] the pie” could create a positive-sum game for the receiving antagonist and thereby induce its cooperation (Touval 1982). Material conflict, on the other hand, refers to coercive measures, such as military force and economic sanctions. The literature typically prescribes the use of such forceful tools in conjunction with cooperative incentives when the latter prove to be ineffective alone. It is also recognized, however, that although coercive means could help to terminate violence temporarily, they fail to create an environment that is conducive to the resolution of a conflict (Laue 1993).

We examine the relationship between the mediation strategies mentioned above and the level of conflict, using the mediation measure pattern and event counts of material cooperative and conflictual behavior between the mediator and antagonists as the independent variables, and several measures of mediation effectiveness as the dependent variable. The models tested have the following general form:

\[
y(t) = a + b_1\text{mediation}(t) + b_2\text{cooperation}_A(t) + b_3\text{conflict}_A(t) + b_4\text{cooperation}_B(t) + b_5\text{conflict}_B(t).
\]

The dependent variables tested were the following:

- scaled conflict difference: scaled conflict \((t + k) - \text{scaled conflict}(t)\)
- conflict frequency difference: conflict frequency \((t + k) - \text{conflict frequency}(t)\)
- scaled conflict lags: scaled conflict \((t + k)\)
- conflict frequency lags: conflict frequency \((t + k)\),

where scaled conflict is the negative of the Goldstein-scaled score, and conflict frequency is the count of material conflict events; both measures are aggregated by month.

The independent variables consider only material cooperation and conflict. The subscripts \(A\) (the stronger antagonist) and \(B\) (the weaker antagonist) denote the targets of the mediator’s cooperative or conflictive behavior. So, for example, in the EURBF case, \(\text{cooperation}_A\) is the number of events in the material cooperation category with either the EU, France, Germany, Italy, or the United Kingdom as the source actor, and Serbia as the target.

Figure 7 summarizes a number of experiments with different formulations of the dependent variable across different lags. It shows the average values of the \(z\) score on the mediation variable across the 10 cases we have studied for various lags \((k = 0\) to \(10\))
Effective mediation, in the sense of violence reduction, should result in negative $z$ scores for all formulations, and the results in Figure 7 show a very consistent pattern. First, the shape of the curve over the lag times is quite similar for the four formulations, with a high positive contemporaneous value (no lag in the independent variables), and then a decline to 0 or negative values with the increasing lag, which levels off around $k > 4$.

When mediation success is measured by the change in conflict levels, the $z$ score on mediation is consistently stronger than when conflict is measured by the level, and the $z$ scores are consistently stronger when conflict is measured by conflict frequency than with scaled conflict. These are averages for the 10 cases, including two Balkans cases that have poor mediation success. Thus, although the average in the best case—difference in conflict measured by event counts—is barely significant, the $z$ scores in several of the individual cases are quite significant.

Figure 7 is arguably treading a fine line between exploratory analysis and a statistical fishing expedition. We argue that this exploration is necessary for at least two reasons. First, we know from the cross-correlation analysis and the qualitative literature that there is a substantial lag between mediation efforts and changes in the level of conflict, and that mediation usually is positively correlated with conflict (i.e., increased conflict is contemporaneously associated with increased mediation). Sorting out these two effects is problematic.

Second, we are still unclear as to whether it is better to study these behaviors using scaled or frequency measures of conflict. There is clearly not a whole lot of difference between the two, although in a series of additional experiments, we found that the frequency measure usually produces slightly stronger relationships with the mediation variable, whether measured through the $R^2$ of the entire equation or the $z$ score on the
mediation variable. This may be because mediation is itself an event-frequency measure, or it may be evidence against the utility of scaled event data. We therefore analyzed both formulations.

Table 1 shows the results of estimating the sticks-or-carrots model for the change in the scaled conflict and conflict frequency variables for the difference of 4 months, that is, scaled conflict(t + 4) – scaled conflict(t), and conflict frequency(t + 4) – conflict frequency(t), respectively. Our tests with lags higher than 4 gave results that were generally similar to that with lag 4. This is consistent with Figure 7, suggesting that mediation takes hold after a lag time of about 4 months and then has a generally consistent effect.

