The Dynamics of Reciprocity, Accountability, and Credibility *

Patrick T. Brandt
School of Economic, Political and Policy Sciences
University of Texas at Dallas
E-mail: pbrandt@utdallas.edu

Michael Colaresi
Department of Political Science
Michigan State University
E-mail: colaresi@msu.edu

John R. Freeman
Department of Political Science
University of Minnesota
E-mail: freeman@polisci.umn.edu

July 26, 2007

Abstract

Do public opinion dynamics play an important role in understanding conflict trajectories between democratic governments and other rival groups? The majority of previous research has assumed either that public opinion is irrelevant to conflict processes or that the relationships are one-way causal chains. In this paper, we argue that neither of these assumptions are theoretically or empirically necessary. Instead, we interpret several theories of opinion dynamics and government behavior as particular causal links in models of reciprocity, accountability and credibility relationships. Theoretical expectations about the character of these linkages are translated into four distinct Bayesian structural time series models. These models allow us to include novel domestic public information where available, as well as relax the strict recursive structure that previous time series models have assumed. The models are fit to events data from the Israeli-Palestinian conflict with provisions for U.S. intervention and public support for peace. We find that a credibility model, which allows domestic public opinion to influence U.S., Palestinian and Israeli behavior within a given month, fits the data best. This credibility model supports research that predicts asymmetric reciprocity between democratic and non-democratic belligerents. For the credibility model there is evidence that more pacific Israeli opinion leads to more immediate hostility by the Palestinians toward the Israelis. The direction of this response suggests a negative feedback mechanism where low level conflict is maintained and momentum toward either all out war or dramatic peace is slowed.

1

1 Introduction

Much progress has been made recently in the analysis of conflict processes. Reciprocal behaviors, uncovered in the Balkans and Middle East, it is argued, are the bases for cooperation and peace. The same is true for the triangular relationships that researchers have uncovered in these and other conflicts. Third party intervention in conflicts, especially by great powers, also can promote cooperation and peace. What has not been studied is how and when public opinion affects conflicts. Does public opinion forbid or encourage local leaders to reciprocate other rival's behavior? Do local publics monitor international conflicts and hold leaders accountable for policies that do not match public preferences? Do leaders use public opinion in a rival country to gauge the credibility of signals they receive from those adversaries? We will not understand how conflict depends on domestic politics, including democratic politics, until we begin to focus on the interplay between the domestic and international arenas.[1]

Up to this point answering these questions has been difficult for several reasons. First, domestic politics is embedded in the conflict system. This is a system of multiple relationships (behavioral equations) between the behaviors of several governments. Any model purporting to capture domestic politics must be of moderate to large scale. Endogeneity is a second problem. Accountability, if it exists, implies that opinion formation and expression are both a cause and a consequence of government policy. Reciprocity and triangularity also imply endogeneity as well. Empirical models that impose strong exogeneity restrictions are thus liable to produce results that are biased. Finally, any link between conflict and domestic political dynamics is likely to exist at a sub-annual level of temporal aggregation. Public opinion likely reacts quickly to hostile actions and leaders may monitor these reactions and rapidly translate them into policies.Analyzing conflicts in terms

---

[1]In this paper we refer to international conflict as conflict between organized, self-identifying groups that may or may not control territory (nations). This has become standard in the event data literature, where many of the actors of interest—rivals—represent nations in this sense, but fall short of unanimous classification as states (e.g. Palestinian Authority, Kosovo, Taiwan, and Hong Kong). For a lucid explanation of the differences between state and pseudo-state actors see Lemke (2007).

of lagged quarterly or yearly patterns of behavior will likely miss these dynamics.

We assess the impact of domestic politics on international and inter-group conflict in a way that solves these three problems. Reciprocity, accountability, and credibility are interpreted as particular causal links between governments' behaviors and public opinion. Expectations about these linkages are described and translated into a framework that captures competing claims about the structure of contemporaneous relationships among the respective variables (belligerents) and competing arguments about conflict dynamics. These specifications can be contrasted with the common practice of using a recursive identification scheme that assumes one-way causality. Our competing specifications are key elements of multi-equation, Bayesian time series models with complex, endogenous relationships among variables. These models are fit to data for the Israeli-Palestinian conflict with provisions for U.S. intervention. This case is especially useful for testing the competing theories because the Kansas Event Data System (KEDS) and other event databases provide temporally disaggregated records of the behaviors of the belligerents and of the U.S. Equally important, the Tami Steinmetz Center for Peace Research regularly polls Israeli citizens about prospects for and impacts of peace initiatives. Using these public opinion data, we believe that for the first time in a major conflict, the impact of public opinion on reciprocity, accountability and credibility can be assessed.[2]

The paper has three parts. Part one derives competing arguments about the impact and origins of public opinion on conflict dynamics. The research design in part two translates the competing arguments into four Bayesian time series models that are presented in part three. We find that a cross-level credibility model that supports asymmetric reciprocity between democratic and non-democratic belligerents fits best. In this model there is evidence that more bellicose Israeli opinion leads to more immediate cooperation by the Palestinians. This supports the credibility model since

---

[2]The novelty and importance of the Steinmetz Center data results from the consistency of the relevant questions over time, periodicity of the polling increments (monthly rather than quarterly or yearly) and relatively large number of time points for which data is available (over 100 months). While public opinion databases for a small number of countries tend to share one or two of these traits, no other source of information that we are aware of fits all three criteria.

Palestinians account for changes in Jewish public opinion. The direction of this response suggests a negative feedback mechanism where low level conflict is maintained and momentum toward all out war or dramatic peace is slowed. Finally, we present two forecasting models. The first, based on the credibility model, is used to forecast nine months from March 2005. The forecasts show a rapid deterioration in Israeli-Palestinian relations. The *ex ante* forecasts predict steadily deteriorating relations between the Israelis and Palestinians even if Kadima and Hamas had not won recent elections. The risk of a significant increase in conflict—based on the confidence regions—is high. The forecasts including the public opinion measure shows that the level of violence and support for peace move in opposite directions. The second forecast, a model that omits the Jewish public opinion series, does much worse because it ignores the predictive impact of public opinion and thus accountability and credibility on the conflict dynamics.

2 Reciprocity, Accountability and Credibility

2.1 Reciprocity

One important branch of international relations and conflict research investigates the role of reciprocity in foreign policy. Building on Axelrod (1984) and repeated non-zero sum games (e.g., iterated prisoners' dilemma), many have found that cooperative diplomacy from one actor begets cooperative diplomacy from the target. This theoretical result suggests that cooperation can be built upon bilateral strategies even when temporary incentives to defect exist. If one group has an opportunity to win disputed territory from another, it may choose not to attack since this could discourage cooperation in the future. The prevalence of these tit-for-tat strategies has been empirically identified in Cold War relations, and triangular relations between the U.S., U.S.S.R. and China (Goldstein and Freeman, 1990), in Indian-Pakistani relations (Ward, 1982), in the Middle East (Goldstein et al., 2001) and in the Balkans (Goldstein and Pevehouse, 1997). Scholars have been able to show that similar action-reaction sequences exist in sub-national conflicts (Moore, 1998; Shellman, Reeves and Stewart, 2006).

Extensions of the reciprocity literature relate bilateral cooperation to third party intervention or mediation (Goldstein and Freeman, 1990; Brandt and Freeman, 2006; Hudson et al., 2006). Goldstein et al. (2001) report that the U.S. was able to alter Iraqi behavior toward its neighbors and Israeli behavior toward the Palestinians in the 1980s. The modeling of these triangular relations has allowed reciprocity-related research to focus not only on spirals of conflict and cooperation, but also to explain more complex multi-actor situations.

Despite the strong evidence of reciprocity and action-reaction sequences in international affairs, there are at least three remaining empirical puzzles. First, there are alternative tests of reciprocity that miss the implications of endogeneity and lagged relationships. Two empirical approaches, both subject to criticisms, have been used to establish reciprocity. One approach, used in almost every empirical model that claims evidence for reciprocity and triangularity, is Granger causality analysis. This method looks only at the bivariate relationship between how *past* actions affect today's actions. This inferential methods ignores the (contemporaneous) endogeneity of multiple belligerents, third-parties, and outside pressures (Goldstein and Pevehouse, 1997; Goldstein et al., 2001, e.g.,). A second approach looks at the contemporaneous and dynamic responses—the structural relationships—between belligerents. Studies without an explicit structural model are akin to a naive recursive model where the shocks to the equations for each actor enter the model in some pre-specified order (Goldstein and Freeman, 1990, e.g.,).[3] In a structural model, a surprise action by one actor can have a simultaneous impact on another actor, but that the targeted actor's unforecasted behavior can only influence the actors with some pre-specified time lag. While this assumption may be useful,we believe that there is no theoretical or practical reasons to assume conflict dynamics form either simple bivariate one-way causal chains, or recursive contemporaneous relationships. Here, we offer an alternative, non-recursive window into reciprocity and other conflict dynamics.[4]

---

[3]The only exception to this assumption that we are aware of can be found in McGinnis and Williams (2001).

[4]Their is considerable confusion surrounding the idea of endogeneity and Granger causality in vector autoregression analysis. The usual practice is to estimate a reduced-form model and then produce impulse response functions to trace out the implied dynamics. In order to give the resulting coefficients and plots a political meaning, assumptions need to be made that map the estimated reduced-form into an interpretable structural-form. A recursive structure assumes that the contemporaneous shocks or innovations are orthogonal and enter the model one equation at a time

Second, what explains deviations from tit-for-tat behavior? Why do some leaders such as Gorbachev during the twilight of the Cold War, step out on a limb and offer cooperation to a rival? Why do other leaders feign cooperation and then attack an adversary as in Operation Barbarosa during World War II? Deviations from reciprocity reflect dramatic historical events. If this unexplained variance follows a predictable pattern, what theories can help us improve our predictions?

Finally, reciprocal behaviors may be asymmetric across groups with one government or authority reacting more strongly than another (Ward, 1982). What explains this action-reaction mismatch? Since democratic leaders must pay attention both to foreign and domestic conditions for political survival (Bueno de Mesquita et al., 2003), and likely rivals know this fact, public foreign policy preferences may explain the reciprocal behaviors and unilateral deviations and dyadic asymmetry.

2.2 The Intersection of Domestic and World Politics

A branch of research extending reciprocity models explains conflict and cooperation as a function of domestic institutions and politics. Putnam (1988), Fearon (1994, 1998), and Bueno de Mesquita et al. (2003) view world politics as a two-level game where elites sit at the insticese of foreign and domestic politics. Leaders must worry about whether their constituents are willing to ratify agreements they negotiate with other groups or support their foreign policies. Within identifiable institutional contexts citizens can hold leaders accountable for their foreign policy choices. Policies deviating from public preferences may be rejected and lead to a leader's electoral punishment. A potential repercussion of this accountability process is that leaders who can generate high audience costs for their policies, thus being punished if their foreign policy fails, can more credibly

based on a triangularization of the error covariance. The confusion arises because most studies also use Granger causality tests to specifically deal with two-way relationships. However, the only two-way relationships that are directly modeled in a Granger causality framework are those where one of the variables operate on the other with a lag. This is why the speed of adjustment (as determined by a subsequent impulse response analysis) in a Granger causality framework is important. But for such an impulse response analysis, we need to make identification restrictions to understand the simultaneous effects. Recursive models represent one set of identification restrictions–a Wold causal chain (Brandt and Williams, 2007, for details, see). Despite their convenience, recursive structures do not necessarily lend themselves to tests of conflict behavior such as reciprocity, accountability and credibility (see below).

communicate resolve in international conflict situations (Fearon, 1994; Martin, 2000). For this reason, some international relations forecasters argue that scenario generation must anticipate public opinion changes (Sylvan, Keller and Haftel, 2004; Sirriyeh, 1995).

Research on the causes of reciprocal behavior suggests that domestic preferences play an important role explaining conflict. McGinnis and Williams (2001) aver citizens' preferences constrain leaders' tit-for-tat interactions. In their "rivalry-as-prison" or Goldilocks hypothesis, the public prefers conflicts to remain in some limited range of interactions. Policies that are too dramatic in either a cooperative (too cold) or bellicose (too hot) direction are viewed as domestically costly and may lead to electoral punishment of the leader. Under these constraints, leaders spend their time reciprocating each other's limited cooperative or bellicose policies. Guisinger and Smith (2002) and McGillivray and Smith (2000) hypothesize that in democracies the public prefers to pocket the gains that accrue from reciprocated cooperation rather than to defect for a one-time payoff that makes long term cooperation impossible. The public is then expected to punish a leader deviating from reciprocity. In these explanations, the public constrains a leader's behavior and supports reciprocal outcomes. This begs the question: can deviations from reciprocity be explained by changes in the public's foreign policy preferences? If so, then public preferences on foreign policy processes could aid in forecasting and understanding international conflict.

2.3 Counterclaims: A Flock of Followers

All scholarly voices do not sing the praises of domestic political explanations. Lippmann (1922), Almond (1965), Morgenthau (1967), and Rosato (2003, pg. 599) question the public's ability to comprehend, process and intelligibly guide foreign policy. Because of knowledge and interest deficits between the public and foreign policy elites, it is hypothesized that leaders can manipulate citizens. These authors point to the rally-'round-the-flag effect as evidence that when a leader tells the public to jump, approval jumps. Thus a leader is not constrained by the public since citizens will follow and rally to the government.This follower model predicts that information on public preferences is not useful in understanding policy.

Evidence on the rally-effect and public knowledge deficits is mixed. While there is some evidence that a rally-'round-the-flag effect exists in the U.S. and Britain, this does not mean that public preferences are meaningless in predicting international events in these cases. Colaresi (2007) shows that the rally varies considerably from crisis-to-crisis in ways that can be predicted from an accountability-based signaling model of foreign policy decision-making. Holsti (1996), Page and Shapiro (1995), Aldrich, Sullivan and Borgida (1989) argue that public opinion is more stable and reasonable than either Almond or Lippmann credit. The empirical question remains, in foreign policy who is the leader and who is the follower?

Challenging the public-as-follower model, Wlezien (1996), Eichenberg and Stoll (2003), Baumgartner and Jones (2005) show that in many circumstances not only do leadership cues fail on defense issues, but the public reacts oppositely. When holding public preferences constant, and a leader changes his policy toward conflict, the public is likely to react not by changing its mind to support more conflict but by pulling in the opposite direction and supporting more cooperation. This is identified by Baumgartner and Jones (2005) as a negative feedback mechanism. This negative feedback is consistent with the McGinnis and Williams (2001) Rivalry-as-Prison effect since the public hems in policy rather than blindly following leadership cues.

There are competing claims about the effects of foreign policy change on public preferences in the public-as-follower model and the proposed negative feedback mechanism. The public-as-follower model suggests that the public is attracted to a leader's policy. A leader's shift to a more cooperative international policy creates more cooperative public preferences. Alternatively the negative feedback mechanism of Baumgartner and Jones (2005) hypothesizes that the public will be repulsed by a change in policy. Instead of following a cooperative policy, it is likely to move in a bellicose direction and holds its leaders accountable for their policies.

2.4   On the Other Side(s): Cross-level Credibility

Rivals and potential mediators may base their policies on the accountability dynamics of their adversary. Putnam (1988), Fearon (1994), and Martin (2000) suggest that leaders look across the

water's edge to gauge the credibility of any bargaining strategies. An important component of Putnam's two-level game framework is that leader B will examine the public preferences constraining leader A to determine if an offer/action is a bluff. The greater the constraints on leader A from leader B's perspective (the farther the public preferences of A are from B's ideal outcome), the greater the credibility of an uncompromising offer from leader A. If the relevant public is unsupportive of peace, regardless of the offers made by leader A, leader B does not have an incentive to offer extreme cooperation.[5] Even if the other side offered to reciprocate, it is unlikely that this cooperation would be ratified. Shlaim (2001) notes that Nasser avoided cooperative gestures toward Israel for fear of both his own public's negative reactions, *as well as the Israeli public's negative reactions*. Conversely, as a public becomes more supportive of collaboration, a peace-seeking rival may suggest greater cooperation, since citizens are now more likely to ratify any agreement. [6] But the dynamics of credibility may not be unidirectional. Page and Shapiro (1995), and Holsti (1996) suggest that the public bases their support of or opposition to cooperation on available information, including a rival's actions.[7] Figure 1 summarizes the causal linkages associated with the ideas of reciprocity, accountability, and credibility for a case where public opinion is available for one of the countries, as we have in our case study.

[Figure 1 about here.]

2.5   From Theory to Models: Problems and Solutions

In this section the preceding arguments are organized into four models of international conflict, which are then applied to the Israeli-Palestinian case. The translation of these theories to testable models of conflict and cooperation is nontrivial. These models suggest a highly dynamic, endogenous, and large scale data generation process. For example, reciprocity, triangularity, accountability and credibility relationships suggest that international actions and public preferences

---

[5]This only holds if public preferences are negatively or unrelated to a rival's actions (see below).

[6]Note that if accountability does not hold and the public are followers/bystanders, then credibility is irrelevant.

[7]Putnam (1988) described this dynamic as foreign-to-domestic reverberation. Rivals that cooperate (attack) may signal to the public that cooperation is safe (unsafe) and (not) profitable.

are dynamic. Previous changes in policy/preferences are likely to affect *current* changes in policy/preferences. These frameworks also suggest that current changes in many variables simultaneously affect other variables. If a leader is worried about political punishment for deviating from public preferences, a surprise change in those public preferences may lead to a simultaneous change in policy to bring conflict and cooperation into line with that new information. Similarly, a watchful public might alter its opinion about a conflict immediately after witnessing a surprise change in a rival's policy toward its country. The temporally aggregated data used in many existing studies often masks this kind of public reaction (thus increasing the role of endogeneity). Finally, focusing on the directed-dyadic behavior of two rivals, allowing for great power intervention, and taking into account the possibility of accountability and credibility, produces a model that includes at least seven or more equations. Allowing lagged effects of variables, the number of parameters in the model will grow exponentially as we add relevant actors and more equations.

To cope with these problems, the earlier ideas are presented as four competing sets of relationships. Schematically, Figure 1 is translated into Table 1. Here the possibility of the different "two-way" interactions of reciprocity and credibility are analyzed. The model of interactions assumes there is a separate behavior equation for each of the dyadic interactions between rivals. In the table, the foreign policies of actor 1 (2) toward actor 2 (1) is denoted by $\text{Policy}_{1 \to 2}$ ($\text{Policy}_{2 \to 1}$), and public opinion in the political jurisdiction of actor 1 as Public_1. The column labeled $\text{Policy}_{Z \to 1}$ represents outside mediation by Z towards 1. The rows are the behavioral relationships (equations) to be modeled for each of the four theories of rivalry and public interactions. The columns are the independent variables, the changes or "shocks" in policy and opinion. A "Yes" represents an expectation that the column variable will influence the row equation, either simultaneously or with a lag, and a "No"-entry indicates an expected zero restriction.

The first set of ideas, the *bystander model*, contends that the public plays no role in foreign policy. Here the public pays little attention to conflict and cooperation events. In this view, previous scholarship on reciprocity and triangular relationships is correct in ignoring public opinion in its empirical specifications. This thesis predicts that 1) a surprise change in conflict and cooperation

either sent or received by one actor will have no systematic affect on public preferences and 2) a shift in the public's preferences will not lead to changes in the policies of its home government toward an adversary.

The second model allows the public to follow but not lead foreign policy toward a rival. This *follower model* mirrors the logic of the patriotic rally-'round-the-flag literature. The public's government acts towards an adversary and the public supports that action. Public preferences react to government policy but government policy does not react to changes in public preferences. Therefore, changes in foreign policy should occur independent of public preference shifts.

In the third model, the public not only reacts to foreign policy but also can lead that policy. Under this *accountability model* government policy reacts to changes in public preferences. If public support of conflict or cooperation toward a rival changes, its government's policies reflects this in some way. This third perspective contends that the public is paying attention not only to what their government does (actor 1), but also to what the rival (actor 2) may be doing. Thus there should be a bi-directional relationship between conflict and cooperation and public opinion.

The final model incorporates the belief that if accountability relationships exist within one actor, its rival is likely to use that information to formulate its own foreign policy toward that country. In this *credibility model* public preferences in one rival are likely to affect the policies of other units towards that rival. If the public has relevant ratification power, as suggested by the accountability model, that information is unlikely to be ignored by the rival.

[Table 1 about here.]

Another process relevant to both accountability and credibility is external third-party intervention (rival Z in Table 1). Following previous work on triangularity, localized conflicts are likely to react to great power interventions by encouraging or discouraging cooperation between the belligerents. To keep the analysis here simple, we assume that the public does not react to the policies of foreign governments in the same way it reacts to the policies of its own government. On the other hand, the credibility model suggests that mediators are likely to take public opinion into ac-

count when calculating the expected costs and benefits of cooperation and conflict. (Goldstein and Freeman, 1990; Brandt and Freeman, 2006).

A key implication of these four models is that in any analysis with temporally aggregated data, the theories of reciprocity, accountability and credibility can not be represented by either a single equation or a recursive multiple equation model. The hypothesized relationships move in two directions and are contemporaneous, not solely lagged. The foreign policy of one group is likely to react to the policies of the rival, and *vice versa.* A recursive model would only allow one rival to react without a lag. Further, the accountability and credibility models predict that the public both leads and responds to foreign policy actions. A recursive model would necessitate contemporaneous and atheoretical zero restrictions on the response time of either public opinion or foreign policy. This of course raises three questions that we answer in the next two sections. First, where is information available that allows us to systematically test the hypothesized accountability and credibility models versus the others? Second, even if we located both relevant foreign policy and public opinion data, how can we empirically analyze these hypothesized non-recursive structural models? And third, how do we test foreign policy and two-level game dynamics in a model with contemporaneous effects?

3 Research Design

3.1 The Case

We analyze Israeli-Palestinian interactions as our test case. This protracted conflict is one of the most enduring of our time. Its significance is widely recognized; *The Economist* (February 21, 2004: 24) writes that this conflict is "where the world's fault-lines meet: divides between rich and poor, secular and religious, Islam and the West."

Political scientists have studied this conflict for many years. Goldstein et al. (2001) find a complex set of behavioral relationships, including evidence of reciprocity between the two rivals as well as of "triangularity" in U.S. behavior toward the Israelis and Palestinians. Schrodt et al.