When ordinary least squares (OLS) regression was used, about half of the cases had significant Durbin-Watson statistics, indicating the presence of first-order, serially correlated residuals. Consequently, the estimates in these tables use the Prais-Winsten transformed regression estimator (Stata prais), although, in general, the pattern of significant coefficients is the same in the OLS and Prais-Winsten estimates. To reduce the size of the table, only coefficients for variables that were significant at a level of $p < .10$ are reported; full results are available from the authors.

Two general patterns are evident from these results. First, the correlations for the equations as a whole are significant on almost all of the regressions. Most of the exceptions involve either the UN, as expected, or the United States mediation on Croatia. In contrast to the cross-correlation analysis, there are significant correlations in all of the Levant cases.

Second, the mediation variable, when significant, is always negative. There are no exceptions to this pattern, confirming that third-party involvement in the form of meditative activity does lower the level of conflict between the antagonists. Regression results with respect to this variable are generally consistent with the cross-correlation values around 4 months after mediation. A striking discrepancy is with the UN mediation in the Serbia-Bosnia conflict. Although cross-correlation surprisingly implied that mediation by the UN further increased the level of conflict between the antagonists, this association is found to be significantly negative in the Prais-Winston regression.

The analysis, however, is less clear on the sticks-or-carrots question. Although $conflict_A$ and $conflict_B$ have negative coefficients in all cases and both formations of the dependent variable, there are mostly positive coefficients on $cooperation_B$, and there is no discernible pattern to $cooperation_A$. These findings suggest that material conflict directed toward the weaker and the stronger antagonists is associated with reduced conflict between fighting parties. The relationship is not significant in all of the cases, but material conflict by a mediator consistently has the desired effect in all instances where it is found to be a significant factor. Mediation accompanied by material cooperation toward the weaker party, on the other hand, is found generally to decrease the level of conflict. However, based on these results, it is difficult to make a generalized

---

9. Exceptions to this observation involve differences in the presence of a significant coefficient in some cases; in no case, however, do we observe the sign of a coefficient changing with the change in the lag time. Detailed results may be obtained from the authors.
### TABLE 1
Sticks-or-Carrots Models ($k = 4$), Prais-Winston Regression

<table>
<thead>
<tr>
<th>Case</th>
<th>Variable</th>
<th>Conflict Frequency Difference</th>
<th></th>
<th></th>
<th>Scaled Conflict Difference</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$R^2$</td>
<td>Estimate</td>
<td>t</td>
<td>$R^2$</td>
<td>Estimate</td>
<td>t</td>
</tr>
<tr>
<td>ISRPAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mediation</td>
<td>0.077***</td>
<td>-0.11</td>
<td>-2.55**</td>
<td></td>
<td>0.042*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cooperation$_B$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conflict$_B$</td>
<td></td>
<td>-2.15</td>
<td>-2.72***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISRLEB</td>
<td>Conflict$_A$</td>
<td>0.125***</td>
<td>-3.40</td>
<td>-3.90***</td>
<td></td>
<td>0.104***</td>
<td>-27.44</td>
</tr>
<tr>
<td>ALLBFR</td>
<td>Mediation</td>
<td>0.164***</td>
<td>-0.07</td>
<td>-3.73***</td>
<td></td>
<td>0.098**</td>
<td>-0.46</td>
</tr>
<tr>
<td></td>
<td>Cooperation$_A$</td>
<td></td>
<td>0.37</td>
<td>1.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USABFR</td>
<td>Mediation</td>
<td>0.271***</td>
<td>-0.14</td>
<td>-4.67***</td>
<td></td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cooperation$_A$</td>
<td></td>
<td>2.06</td>
<td>3.18***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EURBFR</td>
<td>Mediation</td>
<td>0.107**</td>
<td>-1.25</td>
<td>-2.99***</td>
<td></td>
<td>0.172***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cooperation$_A$</td>
<td></td>
<td>-0.96</td>
<td>-1.97**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conflict$_A$</td>
<td></td>
<td>-0.47</td>
<td>-1.82*</td>
<td></td>
<td>-4.64</td>
<td>-1.82*</td>
</tr>
<tr>
<td>UNOBFR</td>
<td>Mediation</td>
<td>0.032</td>
<td>-0.10</td>
<td>-1.77*</td>
<td></td>
<td>0.055</td>
<td></td>
</tr>
<tr>
<td>ALLCRO</td>
<td>Cooperation$_B$</td>
<td>0.135***</td>
<td></td>
<td></td>
<td></td>
<td>0.112**</td>
<td></td>
</tr>
<tr>
<td>USACRO</td>
<td>Cooperation$_B$</td>
<td>0.063</td>
<td>-2.59</td>
<td>-1.74*</td>
<td></td>
<td>-12.05</td>
<td>-1.77*</td>
</tr>
<tr>
<td>EURCRO</td>
<td>Mediation</td>
<td>0.164***</td>
<td>-0.53</td>
<td>-2.88***</td>
<td></td>
<td>0.164***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cooperation$_B$</td>
<td></td>
<td>-1.31</td>
<td>-1.97**</td>
<td></td>
<td>-10.89</td>
<td>-1.95**</td>
</tr>
<tr>
<td>UNOCRO</td>
<td>Cooperation$_B$</td>
<td>0.077*</td>
<td>0.085*</td>
<td></td>
<td></td>
<td>-19.43</td>
<td>-1.71*</td>
</tr>
</tbody>
</table>