(2001) report similar findings about triangularity. The role of the U.S. in this conflict is echoed in other quantitative works like Organski and Lust-Okar (1997) and qualitative work like Sirriyeh (1995). Organski and Lust-Okar include the U.S. as a moderator while Sirriyeh bases his qualitative forecasts for Israeli-Palestinian relations on U.S. intervention.[8]

These studies do not consider the possibilities of accountability or credibility. This is in spite of the fact that leaders like Israel's Yitzhak Rabin commissioned and examined polling data (Auerbach and Greenbaum, 2000). Rabin apparently sought to learn from the polls how best to build credibility at a crucial turning point in the conflict, during the secret negotiation of the Oslo Accords. Existing studies thus cannot help sort out the competing theories discussed above.

The two rivals have somewhat different electoral histories. Israel is considered a mature democracy by POLITY and Freedom House. Since the first Oslo agreement major efforts have been made to create a democratic Palestinian state (Brown, 2003). Elections for the Palestinian Executive Authority and the Palestinian Legislative Council were held in 1996. In late 2005 some local Palestinian elections were held and Hamas won an upset victory in Palestinian Legislative Council elections in January 2006 (national election). So now, electoral forces are relevant for both rivals. In the analysis that follows we study the dynamics of the Israeli-Palestinian conflict up to the spring before these recent elections.

3.2 Data

Like Goldstein et al. (2001) and Schrodt et al. (2001), we use events data to measure the directed behaviors of the Israelis, Palestinians, and the U.S. Events for directed behaviors among the U.S., Israel, and the Palestine (WEIS coding) were extracted from Kansas Event Data System (KEDS) from *Agence France Presse* news stories.[9] The events are scaled with Goldstein scores and aggre-

---

[8]Sylvan, Keller and Haftel (2004)'s experts place less emphasis on the role of the U.S. in this conflict, but acknowledge that their experts might take the role of the U.S. for granted in generating their forecasts.

[9]Results reported here are similar if we use the KEDS events coded from Reuters' news wire reports. On the importance of employing multiple sources see Shellman, Reeves and Stewart (2006).

gated into monthly averages.[10] The variable mnemonics used to represent the directed behavior denote the governments of Israel, Palestine, and the U.S. by I, P, and A, respectively. So $A2P_t$ represents the level of conflict/cooperation directed by the U.S. toward the Palestinians at time t. Our estimation sample is monthly from April 1996 to March 2005. We reserve the remaining 2005 data for out of sample forecasting. Thus, our analysis stops before the recent Israeli and Palestinian elections, the election victory of Kadima, and the election of Hamas.[11]

In addition to being politically important and salient, the Israeli-Palestinian conflict provides a unique source of relevant public opinion data that allows an analysis of accountability and credibility relationships. The public opinion measure is from polls conducted by the Tami Steinmetz Center for Peace Research (TSC). The TSC is a multidisciplinary academic enterprise composed of faculties at Tel Aviv University. These polls should be less prone to the journalistic biases that has been found in U.S. polling (Gaubatz, 2001). We use the TSC's peace index for Jewish respondents only. This index does not yield the kind of detailed insights about specific policies that were obtained from polls like those which Rabin commissioned. But it provides a continuous sounding of the Israeli public's evaluation of their governments effort to create peace and of the prospects for peace.[12] A few of the peace index observations are missing — out of the 108 months in the

---

[10]Monthly averages of Goldstein-scaled events are employed because they place the event data on a scale similar to the public opinion data (which is the average of respondents in the polls) and because we believe that policymakers are concerned with deviations from the average level of an on-going conflict. The dynamics of the average and total Goldstein scaled data are similar. For an analysis using totals instead of averages see Brandt and Freeman (2006).

[11]While the post-December 2005 period is of keen interest, given the short time frame there is not much that could be said definitively for this period since there is little event data (less than 24 months) at the current time.

[12]The peace index is composed of two questions. The first is, "In general do you consider yourself a supporter or opponent of the peace process between Israel and the Arabs?" The possible responses : greatly opposed (0), somewhat opposed (1), in the middle (2), somewhat supportive (3), greatly supportive (4) and don't know/no opinion. The second question is: "Do you believe that in coming years there will be peace between Israel and the Arabs?" The responses are: certain there will be peace (4), think there will be peace (3), in the middle (2), think there will not be peace (1), certain there will be no peace (0), and don't know/no opinion. The two scores for each respondent are averaged and then multiplied by 25. So each person's final score is between 0 and 100. The index's monthly values are averages over about 500 respondents per survey. It is important to acknowledge that a slight wording change recently was made

estimation sample, eight have missing values in 2003.[13] The missing values were inputed via an ARIMA model. The Jewish peace index, denoted as JPI_t, and the six monthly average event data series are shown in Figure 2.

[Figure 2 about here.]

To capture coalitional, electoral forces, and trends in violence nine exogenous dummy and trend variables are employed. Three of these are dummy variables for the identities of the Israeli prime ministers in each month (one each for Netanyahu, Barak, and Sharon, with Rabin/Peres treated as the reference category). These variables capture changes in conflict dynamics due to leader (coalitional) idiosyncrasies in Israel (Sylvan, Keller and Haftel, 2004; Sirriyeh, 1995). For each of the four prime ministerial regimes a separate time counter that starts at the value 1 in the month after each Israeli election and increases monotonically until the time of the next constitutionally mandated election (and is zero otherwise) is used. These four trend variables capture electorally-motivated cooperation and especially electorally-motivated violence in the run-up to elections.[14] Finally, two dummy variables are included to capture changes in the trends of the mean level of conflict. The first is for the period from the start of the second Intifada to the start of the Battle of Jenin (October 2000–April 2002). The second is a dummy variable for the post-Battle of Jenin period (May 2002–March 2005).[15]

_____

in one of the questions with the term "peace process" replaced by "negotiations."

[13]This imputation uses an ARIMA(9,0,0) model for the data before the break. The ARIMA specification was chosen as the most parsimonious fit using the Box-Jenkins method.

[14]The election counters are reset at the May, 29, 1996 general election, the May 17, 1999, February 2, 2001, and January 28, 2003 direct prime ministerial elections. Another election was held on March 28, 2006, which is outside the sample. Illustrative of electorally-motivated violence were the suicide bombings against Israeli civilians that Hamas reportedly engineered just before the 1996 election. These bombings were intended to help defeat Peres and help the more hawkish Netanyahu. *The Economist* January 28, 2006, p. 11. The time counters create linear drift in the (moving) behavioral equilibrium among the three actors.

[15]The Battle of Jenin occurred in April of 2002 and involved the largest scale of military force in the West Bank since the Six Days War.

3.3 Model

A structural Bayesian time series approach is used to evaluate the four theories in Table 1. This approach addresses the problems of model scale, endogeneity, and specification uncertainty. The appendix describes the actual model, a Bayesian Structural Vector Autoregression (B-SVAR). A fuller explication this model and explanation of its value in macropolitical analyses can be found in Brandt and Freeman (2006, 2007). The B-SVAR model has seven equations—six for the dyadic interactions for the Israelis, Palestinians, and U.S., and one for the Jewish opinion about the peace process. Test statistics support using two lagged values of each of these seven endogenous variables.[16] Each equation in the B-SVAR includes the nine exogenous variables described earlier. The next sections describe elements of the empirical model that allow us to capture the essence of the theoretical models.

3.3.1 Structural Identification

The theories about reciprocity, accountability and credibility can be represented as competing claims about the *contemporaneous* and *lagged* relationships among seven variables. The claims about contemporaneous relationships concern the speed of the response of some variables to shocks in others, especially about 1) the immediate effects of shocks in Jewish opinion on directed behaviors of Israel and Palestine, and 2) the immediate reaction of Jewish opinion to these behaviors. Operationally, these claims are restrictions in the matrix of coefficients for contemporaneous relationships in the model. The four theories allow for the possibility of reciprocity and triangularity between the Israelis, Palestinians and the U.S. since no restrictions are imposed on any of the respective coefficients for these contemporaneous relationships. The theories do differ in the zero restrictions they imply for the contemporaneous relationships for Jewish opinion.

Previous empirical models that purport to analyze reciprocity, accountability and credibility relationships must account for these hypothesized two-way relationships. For comparison, we

---

[16]F-statistics and Akaike information criteria support using 2 lags. We also looked at the posterior distribution summaries for models with more lags and found that they fit worse than two lag models.

present a recursive structural model of the variables of interest at the top of Table 2. This is the default pattern of identification restrictions used in the interpretation of event data analyses and existing vector autoregression models.[17] The sets of seven rows each in Table 2 represents the contemporaneous relationships among the variables of interest (the A_0 matrix discussed in the Appendix) for the distinct contemporaneous identification schemes. The rows of Table 2 correspond to the equations for each directed-dyad or Jewish public opinion equation. The columns are the contemporaneous residuals that enter each equation. The X's (or other letters) in the cells represent the "free" parameters or those estimated while the empty cells are zero restrictions. A zero restriction indicates no contemporaneous relationship is hypothesized between a column variable residual and a given row equation. Note that in the recursive scheme all relationships above the diagonal are assumed to be zero. This implies that we have no expectation of any contemporaneous two-way relationships, or that the errors propagate in a recursive chain through the equations.. To the extent that there is a contemporaneous, within month for example, correlation between changes in one equation and another, that correlation is assigned by the recursive model in *one direction only*.

All four of the other models imply a constellation of zero restrictions and "free" parameters that make the matrix of contemporaneous relationships non-recursive and over-identified.[18] Consider first the bystander model. Like the others, this model allows for contemporaneous reciprocal and triangular relationships in directed behaviors. The bystander model holds that the Israeli public does not react immediately to the behavior of its own government nor of the Palestinians and, at the same time, that none of the governments react immediately to changes in Jewish public opinion (as implied by the empty cells in the last column and the last row). The public does not pay attention to the behaviors of the three governments and the three governments ignore changes

---

[17]For example programs like STATA and RATS assume a recursive ordering of variables when orthoganalized impulse response functions are calculated for a reduced form model, like those in the current literature.

[18]In this seven equation (variable) model given the formulation of the B-SVAR, there are at most 28 free parameters in this matrix of coefficients, A_0. All four models imply fewer than this number of free parameters, or more zero restrictions than are necessary to identify the model.

in public opinion. The absence of these contemporaneous relationships are represented as blanks are the zero order restrictions imposed by this first theory.[19]

[Table 2 about here.]

The other theories allow for additional contemporaneous relationships in the a B-SVAR model. These theories are presented in the remaining three model blocks of Table 2. Moreover, they are nested: the follower model is implied by the accountability model which is implied by the credibility model. The block of relationships for each model are thus similar to the bystander model for the contemporaneous reciprocal and triangular behaviors. The follower, accountability and credibility models differ by allowing additional contemporaneous relationships. The possibility that Jewish citizens monitor and react immediately to the actions of their government toward the Palestinians but the Israeli government does not react to shifts in Jewish opinion is represented by an F in the seventh row and first column of the follower block. This relationship and the possibilities that Jewish citizens react immediately to the actions of the Palestinians toward the Israelis and that their government reacts immediately to shifts in Jewish opinion are denoted by the A's in the accountability block. Finally, the theory that all these contemporaneous relationships exist and that 1) the Israeli, Palestinian, and American governments monitor and react immediately to shifts in Jewish opinion and 2) Jewish opinion immediately reacts to the directed behaviors of the Israelis and Palestinians are denoted by the C's in credibility block of Table 2. Each of these blocks is the structural identification for one of the four B-SVAR models presented below.