**NOTE:** ISRPAL = conflict between Israel and Palestinians, mediated by the United States, the European Union, and the United Nations; ISRLEB = conflict between Israel and Lebanon, mediated by the United States, the European Union, and the United Nations; ALLBFR = conflict between Serbia and Bosnia, mediated by the United States, the European Union, France, Germany, Italy, the United Kingdom, and the United Nations; USABFR = conflict between Serbia and Bosnia, mediated by the United States; EURBFR = conflict between Serbia and Bosnia, mediated by the United States; UNOBFR = conflict between Serbia and Bosnia, mediated by the European Union, France, Germany, Italy, and the United Kingdom; ALLCRO = conflict between Serbia and Croatia, mediated by the United States, the European Union, France, Germany, Italy, the United Kingdom, and the United Nations; USACRO = conflict between Serbia and Croatia, mediated by the United States; EURCRO = conflict between Serbia and Croatia, mediated by the European Union, France, Germany, Italy, and the United Kingdom; UNOCRO = conflict between Serbia and Croatia, mediated by the United Nations.

*Significant at .10 level, two-tailed. **Significant at .05 level, two-tailed. ***Significant at .01 level, two-tailed.

Conclusion because the variables that are significant and the pattern of findings differ substantially across the different cases of mediation.
The ISRPAL case is particularly idiosyncratic. It is the only case where material cooperation toward the weaker party is associated with increased conflict (and that is only when the scaled measure is used). In the Palestinian case, increased material cooperation is generally found in the post-Oslo period and a shorter period shortly after Israel’s 1982 invasion of Lebanon. These are periods of rising total conflict, which would explain the positive coefficient. On the other hand, conflictive behavior toward the Palestinians increases the level of cooperation between the latter and Israel; this relationship is consistent with the only other case where conflict is found to be significant, EURCRO.

A pattern that is hard to explain across the different mediation cases in the Serbia-Bosnia conflict is the role of material cooperation, cooperationA, between the mediators and Serbia. The direction of the relationship is not consistent across the cases where the variable is found to be significant (all but UNOBFR). When mediation by all of the potential mediators or the United States by itself is accompanied by material cooperation with Serbia, conflict between Serbia and Bosnia increases following the mediation. In the case of mediation by the European actors, on the other hand, cooperative behavior towards Serbia is associated with a decrease in the level of conflict.

Interpretation of the findings in the Serbia-Croatia conflict is more straightforward. Cooperation between the mediators and Croatia is associated with a significant increase in cooperation between the antagonists in all the cases except EURCRO, where cooperationB is not significant at all. It is conflict, and not cooperation, between the European actors and both of the antagonists that is associated with reduced conflict in case of European mediation. It is interesting to note that in the conflicts in which Serbia is involved, material conflict between the European mediators and this stronger antagonist is negatively correlated with the level of conflict between the antagonists following mediation.

Based on this analysis, mediation efforts in the Israel-Lebanon case appear to be similar to those in the Balkans. As we expected, Israel-Palestine is quite different. We speculate this is in part a function of the close relationship between Israel and the United States, which is not mirrored in the other cases. It may be further complicated by the diverse set of groups that we code as “Palestinian.” An analysis that differentiated between the groups associated with Arafat—Fatah and the Palestinian National Authority following Oslo—and various more militant groups might clarify this relationship. Our data set actually contains codes for the more common Palestinian groups such as Hamas, the Popular Front for the Liberation of Palestine (PFLP), and the Democratic Front for the Liberation of Palestine (DFLP); in a future analysis, we could analyze these separately.

CONCLUSIONS AND FUTURE DIRECTIONS

This analysis was intended to illustrate three main points. First, it is possible to formulate meaningful hypotheses about the dynamics of mediation processes (as distinct...
from their structural characteristics) and test these hypotheses using statistical methods. Second, by using a combination of machine-coded event data and relatively simple definitions of event sequences, it is possible to derive credible measures of behavior relevant to this study in a transparent and reproducible manner. Finally, time-series case studies can provide insight into how differences in contextual variables across cases might produce variation across mediation processes and outcomes.

This is the first major analytical work from a larger project, and we regard these results as illustrative rather than conclusive. The hypotheses that we have studied here do not capture many of the nuances (or inconsistencies) in the existing theoretical literature on mediation. In this study, we have used data coded in the cold war-era, WEIS framework, which assumes a "Westphalian-Clausewitzian" political worldview of sovereign states reacting to each other through diplomacy and military threats rather than dealing with contemporary issues, such as ethnic conflict.

We are currently implementing a new coding scheme called Conflict and Mediation Event Observations (CAMEO). CAMEO is specifically designed to code events relevant to the mediation of violent conflict and is being implemented using the new, open-source Textual Analysis by Augmented Replacement Instructions (TABARI) automated coder from the KEDS project. Following the example of the Integrated Data for Events Analysis (IDEA) codebook from the Protocol for the Assessment of Nonviolent Direct Action (PANDA) project (http://vranet.com/idea/), the CAMEO codebook will exist in a printed and Web-based format. We have also followed the lead of IDEA in introducing 4-digit tertiary coding categories that focus on very specific types of behavior, for example, differentiating agreement to or rejection of cease-fires, peacekeeping, and conflict settlement. In contrast to IDEA, we have not retained strict backwards compatibility with WEIS. Instead, we have combined some of the WEIS categories and provided far more detail on others (notably our expansion of the WEIS 23 "force" category into four distinct primary categories). Nonetheless, it should be possible to map most of the CAMEO secondary categories unambiguously into a WEIS secondary category.

Patterns of mediation behaviors (or any other political behavior) are likely to have a significant empirical component that is distinct from the theoretical considerations of the academic literature on the subject. Therefore, it is important to experiment with coding systems rather than to attempt to establish these a priori. Because of the high degree of selectivity of news reporting, a behavior that is important in a case study—the analytical approach that still informs most of the mediation literature—will necessarily show up as a useful statistical indicator. Similarly, good exploratory analysis of the event data may reveal indicators not found in the theoretical literature, often because these are surrogates for other variables. This is not to say that statistical studies should be atheoretical, but the development of useful statistical models will, in part, be an empirical exercise of matching methods to data.

The analysis reported here has made us acutely aware of the problem of analyzing processes that are known to have a stochastic time lag between the action and the effect. This has emerged as a major complication when we are using conventional time...
series methods that involve the correlation of variables at a fixed lag \((t - k)\). Because we do not know of any "natural" time lag or set of lags to apply in these models, we are instead left with at least three options, none of which we find wholly palatable.

The first method that we employed in the time-series analysis here is to use exploratory methods, such as cross-correlation and experimentation with plausible alternative lag structures, to get a general idea of the time lag where the effects of the relationship seem to be strongest, and do the analysis with a small number of lags. This runs the risk, however, of overfitting the data, and the choice of the time lag is somewhat arbitrary, particularly when the effects are spread across a number of months, as we have seen in this case.

The second method would be to use a range of credible lags for all of the independent variables. This is the approach used in vector autoregression (VAR), which has been employed in a number of studies that use event data (e.g., Goldstein and Freeman 1990; Goldstein and Pevehouse 1997). The disadvantage of VAR is that it puts one into the realm of diffuse parameter structures with indeterminate values. When the independent variables are auto-correlated (as the sequences studied here generally are), collinearity expands the standard errors of the VAR coefficients to the point where they cannot be interpreted substantively. On the positive side—at least for the analyst—the decades-long crises that we are examining provide sufficient degrees of freedom, even at monthly levels of aggregation, that fairly elaborate VAR models can be estimated.