3.3.2 Dynamics

The lagged relationships between the variables describe how changes in directed behavior and opinion are related through time. What distinguishes the B-SVAR model is that these relationships explicitly depend on the contemporaneous causal relationships. The responses of the system to

---

[19]Of course, the data may not move the free coefficients denoted by the X's off zero if there are no such relationships. But the zero order restrictions impose zeros in the posterior for the coefficients denoted by blanks regardless of the information in the data.

changes in the endogenous variables or shocks, as revealed by an analysis of its reduced form, reflects both the lagged relationships and the (competing) restrictions on the contemporaneous relationships between the variables.[20]

Unlocking the dynamics of the system is difficult due to the problem of scale. The model contains a large number of parameters and the parameter uncertainty makes it difficult to draw causal inferences and policy analysis.[21] To cope with this problem, we employ a modified version of the Sims and Zha (1998) prior. This prior addresses model scale by putting lower probability on coefficients of the lagged effects, especially those at the most distant (largest) lags. But rather than imposing exact (possibly incorrect) restrictions on these coefficients such as deleting lagged variables, the prior imposes a set of inexact restrictions on the lagged coefficients. These inexact restrictions are prior beliefs that many of the coefficients in the model—especially those for higher lags—have a prior mean of zero and small variances. The prior on the model coefficients is then correlated across equations in a way that *depends on the contemporaneous relationships among the variables*. This allows beliefs about the structure of the system to be included in the prior. Finally, the prior is centered on a random walk model, allowing beliefs about degrees of persistence in behavior.[22] Details of this prior are in Sims and Zha (1998) and Brandt and Freeman (2006, 2007).[23]

---

[20]If A_+ is the matrix of coefficients on lagged values of the variables in the structural model and A_0 is the matrix of coefficients for the contemporaneous relationships among the variables, then the reduced form coefficients of the model are $B = -A_+A_0^{-1}$.

[21]The weekly model of Goldstein et al. (2001, 607, fn. 33) has 24 variables each with nine lags. Thus, it has 217 coefficients per equation or more then 5000 coefficients overall. Because of such scale, meaningful causal inferences about the dynamic responses of their system will be difficult to make. For this reason investigators usually do not even try to analyze dynamics, conduct innovation accounting or produce forecasts. There is good reason to believe that most of the coefficients for the high frequency dynamics in their model are close to zero. To ignore this belief is to use a *flat prior* or allow the prior variance on the coefficients on all lags to be the same.

[22]This prior serves as a benchmark and does not mean that the data follow random walks. If it is inconsistent with the data, the estimated posterior will reflect this.

[23]The hyperparameter values for the Sims-Zha prior were set at values based on experience with events data and discussions with leading international relations scholars like Philip Schrodt (Brandt and Freeman, 2006). The hyper-

3.3.3 Model evaluation

Model scale also creates challenges for model evaluation. Complex, highly parameterized models are bound to overfit the data making conventional (frequentist) fit statistics less useful. The current time series literature and Bayesian statistics employs a Bayesian measures to summarize posterior model fit (Sargent, Williams and Zha, 2006; Sims and Zha, 2006; Brandt and Freeman, 2007). This measure, the log marginal data density (log MDD) measures the log posterior density for the sample data (Chib, 1995). For the log MDD measure, one prefers larger values, since they mean that the model has a higher posterior probability of generating the data. This measure also accounts for the differences in the number of structural parameters in the various models because it integrates over each parameter in the computation of the marginal likelihood—so models with excess parameters will be penalized if they do not fit well. Differences in log MDDs across models are log Bayes factors which compare the posterior odds of a model to its prior odds or the weight of the evidence for one model over another. The log Bayes factor allows us to compare support for the four models in a manner that is consistent with our statistical approach (Kass and Raftery, 1995; Gill, 2002).[24]

Equally important, we analyze the implied dynamics of each model to see which produces the most plausible set of key impulse responses. Using Bayesian methods provides meaningful error bands for these responses. For example, if the bystander model receives strong support, its log(MDD) value should be larger than the log(MDD) values for the other models. And if the bystander model receives the most support, the error bands for the responses of JPI to shocks in both I2P and in P2I as well as the error bands for the responses of I2P and P2I to shocks in JPI should include zero.

---

parameter values used in all four models were $\lambda_0 = 0.8$, $\lambda_1 = 0.1$, $\lambda_2 = 1$, $\lambda_3 = 1$, $\lambda_4 = 0.1$, $\lambda_5 = 0.05$, $\mu_5 = 0$, and $\mu_6 = 5$. Other values yield qualitatively similar results and inferences to those reported here. These results are available from the authors.

[24]It is important when using log Bayes factors to assess the sensitivity of the results to the prior specification for each model. In the present case, the posterior is insensitive to the choice of the prior hyperparameters. Also, the log MDD results are similar across the two MCMC chains for each model.

For the best fitting model, two sets of out of sample forecasts from April–December 2005 are presented. One set of forecasts includes the JPI series, while the other omits this series from the model. These forecasts are original in terms of their time span and because past studies of the Israeli-Palestinian conflict have not been able to produce specific timing predictions. They suffer from "off-on-timing" problems (Bueno de Mesquita, 1997). Our use of Bayesian methods allows us to provide meaningful error bands for these forecasts as well.[25]

4   Results

4.1   Model fit and selection

Table 3 presents the posterior fit measures for the five different structural models—recursive, bystander, follower, accountability, and credibility—outlined previously. Moving down the rows of this table, each model contains more contemporaneous parameters (in the its A_0 matrices), with the exception of the recursive model. The best fitting model, with the largest log MDD measure, is the credibility model. The final four columns of Table 3 are the log Bayes factors. There is are difference between the recursive, bystander, follower, and accountability models (log Bayes factors less than 2 are weak evidence for one model versus another). The credibility model has large Bayes factors when compared to the other four models, with values between 6.72 and 17.08. This is very strong evidence for the credibility model over the other contemporaneous specifications.[26]

[Table 3 about here.]

---

[25]Measure of forecast uncertainty (error bands) are absent in the work of the Political Risk Service (PRS), International Crisis Group (ICG), and in more sophisticated analyses using hidden Markov models (Schrodt, 2000; Schrodt and Gerner, 2000).

[26]We also ran models that 1) excluded the exogenous control variables from the B-SVAR models and 2) included more lags of the endogenous variables. These models produce the same inference, namely that the credibility model is superior to the other specifications.

In terms of model fit the credibility model offers the best explanation of the data. The restrictions on the contemporaneous impacts of Jewish support for the peace process implied in the bystander, follower, and accountability models are not as likely as those in the credibility model.

The reason for the better fit of the credibility model is due to its ability to capture the contemporaneous reciprocity, accountability and credibility relationships among the belligerents and the Jewish public. Next, we turn to looking at the dynamics that follow from these contemporaneous effects. This allows us to look at the direction of the responses to unanticipated shocks to the model.

4.2   Impulse response dynamics

While the fit of the credibility model is best, we are also concerned with the conflict and opinion dynamics of this model. Impulse response analysis, tracing out the response of the system of equations to shocks in selected variables, summarizes the complex dynamics of the endogenous variables and the estimated structural models. The full set of impulse responses for the systems estimated here include 245 responses (49 responses per model $\times$ 5 models). Rather than present all these responses, we focus on those of greatest substantive interest: the responses of I2P and P2I to American actions and Jewish public opinion shocks and the responses of Jewish public opinion to conflict shocks. These are the most theoretically relevant comparisons across the contemporaneous specifications in Table 2. The starting points for the impulse responses are the contemporaneous effects identified in the last section (via the identification of A_0^{-1} for each model). The results presented in this section thus allow us to compare the dynamic implications of the model specifications.

Given correlated uncertainty about the dynamics of these models, the error bands for the impulse responses are computed using an eigenvector decomposition method (Brandt and Freeman, 2006).[27] The reported error bands are 90% (posterior) confidence regions around the median es-

---

[27]The eigenvector decomposition of the variance of the impulse responses decomposes the variation of the responses over time. This is a better summary of the overall shape, skewness and location of the error bands since it accounts

timates. They provide a summary of the general trend and shape of the responses. The impulse responses presented here are for positive shocks to the B-SVAR model system of equations. The responses are presented for positive shocks since the identification and estimation are invariant to the signs of the shocks. Positive shocks are chosen since they can naturally be interpreted as "surprise" movements towards peace, less conflict, or increases in the support for the peace process.

Figure 3 are the responses of the I2P, P2I, A2I, and A2P equations to shocks in the same variables. The rows in this figure are the responses of these equations to shocks in the column variables.[28] The bystander (credibility) model responses and error bands are shown with solid (dashed) lines. To interpret the dynamics, start with the bystander model (solid lines). The response of I2P to a positive shock in P2I is positive and its error bands do not include zero. The bystander model thus implies reciprocity by the Israelis to surges of cooperation by the Palestinians. Now consider the second row of the figure for the bystander model where a shock in I2P is positive. Here the response in P2I is briefly positive but the error bands soon span and include zero. This indicates, implausibly, that the Palestinian response to Israeli conflict innovations is very short-lived to nonexistent.

[Figure 3 about here.]

Responses for the credibility model are more consistent with knowledge of the Israeli-Palestinian conflict. Reading across the top row of Figure 3 for this model (dashed lines), the shock in P2I is positive and the response again is reciprocal or weakly negative in I2P. But row two shows that for this model a positive innovation in I2P produces a clear, sustained inverse response in P2I. This finding that only the more democratic member of the conflict dyad reciprocates cooperation (and conflict) is consistent with the work of Guisinger and Smith (2002); McGillivray and Smith (2000) who suggest that democratic institutions create incentives for leaders to use reciprocal strategies.

---

for the correlation in the responses. The variation explained by the first eigenvector of each response is between 65% to 100% of the variation in the impulse responses. (for details see Brandt and Freeman, 2006). Error bands computed from the empirical percentiles of the posterior responses generate similar inferences.

[28]These are one standard deviation shocks of the residuals from the respective equation in the B-SVAR model.

Conversely, their model implies that when the public finds it difficult to hold leaders accountable, elites may defect from cooperation. Therefore, we expect to see a pattern of non-democratic conflict in response to cooperative gestures from a rival. This "tit-for-take-that" behavior is consistent with an interpretation of Palestinian policy making that emphasizes the weakness of their central authority. Even for the data from before the Hamas election victory in 2006, these result show that when Israel cooperates the Palestinian Authority fails to rein in militants who might attempt to scuttle the peace process. When Israel is hostile, the Palestinians can not escalate in kind due to the their lesser military capabilities. These dynamics are only apparent when underlying domestic accountability and credibility mechanisms are explicitly modeled.

Also this figure demonstrates some evidence of triangularity. Surprise cooperation by the Americans generates nearly no response by the Israelis towards the Palestinians or by the Palestinians towards the Israelis (per the upper right two-by-two set of responses). American reaction towards Israeli and the Palestinians (the bottom left two-by-two set of responses) are generally positive and with a lag. This is evidence that U.S. actions do not drive changes in I2P and P2I, but that American actions are supportive of surprise cooperation by the Israelis and Palestinians.

While the reciprocal and non-reciprocal responses in Figure 3 are consistent with models of credibility and accountability, they do not directly demonstrate accountability. Accountability in the dynamic model presented here means that changes in Jewish public opinion respond to changes in conflict (and vice versa). Figure 4 shows the responses of Jewish support for the peace process (JPI) to unexpected positive or cooperative one standard deviation shocks in Israeli actions toward the Palestinians and vice versa. The solid (dashed) lines are the responses from the follower (credibility) model. For the follower model, there is a weak negative response—indicating an inverse public reaction to increased cooperation. In the credibility model response in Figure 4, the median response in JPI to a positive shock in I2P is more than twice as large as that of the follower model. The follower model allows only for a public reaction to the level of conflict, while the credibility model also includes contemporaneous relationships between Jewish opinion and the behaviors of both belligerents. The reason for the different magnitudes of the responses is omitted endogene-

ity. In the credibility model which allows endogeneity between the Jewish support for the peace process and conflict, the presence of accountability and credibility mechanisms generates a strong public opinion response to changes in the level of cooperation. The response is the opposite of what a follower model of public opinion would expect. Instead of echoing elite policy, public opinion constrains and reacts inversely to policy, supporting Wlezien (1996), Eichenberg and Stoll (2003), and Baumgartner and Jones (2005).

[Figure 4 about here.]

The final component of the reciprocity-accountability-credibility causal linkage is the dynamic response of the Israelis and the Palestinians to changes in Jewish opinion about the peace process. If the Palestinian reaction to the Israelis (P2I) responds to shocks in public opinion (JPI), then this demonstrates credibility. Figure 5 shows the the accountability and credibility models' impulse responses for the I2P, P2I, A2I, and A2P equations for innovations in Jewish opinion. The main difference between of the credibility model and the others are the impacts of surprise changes in JPI on the other equations in the system that represent triangular relationships, specifically the impacts of JPI on I2P, P2I, A2I and A2P. In figure 5 the rows are the responses of the I2P, P2I, A2I and A2P equations to a surprise increase in Jewish support for the peace process. The comparison here is between a model that allows for contemporaneous political accountability (but not credibility) and a model that allows for both political accountability and cross-rival credibility.

[Figure 5 about here.]

The error bands for the response of Israeli policy toward the Palestinians to an innovation in JPI include zeros for the accountability, but not the credibility model.For the accountability model increases in Jewish support for the peace process produce no Palestinian response toward the Israelis over 24 months. In contrast, for the credibility model, an innovation in JPI initially produces a lagged response, since the P2I response and its error bands drop below zero after five months. Therefore, in a model that ignores the two-way relationship between opinion innovations and international actions, the congruence between Jewish opinion and Palestinian policy toward Israeli is

incorrectly estimated. On the other hand, in the credibility model—where the instantaneous feed-back between opinion and events is explicitly modeled—a cross-level (domestic to international) pattern of asymmetric interaction is evident. The Palestinians react with a slight lag to public support for peace with conflict and opposition to peace with cooperation. This response, like the asymmetrical reciprocity finding discussed above (Figure 3), is that Palestinian actions counteract rather than accelerate changes in public opinion.

Finally, there are similar responses on U.S. reactions to the belligerents when Jewish support for the peace process increases. The U.S. response toward Israeli cools when Jewish support for the peace process increases; in contrast U.S. reaction toward the Palestinians is initially negative, but then rebounds and is cooperative. This latter reaction is slower and smaller in the credibility model.

4.3 Forecasts

One way to understand the implications of public opinion in a two-level game of conflict is to look at the forecasting performance of models with and without a public opinion variable. This final exercise is a comparison of the reduced form Bayesian VAR forecast performance of the models used here. Instead of comparing different contemporaneous structural specifications we ask instead, "what happens if we include or exclude the public opinion measure from the system?"[29]

Two different monthly forecasts for April 2005 to December 2005 were estimated. The forecasts, based on the posterior sample of the coefficients, account for parameter uncertainty and uncertainty about structural shocks. The forecasts assume no structural changes in the model and the election counters and other dummy variables are continuations of those that ended in the estimation sample in March 2005.[30] They are based on the state of the world in March 2005 and do

---

[29]One could do this exercise with the B-SVAR rather than the B-VAR models. The results will be the same, since the SVAR version just decomposes the reduced form residuals. This decomposition is critical for the earlier model selection, log Bayes factor analysis, and interpretation of the impulse responses. Forecast, however, is a reduced form exercise and is invariant to structural specifications.

[30]That is, the election counters and dummies continue forward just as they would if we coded the data through

not include events or information about events as Ariel Sharon's split from Likud and formation of Kadima, his illness and removal from office, the outcomes of the recent Palestinian and Israeli elections, or the recent Fatah-Hamas conflicts. These are what an analyst using the credibility model might have forecast *ex ante* in March 2005. For the seven equation model including the Jewish peace index (JPI) variable in each equation and a separate equation for for JPI as a function of the other variables in the model we computed a set of forecasts. We then estimated a second set of forecasts for a model with no Jewish peace index variable or equation in the model.

We answer the central question in this section with the results in Figure 6 which show the nine month forecasts from April 2005 for the models with and without the JPI variable. The figure shows the median forecasts with 90% posterior error bands that provide an assessment of the risk or uncertainty of the various paths of the variables over time. The KEDS data for April-December 2005 are superimposed on the figure using dashed lines.[31] The first row of forecasts include the JPI equation and the lags of JPI in each equation. The second row shows the forecast results excluding the JPI variable from each equation and omitting the JPI equation from the model.

For the model that includes JPI, the forecasts for Israeli and Palestinian dyadic actions (I2P and P2I) rapidly deteriorate over the course of 24 months. The risk—based on the confidence region—places most of the posterior probability in the region of greatly increased conflict which matches the data over this period. These forecasts are for some of the most severe conflict since the second Intifada. The forecasts for American actions toward the Israelis and Palestinians show less cooperation as well over the 24 months—a conclusion which is borne out by a comparison to the actual data for this period. The changes in Israeli and Palestinian actions towards the Americans change very little over the forecast period and are within the confidence regions of the forecasts.

The forecasts including the Jewish public opinion is for a steady increases over the 24 months. By December 2005, the forecasted median level of JPI has a confidence region over its historical

---

December 2005. The Kadima and Hamas election victories come in January 2006 and thus are outside our forecasting window.

[31] These data were not used in the estimation, so this exercise serves as a cross-validation of our vector autoregression models.

(sample) maximum of 67 (standard deviation of JPI is 5.3). This trend in support of peace by the Israeli public is closely related to the trends in Israeli-Palestinian violence. In contrast the actual level of the peace index declines. While the forecasts support more optimism about peace, the actual data move in the opposite direction, reflecting the actual relations between the two rivals.

[Figure 6 about here.]

For the model omitting the JPI variable the forecast performance is best described as "poor": the 90% confidence regions rarely cover the actual data. Further, even the forecasted trends in the data are hard to discern. The Israeli and Palestinian actual series for the main series of interest— I2P and P2I—are mainly outside the 90% forecast interval. The same is true for the other series in the model. Finally, the forecast errors for the model including the Jewish public opinion data are much smaller. The forecast root mean squared errors for the I2P and P2I series are 11-15 times smaller in the model that includes the Jewish peace index versus the model that excludes this equation and variable. This indicates the strong predictive capacity of the public opinion measure in this case and its general relevance to understanding and interpreting reciprocity, accountability and credibility in the earlier structural models. Table 4 presents the root mean squared forecast error for each equation in each model. There is a large improvement in the forecast error when the JPI series is included in the model. This improvement affects all of the series in the model, in particular, those involving dyadic interactions with the U.S.

[Table 4 about here.]

5 Conclusion

This paper makes two contributions towards increasing our ability to understand and forecast conflict and cooperation between rivals. First, we illustrate the complex interconnections between public opinion and inter-group conflict/cooperation and international mediation. The results support the conclusion that Israeli public opinion is one key component in the system of conflict that

28

links Israeli and Palestinian actors. Our credibility model, which allows for simultaneous responses in and between conflict/cooperation events and Israeli public opinion, provides the best fit to the in-sample data and produces reasonable *ex ante* forecasts for 2005. Additionally, the introduction of Israeli public opinion information into the analysis substantially improves forecasts of conflict and cooperation, as compared to models that ignore public opinion information.

Second, we have applied new Bayesian structural time series tools that allow us to relax the atheoretical, restrictive, recursive assumptions of previous models. For example, to date, event data analyses purporting to test reciprocity hypotheses in hot spots around the globe premised their findings on uni-directional, causality assumptions. Our theoretically-informed models uniquely identify asymmetric reciprocity whereby the more democratic Israel reciprocates while the Palestinian Authority does not reciprocate.

This adds sophisticated empirical support to the credibility model's well-developed micro-foundations (Fearon, 1994). New experimental work on human subjects also supports the idea of audience costs, the key concept on which the credibility model rests (Tomz, 2005). This macro dynamic analysis complements both lines of work since it supplies both a contemporaneous and lagged temporal structure to the credibility processes and it analyzes the behavior of actual belligerents in an important conflict. The impulse response and forecast analyses show how the credibility model can be used to perform substantively and theoretically useful counterfactual analyses as well as to produce early warnings.

The evidence and forecasts this analysis uncovers should be of interest to scholars studying not only localized conflicts around the globe but also general international relations theorists. One of the central debates in IR theory involves the role of domestic politics in the foreign policy process. (see Colaresi, 2007, for a summary). Is the public relevant or does foreign policy operate independently from its' domestic context? Our findings that 1) the public does not purely follow elite cues, and 2) local rivals (and external interveners) likely pay attention to changes in public opinion are consistent with theories that take the dynamic two-way linkages between domestic and world politics seriously and are inconsistent with alternative models of conflict processes that treat

public opinion as irrelevant.

Our study also strongly suggests that additional cases of conflict should be studied in future research. We now have evidence that collecting consistent public opinion information within rivalries over a significant period of time aids in both the analysis and forecasting of conflict and cooperation. This implies that if the same type of information could be collected in other democracies that are involved in rivalries (e.g. India, Greece, Cyprus, Taiwan, and South Korea) our understanding of those specific conflicts as well as conflict and cooperation processes in general would be expanded. These extensions could tell us if credibility is a unique feature of the Levant or, as theory predicts, a common characteristic of democratic rivals.

Bringing the power of Bayesian time series analysis to bear on these questions will produce deeper insights on comparative conflict dynamics. This will answer whether there exists a negative feedback mechanism that prevents both all out war and dramatic peace in other parts of the globe. Further, future research has the possibility to uncover additional evidence as to whether democracies, but not non-democracies, are more likely to reciprocate conflict and cooperation. More precise knowledge of asymmetric reciprocity between democracies and non-democracies, as well as cross-level credibility relationships would enhance our ability to provide *ex ante* forecasts (early warnings) to policy makers as well as *ex post* analysis of foreign policy behavior.