The final method would be to use a method such as event history models or Poisson regression, where the stochastic delay between the treatment and response is the dependent variable. In Schrodt et al. (2001), we estimated a number of Cox proportional hazard models on the data sets used in this study. These models used the time periods between violence and mediation, mediation and agreement, and agreement and subsequent violence as the dependent variables, and the same measures of actor interactions that we used in the sticks-or-carrots analysis as the independent variables. In contrast to the analyses presented here, the Cox models showed few consistent patterns of behavior beyond those already evident from the time-series analysis (e.g., that UN mediation did not work). We have not experimented with parametric variants on the basic Cox model, for example, using the Weibull or Gompertz distributions, and it is possible that those might produce more informative results.

We suspect that the issue of specifying lagged responses has not received a great deal of attention in the existing quantitative literature in international politics because most of these studies have used data aggregated by year. That time period is sufficiently long so that most responses will appear to occur either contemporaneously or, at most, with a lag of one period. Event data, in contrast, can effectively be aggregated to a month or even a week. At this level of detail, there is usually a substantial difference between the time that a change in behavior occurs and the time its effects are observed.
APPENDIX A
World Event Interaction Survey (WEIS) Cue Categories and Average Goldstein Scale Scores

<table>
<thead>
<tr>
<th>Code</th>
<th>Content</th>
<th>Average Goldstein Scale Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Yield</td>
<td>1.1</td>
</tr>
<tr>
<td>02</td>
<td>Comment</td>
<td>-0.1</td>
</tr>
<tr>
<td>03</td>
<td>Consult</td>
<td>1.9</td>
</tr>
<tr>
<td>04</td>
<td>Approve</td>
<td>3.5</td>
</tr>
<tr>
<td>05</td>
<td>Promise</td>
<td>4.3</td>
</tr>
<tr>
<td>06</td>
<td>Grant</td>
<td>2.2</td>
</tr>
<tr>
<td>07</td>
<td>Reward</td>
<td>7.4</td>
</tr>
<tr>
<td>08</td>
<td>Agree</td>
<td>4.8</td>
</tr>
<tr>
<td>09</td>
<td>Request</td>
<td>1.6</td>
</tr>
<tr>
<td>10</td>
<td>Propose</td>
<td>0.7</td>
</tr>
<tr>
<td>11</td>
<td>Reject</td>
<td>-4.0</td>
</tr>
<tr>
<td>12</td>
<td>Accuse</td>
<td>-2.8</td>
</tr>
<tr>
<td>13</td>
<td>Protest</td>
<td>-2.1</td>
</tr>
<tr>
<td>14</td>
<td>Deny</td>
<td>-1.0</td>
</tr>
<tr>
<td>15</td>
<td>Demand</td>
<td>-4.0</td>
</tr>
<tr>
<td>16</td>
<td>Warn</td>
<td>-3.0</td>
</tr>
<tr>
<td>17</td>
<td>Threaten</td>
<td>-6.0</td>
</tr>
<tr>
<td>18</td>
<td>Demonstrate</td>
<td>-6.4</td>
</tr>
<tr>
<td>19</td>
<td>Reduce Relation</td>
<td>-4.5</td>
</tr>
<tr>
<td>20</td>
<td>Expel</td>
<td>-5.0</td>
</tr>
<tr>
<td>21</td>
<td>Seize</td>
<td>-6.8</td>
</tr>
<tr>
<td>22</td>
<td>Force</td>
<td>-9.0</td>
</tr>
</tbody>
</table>


APPENDIX B
Cross-Correlation

Cross-correlation is useful in determining if a behavior has a long-term effect when the timing of that effect is not specified theoretically. The technique has not been widely used in conflict studies, and some explanation is in order.

The cross-correlation function is similar—but not identical—to computing the Pearson product moment “r” between x, and y ± k for various values of k. Both statistics have the following form:

\[ r = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x) \text{Var}(y)}} \]

In a cross-correlation, Var(x) and Var(y) are estimated from the entire sample, whereas in a Pearson product moment, these variances are computed only on the cases that were used to compute the covariance. The resulting cross-correlograms are not a time series giving the effect of a single mediation on subsequent behavior; they are a correlation of the mediation with prior and future behavior for the entire time period. For additional information on cross-correlation, see Kendall (1973, 129), Chatfield (1989, 136), and Gottman (1981, 318).
The approximate critical value of the cross-correlation coefficient at the 5% two-tailed significance level is plus or minus $2/\sqrt{N}$, which is roughly 0.13 for the Levant case and 0.18 for the Balkans case. However, that approximation assumes that the variables have neither trend nor autocorrelation. Because the correlograms in Figures 4, 5, and 6 were computed on the raw series—which have some trend and autocorrelation—the significance of the correlation may be overestimated (see Chatfield 1989, 137-40).