Appendix

Model

The Bayesian structural vector autoregression model employed here is based on a system of equations for each dyadic conflict measure and the Jewish public opinion data. It has one equation for each of these endogenous variables. Each endogenous variable is a function of contemporaneous (time "0") and $p = 2$ past (lagged) values of all of the endogenous variables in the system. The dynamic simultaneous equation model is written in matrix notation as

$$\underset{1\times m}{y_t}\ \underset{m\times m}{A_0}\ +\ \sum_{\ell=1}^{p}\underset{1\times m}{y_{t-\ell}}\ \underset{m\times m}{A_\ell}\ =\ \underset{1\times k}{Z_t}\ \underset{k\times m}{D}\ +\ \underset{1\times m}{\epsilon_t}\ ,\quad t = 1, 2, \ldots, T, \tag{1}$$

with each vector's and matrix's dimensions noted below the matrix. This is an m-dimensional VAR for a sample size of T, with y_t a vector of observations for m variables at time t; A_ℓ the coefficient matrix for lags $\ell = 1, \ldots, p$, $p = 2$ the maximum number of lags (assumed known); Z_t a matrix of exogenous variables for the Israeli election counters, Israeli prime ministerial regimes, and conflict trends for the second Intifada and the post-Battle of Jenin period, and a constant; D is a matrix of coefficients for the exogenous variables; and, ϵ_t a vector of i.i.d. normal *structural shocks*:

$$E[\epsilon_t|y_{t-s}, s > 0] = \underset{1\times m}{0}, \quad \text{and} \quad E[\epsilon_t'\epsilon_t|y_{t-s}, s > 0] = \underset{m\times m}{I}.$$

Two sets of coefficients in it need to be distinguished. The first are the coefficients for the lagged values of each variable, $A_\ell, \ell = 1, \ldots, p$. These coefficients describe how the dynamics of past values are related to the current values of each variable. The second are the coefficients for the *contemporaneous* relationships, (the "structure") among the variables, A_0. The matrix of A_0 coefficients describes how the variables are interrelated to each other in each time period (thus the time "0" impact). The free parameters of the A_0 matrices are defined in the model blocks in Table 2.

The prior for the A_0 and A_+ parameters is specified for (column major) vectorized $a_0 = vec(A_0)$ and $a_+ = vec(A_+)$ where A_+ is a column major stacking of the parameters $A_\ell, \ell =$

31

$1, \ldots, p$:

$$\pi(a) = \pi(a_0)\phi(\tilde{a}_+, \Psi) \tag{2}$$

where the tilde denotes the mean parameters in the prior for a_+, $\phi(\cdot, \cdot)$ is a normal distribution, and Ψ is the prior covariance matrix for $\tilde{a}_+$.

The posterior density for the model parameters is then formed by combining the likelihood for equation (1) and the prior in equation (2):

$$Pr(A_0, A_\ell, \ell = 1, \ldots, p) \propto \phi(a_+ a_0 | Y)\phi(\widetilde{a_+}, \Psi)\pi(a_0) \tag{3}$$

The Bayesian posterior estimates are obtained as detailed in Brandt and Freeman (2007) and Waggoner and Zha (2003). Posterior estimates are found using a Markov Chain Monte Carlo (MCMC) Gibbs sampler algorithm for the equations for the structural model. The estimates reported here are based on a Gibbs sampler with a burn-in of 20000 iterations and 500000 iterations thinned every 5th for the final sample from 200000 draws from two parallel MCMC chains. The posterior estimates pass standard convergence diagnostics such as the Geweke tests and Gelman and Rubin's PSRF.

Impulse responses and forecasts

Details about the impulse response computations are in (Brandt and Freeman, 2006). The responses here are based on one chain (100000 draws) from the posterior sample of the B-SVAR model.

The forecasts are computed by translating the structural model into a reduced form model. The reduced form version of the model,

$$y_t = Z_t C + y_{t-1}B_1 + \cdots + y_{t-p}B_p + u_t, \quad t = 1, 2\ldots, T, \tag{4}$$

is an m-dimensional VAR model for each observation in the sample, with y_t an $1 \times m$ vector of observations at time t, B_ℓ the $m \times m$ coefficient matrix for the ℓth lag, and $p = 2$, the maximum number of lags. In this formulation, all of the contemporaneous effects (which are in the A_0 matrix of the SVAR) are included in the covariance of the reduced form residuals, u_t.

The reduced form in equation (4) is derived from the SVAR model by post-multiplying equation (1) by A_0^{-1}. Thus, the reduced form parameters are transformed from the structural equation parameters via

$$C = DA_0^{-1} \quad B_\ell = -A_\ell A_0^{-1}, \quad \ell = 1, 2, \ldots, p, \quad u_t = \epsilon_t A_0^{-1} \tag{5}$$

where the last term in equation (5) indicates how linear combinations of structural shocks are embedded in the reduced form residuals. Equation (5) shows that restricting elements of A_0 to be zero restricts the linear combinations that describe the reduced form dynamics of the system of equations via the resulting restrictions on B_ℓ and u_t.

The posterior sample of the *ex ante* forecasts is constructed using the following steps:

1. Draw A_0 and A_+ using the Gibbs sampler for the structural model.

2. Compute the reduced form coefficients in equation (5) from the draws of A_0 and A_+.

3. Forecast j periods using equation (4). In these forecasts, the uncertainty of the structural shocks, ϵ_t enters the system by adding a set of reduced form shocks, $u_t \sim N(0, (A_0 A_0)^{-1})$ to the forecasts.

4. Repeat steps 1–3 N times.

The N posterior forecasts are then used to compute the pointwise error bands for the forecasts. The full posterior sample of 200000 draws is used for computing the forecasts.

The exogenous variables (time counters and Israeli prime ministerial regimes) were set based on the values at the end of the sample. That is, trend counters were allowed to continue and no changes in prime ministerial control were made.

References

Aldrich, John H., John L. Sullivan and Eugene Borgida. 1989. "Foreign Affairs and Issue Voting: Do Presidential Candidates "Waltz Before a Blind Audience?"." *American Political Science Review* 83(1):123–142.

Almond, Gabriel A. 1965. *The American People and Foreign Policy*. New York: Praeger.

Auerbach, Y. and C.W. Greenbaum. 2000. "Assessing Leader Credibility during a Peace Process: Rabin's Private Polls." *Journal of Peace Research* 37(1):31–50.

Axelrod, Robert. 1984. *The evolution of cooperation*. New York: Basic Books.

Baumgartner, Frank and Bryan Jones. 2005. *The Politics of Attention*. Chicago, Ill: University of Chicago Press.

Brandt, Patrick T. and John R. Freeman. 2006. "Advances in Bayesian Time Series Modeling and the Study of Politics: Theory Testing, Forecasting and Policy Analysis." *Political Analysis* 14(1):1–36.

Brandt, Patrick T. and John R. Freeman. 2007. "Modeling Macro Political Dynamics." working paper.

Brandt, Patrick T. and John T. Williams. 2007. *Multiple Time Series Models*. Beverly Hills: Sage.

Brown, Nathan J. 2003. *Palestinian Politics: After the Oslo Accords*. Berkeley: University of California Press.

Bueno de Mesquita, Bruce. 1997. "A Decision Making Model: Its Structure and Form." *International Interactions* 23(2):235–266.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph Siverson and James Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.

Chib, Siddartha. 1995. "Marginal Likelihood from the Gibbs Output." *Journal of the American Statistical Association* 90(432):1313–1321.

Colaresi, Michael. 2007. "The Benefit of the Doubt: Testing an Informational Theory of the Rally." *International Organization* 61:99–143.

Eichenberg, Richard and Richard J. Stoll. 2003. "Representing Defense: Democratic Control of the Defense Budget in the United States and Western Europe." *Journal of Conflict Resolution* 47(4):399–422.

Fearon, James D. 1994. "Domestic political audiences and the escalation of international disputes." *American Political Science Review* 88(3):577–593.

Fearon, James D. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52(2):269–306.

Gaubatz, Kurt T. 2001. "Taking the Pulse of Polling: Public Opinion Analysis as a Dependent Variable." Unpublished manuscript. Old Dominion University.

Gill, Jeff. 2002. *Bayesian Methods: A Social and Behavioral Sciences Approach*. New York: Chapman and Hall.

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 Jon C. Pevehouse. 1997. "Reciprocity, Bullying, and International Cooperation: A Time Series Analysis of the Bosnia Conflict." *American Political Science Review* 91(3):515–529.

Goldstein, Joshua S., Jon C. Pevehouse, Deborah J. Gerner and Shibley Telhami. 2001. "Reciprocity, Triangularity and Cooperation in the Middle East, 1979-1997." *Journal of Conflict Resolution* 45(5):594–620.

Guisinger, Alexandra and Alastair Smith. 2002. "Honest Threats: The Interaction of Reputation and Political Institutions in International Crises." *Journal of Conflict Resolution* 46(2):175–200.

Holsti, Ole. 1996. *Public Opinion and American Foreign Policy*. Ann Arbor: University of Michigan Press.

Hudson, Valerie, Phil Schrodt, Ray Whitmer and Adam Shanko. 2006. "A New Kind of Social Science: Analyzing Israeli-Palestinian Event Data Using Reverse Wolfram Models.".

Kass, Robert E. and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90(430):773–795.

Lemke, Douglas. 2007. "The Births and Deaths of Proto States." Unpublished manuscript. Penn State University.

Lippmann, Walter. 1922. *Public Opinion*. New York: Free Press.

Martin, Lisa. 2000. *Democratic Commitments*. Princeton, NJ: Princeton University Press.

McGillivray, Fiona and Alastair Smith. 2000. "Trust and Cooperation Through Agent-Specific Punishments." *International Organization* 54(4):809–824.

McGinnis, Michael D. and John T. Williams. 2001. *Compound Dilemmas: Democracy, Collective Action, and Superpower Rivalry*. Ann Arbor: University of Michigan Press.

Moore, William H. 1998. "Repression and Dissent: Substitution, Context and Timing." *American Journal of Political Science* 42(2):851–873.

Morgenthau, Hans. 1967. *Politics Among Nations*. Forth ed. New York: Knopf.

Organski, A.F.K. and E. Lust-Okar. 1997. "The tug of war over the status of Jerusalem: Leaders, strategies, and outcomes." *International Interactions* 23(4):333–350.

Page, Benjamin and Robert Shapiro. 1995. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press.

Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 42(3):427–60.

Rosato, Sebastian. 2003. "The Flawed Logic of the Democratic Peace." *American Political Science Reivew* 97(4):585–602.

Sargent, Thomas, Noah Williams and Tao Zha. 2006. "Shocks and Government Beliefs: The Rise and Fall of American Inflation." *American Economic Review* 96(4):1193–1224.

Schrodt, Philip A. 2000. "Forecasting Conflict in the Balkans Using Hidden Markov Models." Paper presented at the annual meeting of the American Political Science Association, Washington, D.C.

Schrodt, Philip A. and Deborah J. Gerner. 2000. "Cluster Based Early Warning Indicators for Political Change in the Contemporary Levant." *American Political Science Review* 94(4):803–817.

Schrodt, Philip A., Deborah J. Gerner, Rajaa Abu-Jabr, Oemeur Yilmaz and Erin M. Simpson. 2001. "Analyzing the Dynamics of International Mediation Processes In the Middle East and Balkans." Paper presented at the Annual Meeting of the American Political Science Association, San Francisco.

Shellman, Steven M., Andrew Reeves and Brandon Stewart. 2006. "Fair & Balanced or Fit to Print? The Effects of Media Sources on Statistical Inferences.".

Shlaim, Avi. 2001. *The Iron Wall: Israel and the Arab World*. New York: Norton and Company.