In the past, the recommended statistical solution to this problem has been to pound on the data until it fits some analytical model with known properties because, in the absence of such pounding, one would know nothing about the estimators. However, contemporary computationally intensive statistical analysis allows one to approximate the properties of estimators empirically, based on a much broader set of assumptions. These empirical approximations may, in fact, be more accurate than the analytical approximations, which often are only valid asymptotically.

In addition, the traditional approach of applying a sequence of standard time-series transformations that improve the asymptotic statistical characteristics of the estimators take the data and the analysis further from anything that an analyst can actually understand. For example, when trend and autocorrelation are removed from the time series for U.S. mediation in the Serbia-Bosnia conflict (using detrending, then first-difference), the resulting correlogram has significant positive correlations at lags of -17, -10, -7, -4, -2, -1 and a lead of +1, and negative correlations at leads of +6 and +13. These results are generally tell the same story as the untransformed data: U.S. mediation responds to past periods of high conflict and reduces the level of later conflict.

However, the two series on which such a statistically correct correlogram was computed are almost impossible to explain, and one cannot say that the correlogram implies that U.S. mediation reduces conflict only at 6 months and 13 months. This is true only for the detrended and differenced values of that series, a set of transformations that is nearly meaningless from the perspective of figuring out the underlying behavior.

Given the advances in computing power, an alternative approach to this problem is to estimate the critical values of the significance test using a Monte Carlo simulation. Table 1A provides some examples of this, with approximations of the 5% significance bounds based on 30,000 Monte Carlo trials. The $2/\sqrt{N}$ and "no autocorrelation" rows compare the standard approximation to the Monte Carlo estimate under the null hypothesis that neither series contains autocorrelation. Based on these experiments, the approximation is conservative by about 0.02. The next two rows show critical values for data generated with autocorrelation structure estimated in the Bosnian and Israel-Palestinian data. Despite significant autocorrelation, the bounds for the Israel-Palestinian case are just the same as those in the $2/\sqrt{N}$ approximation. The much higher autocorrelation found in the Bosnian case, in contrast, widens the bounds substantially.

Using the Monte Carlo approximation, it is also possible to estimate empirically the significance bounds for conditions that would be virtually impossible to estimate analytically. The final two rows of Table 1A confirm our assertion in the discussion that whereas the individual cross-correlations for the U.S. mediation in Bosnia are only barely significant, the probability that multiple, consecutive cross-correlations would all have a high value is much less likely. The bounds for $\min = -1, 0, 1(CCF_{k, j})$ show the 5% critical value for the minimum of three consecutive cross-correlations. As expected, the United States-Bosnia case is well above this level. By way of contrast, $\max = -1, 0, 1(CCF_{k, j})$ shows the 5% critical value for the maximum of three consecutive cross-correlations; none of our correlations passes this much stricter criterion.
TABLE 1A
Monte-Carlo Critical Value Estimates of 5%, Two-Tailed Significance Level
for Cross-Correlation under Various Assumptions

<table>
<thead>
<tr>
<th></th>
<th>N = 128</th>
<th>N = 256</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/√N</td>
<td>0.18</td>
<td>0.13</td>
</tr>
<tr>
<td>No autocorrelation</td>
<td>0.16</td>
<td>0.11</td>
</tr>
<tr>
<td>y_t = 0.85 y_{t-1} + e</td>
<td>0.22</td>
<td>0.16</td>
</tr>
<tr>
<td>x_t = 0.38 x_{t-1} + e</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>y_t = 0.23 y_{t-1} + e</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>x_t = 0.57 x_{t-1} + e</td>
<td>0.20</td>
<td>0.14</td>
</tr>
<tr>
<td>min_j = -1, 0, 1(CCF_{k+j})</td>
<td>0.07</td>
<td>0.05</td>
</tr>
<tr>
<td>max_j = -1, 0, 1(CCF_{k+j})</td>
<td>0.20</td>
<td>0.14</td>
</tr>
</tbody>
</table>

REFERENCES


McClelland, Charles A. 1970. Some effects on theory from the international event analysis movement. Mimeo, University of Southern California, February.