Sims, Christopher A. and Tao A. Zha. 1998. "Bayesian Methods for Dynamic Multivariate Models." *International Economic Review* 39(4):949–968.

Sims, Christopher A. and Tao A. Zha. 2006. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* 96(1):54–81.

Sirriyeh, Hussein. 1995. "Is a Palestinian State Politically Possible?" *Israel Affairs* 2(1):46–58.

Sylvan, Donald A., Jonathan W. Keller and Yoram Z. Haftel. 2004. "Forecasting Israeli-Palestinian Relations." *Journal of Peace Research* 41(4):445–463.

Tomz, Michael. 2005. "Domestic Audience Costs in International Relations: An Experimental Approach." Unpublished, Stanford University.

Waggoner, Daniel F. and Tao A. Zha. 2003. "A Gibbs sampler for structural vector autoregressions." *Journal of Economic Dynamics & Control* 28:349–366.

Ward, Michael D. 1982. "Cooperation and Conflict in Foreign Policy Behavior: Reaction and Memory." *International Studies Quarterly* 26(1):87–126.

Wlezien, Christopher. 1996. "Dynamics of Representation: The Case of U.S. Spending on Defense." *British Journal of Political Science* 26.
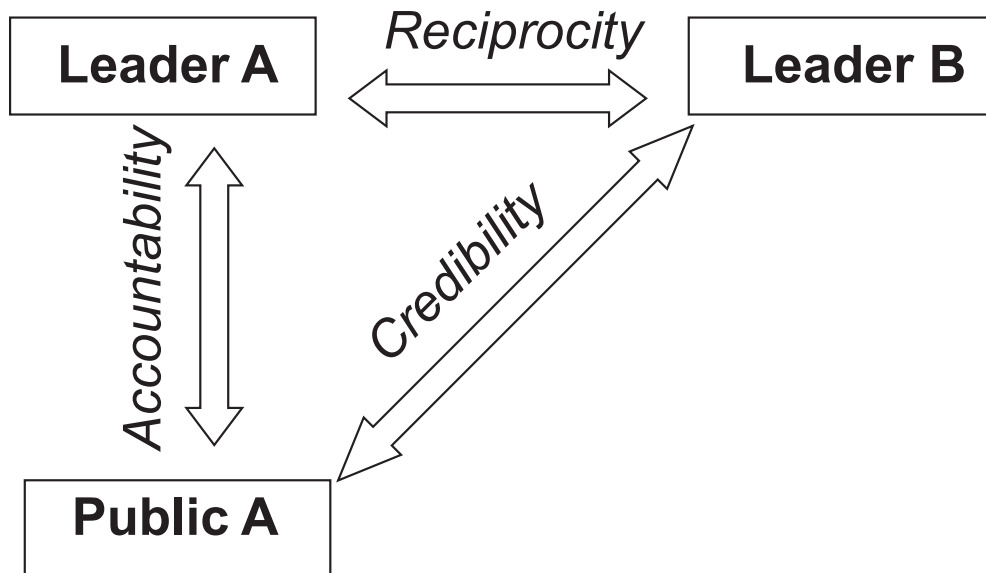
Figure 1: A Road Map of Two-way Streets: Reciprocity, Accountability, and Credibility
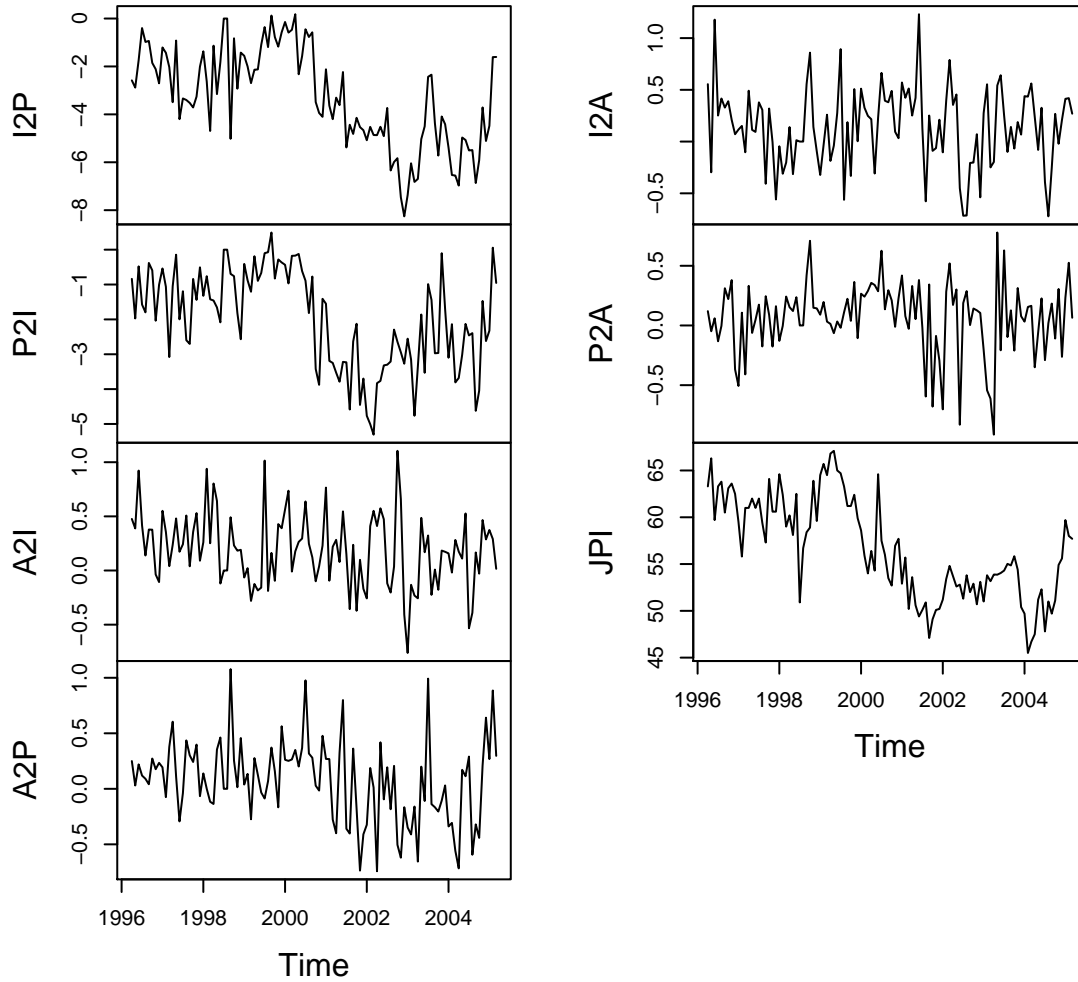
Figure 2: Average Monthly Relations Between Israel, Palestine, and U.S. and Jewish Public Opinion Data, April 1996-March 2005
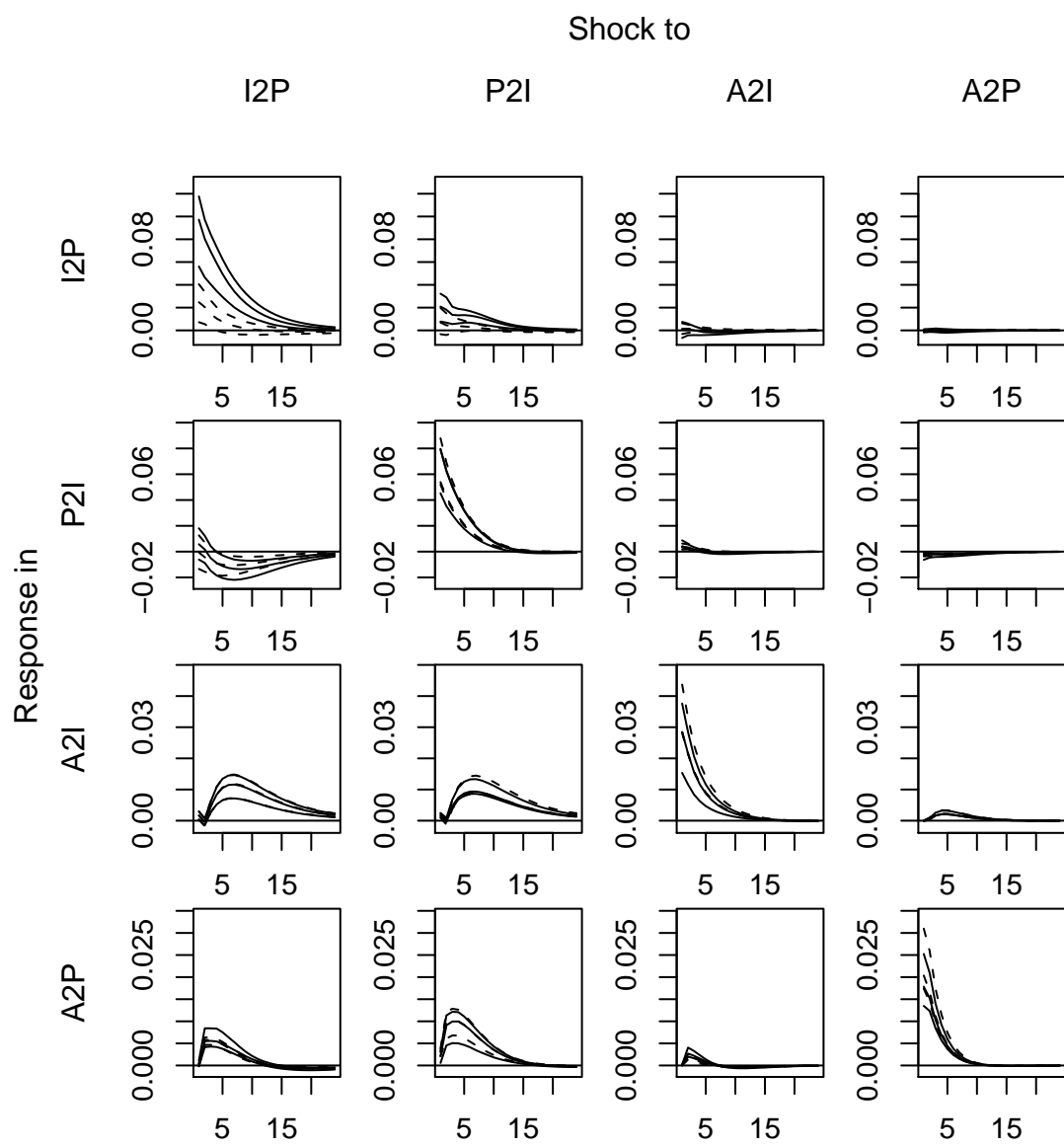
Figure 3: Reciprocity responses for the Bystander and Credibility models. Bystander model results are shown with solid lines and credibility model results are shown with dashed lines. Responses are median estimates of positive one standard deviation shocks to each equation with 90% error bands computed by eigenvector decomposition method over 24 months.

Shock to

I2P   P2I

Response in
JPI

0.00

−0.05

−0.10

−0.15

5 10   20   5 10   20
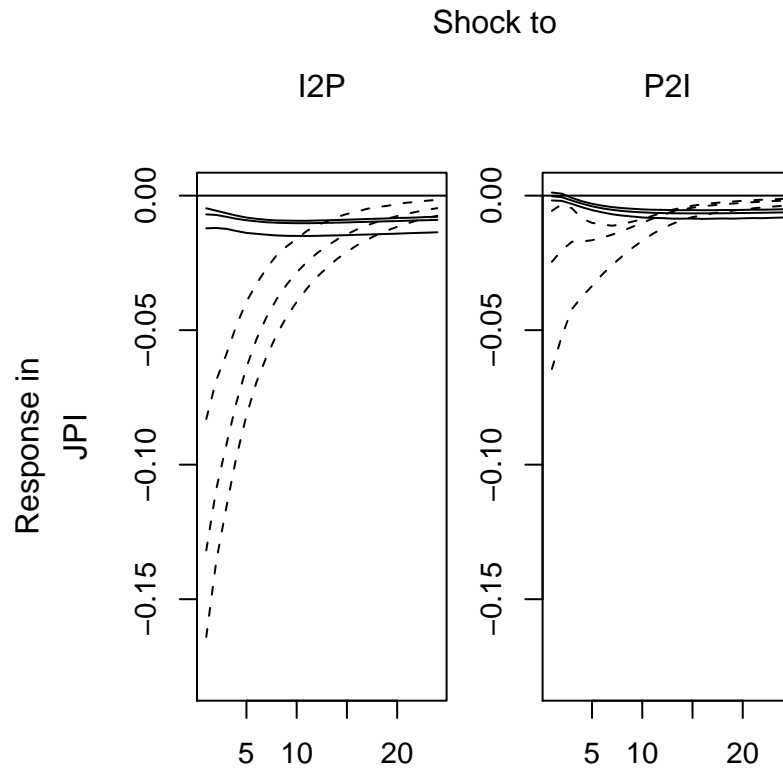
Figure 4: Responses of the Jewish Peace Index to innovations in Israeli actions towards the Palestinians. Follower model results are shown with solid lines and credibility model results are shown with dashed lines. Responses are median estimates of positive one standard deviation shocks to the column variable with 90% error bands computed by eigenvector decomposition method over 24 months.

40
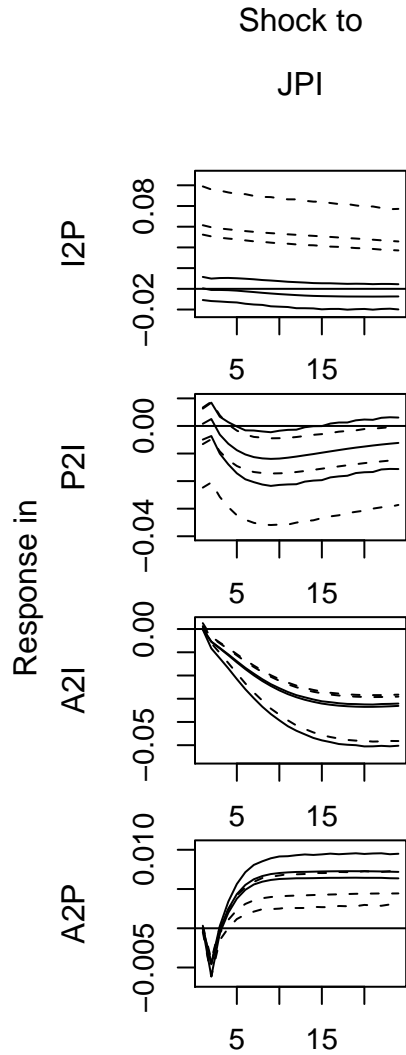
Figure 5: Israeli and Palestinian responses to innovations in the Jewish Peace Index. Accountability model responses are shown with solid lines and Credibility model results are shown with dashed lines. Responses are median estimates with 90% error bands computed by eigenvector decomposition method over 24 months. JPI shocks are negative one standard deviation changes for all equations in both models.
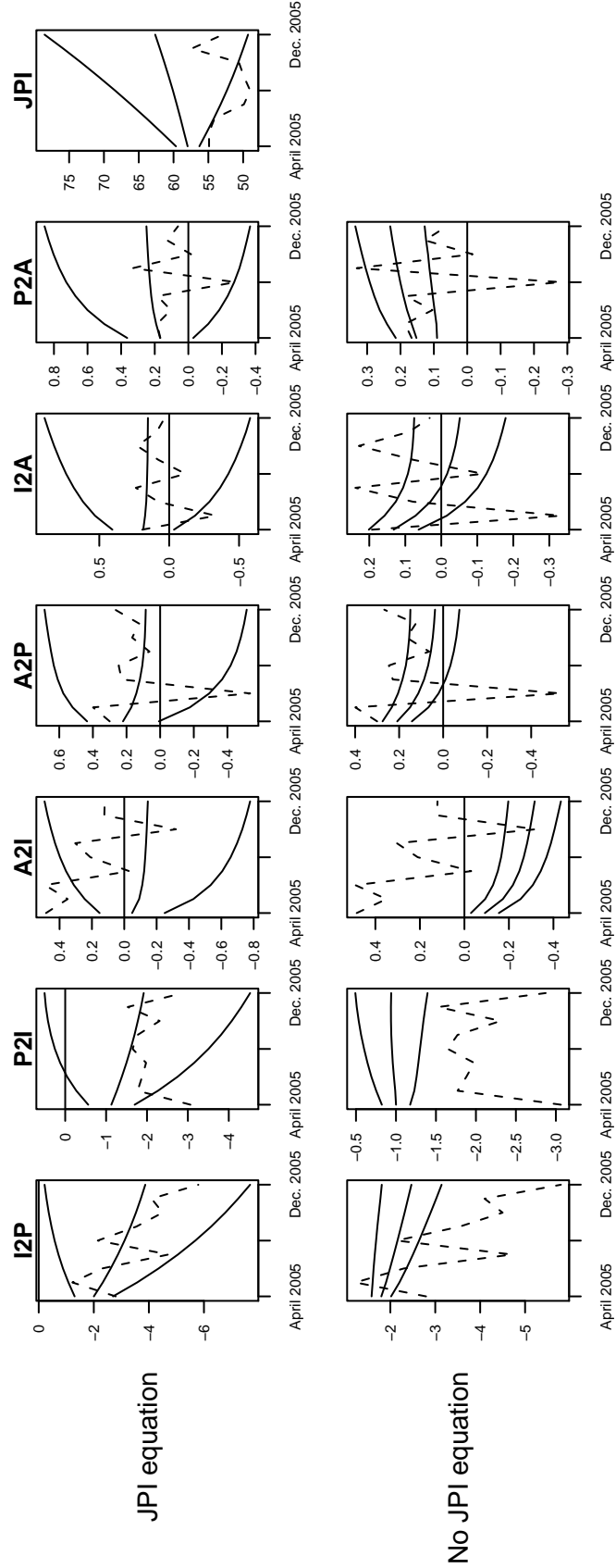
Figure 6: Bayesian VAR forecasts with and without the Jewish Peace Index (JPI) equation, April 2005–December 2005. Solid lines are Bayesian forecast median estimates with 90% pointwise posterior confidence regions. Dashed lines are actual series over this period. Top row are forecasts for a model that includes the JPI equation; bottom row excludes the JPI equation.

| Model	Response in	Shock in			
		$Policy_{1\to2}$	$Policy_{2\to1}$	$Policy_{Z\to1}$	$Public_1$
Bystander	$Policy_{1\to2}$	-	Yes	Yes	No
	$Policy_{2\to1}$	Yes	-	Yes	No
	$Policy_{Z\to1}$	Yes	No	-	No
	$Public_1$	No	No	No	-
Follower	$Policy_{1\to2}$	-	Yes	Yes	No
	$Policy_{2\to1}$	Yes	-	Yes	No
	$Policy_{Z\to1}$	Yes	No	-	No
	$Public_1$	Yes	No	No	-
Accountability	$Policy_{1\to2}$	-	Yes	Yes	Yes
	$Policy_{2\to1}$	Yes	-	Yes	No
	$Policy_{Z\to1}$	Yes	No	-	No
	$Public_1$	Yes	Yes	No	-
Credibility	$Policy_{1\to2}$	-	Yes	Yes	Yes
	$Policy_{2\to1}$	Yes	-	Yes	Yes
	$Policy_{Z\to1}$	Yes	No	-	Yes
	$Public_1$	Yes	Yes	No	-

Table 1: Four Models of Foreign Policy Behavior: Reaction in Row to Column

Model block	Variable	$I2P_t$	$P2I_t$	$A2I_t$	$A2P_t$	$I2A_t$	$P2A_t$	JPI_t
Recursive	$I2P_t$	X						
	$P2I_t$	X	X					
	$A2I_t$	X	X	X				
	$A2P_t$	X	X	X	X			
	$I2A_t$	X	X	X	X	X		
	$P2A_t$	X	X	X	X	X	X	
	JPI_t	X	X	X	X	X	X	X
Bystander	$I2P_t$	X	X	X				
	$P2I_t$	X	X		X			
	$A2I_t$	X		X		X		
	$A2P_t$		X		X		X	
	$I2A_t$			X		X		
	$P2A_t$				X		X	
	JPI_t							X
Follower	$I2P_t$	X	X	X				
	$P2I_t$	X	X		X			
	$A2I_t$	X		X		X		
	$A2P_t$		X		X		X	
	$I2A_t$			X		X		
	$P2A_t$				X		X	
	JPI_t	F						X
Accountability	$I2P_t$	X	X	X				A
	$P2I_t$	X	X		X			
	$A2I_t$	X		X		X		
	$A2P_t$		X		X		X	
	$I2A_t$			X		X		
	$P2A_t$				X		X	
	JPI_t	A	A					X
Credibility	$I2P_t$	X	X	X				C
	$P2I_t$	X	X		X			C
	$A2I_t$	X		X		X		C
	$A2P_t$		X		X		X	C
	$I2A_t$			X		X		
	$P2A_t$				X		X	
	JPI_t	C	C					X

Table 2: Contemporaneous Relationships for recursive, bystander, follower, accountability and credibility models Each model block specifies the contemporaneous relationships and restrictions for the seven equations in the associated B-SVAR model (A_0 matrix). Rows correspond to equations and the columns to variables or changes in the variable in a contemporaneous equation. The X's in the cells represent the "free" parameters or those estimated while the empty cells are zero restrictions. A zero restriction indicates no contemporaneous relationship is hypothesized between a column variable and a given row equation.

| Model | Log MDD | Log Bayes Factor versus | | | |
		Recursive Model	Bystander Model	Follower Model	Accountability Model
Recursive	411.66				
Bystander	414.69	3.03			
Follower	417.96	6.30	2.73		
Accountability	422.02	10.36	7.06	4.06	
Credibility	428.74	17.08	13.78	10.78	6.72

Table 3: Posterior model summaries. Log Bayes factors are found by the difference of the row model from the related column model.

	With JPI	Without JPI	RMSE ratio
I2P	1.93	20.82	10.79
P2I	1.38	20.33	14.73
A2I	0.49	20.07	40.96
A2P	0.40	19.95	49.88
I2A	0.40	19.98	49.95
P2A	0.36	19.92	55.33
JPI	10.06		

Table 4: Root mean square error of the forecasts versus the actual data for models with and without the JPI equation. Estimates based on an average over 200,000 Bayesian posterior forecasts from April-December 2005. Final column shows the ratio of the RMSEs for each equation for the model without the JPI equation to that with the JPI equation.