

Reciprocity, Accountability and Credibility In International Relations*

Patrick T. Brandt
School of Social 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

May 31, 2006

Abstract

Do public opinion dynamics play an important role in understanding conflict dynamics between democracies and their international rivals? These opinion dynamics and government behavior are interpreted as particular causal links in models of reciprocity, accountability and credibility. Theoretical expectations about the character of these linkages are translated into four distinct Bayesian structural time series models. The models are fit to events data from the Israeli-Palestinian conflict with provisions for U.S. intervention. We find that a credibility model 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 in which low level conflict is maintained and momentum toward either all out war or dramatic peace is slowed. Finally, we use the credibility model to forecast *ex ante* 24 months ahead from March 2005. These forecasts show rapid deterioration of Israeli-Palestinian relations (even if recent election outcomes in the Levant had not occurred). The forecasts illustrate that the level of violence and support for peace move in opposite directions.

*Freeman and Brandt's research is supported by the National Science Foundation under Grants Nos. SES-0351179, SES-0351205, SES-0540816. We thank Will Moore, Jon Pevehouse and Phil Schrodt for their advice on this project and also the Tami Steinmetz Center for Peace Research for supplying us with public opinion data. An earlier version of this project was presented at the Annual Meeting of the International Studies Association, San Diego, March 22-25, 2006. The authors are solely responsible for the paper's contents.

1 Introduction

Much progress has been made recently in the analysis of international conflict. 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, can promote cooperation and peace. What has not been studied is how and when public opinion in belligerent countries affects conflicts. Does public opinion forbid or encourage local leaders to reciprocate other state'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?

Answering these questions is difficult for several reasons. First, domestic politics is embedded in the conflict system. This is a system of multiple relationships (behavioral equations) of 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. A model that imposes strong exogeneity restrictions therefore is liable to produce biased results. Finally, any link between conflict and domestic political dynamics likely exists 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. Electoral calendars increase the public interest in foreign policies and heighten official's interest in public evaluations of these policies. Analyzing conflicts in terms of quarterly or yearly patterns of behavior will miss these dynamics.

We assess the impact of domestic politics on international conflict in a way that solves these three problems. Reciprocity, accountability, and credibility are interpreted as particular causal links between government behavior 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 and competing beliefs about con-

flict dynamics. The competing specifications are key elements of multi-equation, Bayesian time series models with complex, endogenous relationships between 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 events 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. For the first time in a major conflict, the impact of public opinion on reciprocity, accountability and credibility can be assessed.¹

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 which are estimated 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 pacific Israeli opinion leads to more immediate conflict by the Palestinians. This supports the credibility model since Palestinians rationally 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. The credibility model is used to forecast 24 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 elections in 2005-6. The risk of a significant increase in conflict—based on the confidence regions—is high. The public opinion forecast shows that the level of violence and support for peace move in opposite directions.

¹The 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 (over 100 months). While public opinion databases for other countries tend to share one or two of these traits, no other source of information that we are aware of satisfies all three criteria.

2 Reciprocity, Accountability and Credibility

2.1 Reciprocity

A branch of international relations research investigates the role of reciprocity in foreign policy. Building on Axelrod (1984) and repeated non-zero sum games (the 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 state 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), 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, 2006a; 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 two remaining empirical puzzles. First, what explains deviation 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 international events. If this unexplained variance follows a predictable pattern, what theories can help us improve our predictions? Second, reciprocal behaviors may be asymmetric

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

2.2 Domestic Politics and International Conflict

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 international relations as a two-level game where elites sit at the interstices of international and domestic politics. Leaders must worry about whether their constituents will ratify agreements they negotiate with other states 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 public disapproval of a leader. 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 communicate resolve in international conflict situations (Fearon, 1994; Martin, 2000) 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 that citizens' preferences constrain leaders' tit-for-tat interactions. In their "rivalry-as-prison" or Goldilocks hypothesis, the public prefers international 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 the public disapproval of a leader. Under these constraints, leaders spend their time reciprocating each other's limited cooperative or bellicose policies. Guisinger and Smith (2002) and McGilliray and Smith (2000) hypothesize that in democracies the public

prefers to pocket the gains that accrue from reciprocated international cooperation rather than to defect for a one-time payoff that makes long term cooperation impossible. The public then punishes a leader deviating from reciprocity. In these explanations, the public constrains a leader's behavior and supports reciprocal international 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), and Morgenthau (1967) 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, citizens are hypothesized to be manipulated by leaders. 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 forecasting 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 countries. Colaresi (Forthcoming) 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 model 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: Cross-level Credibility

Rival states 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 a leader in State B will examine the public preferences constraining State 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 State B's ideal outcome), the greater the credibility of an uncompromising offer from the 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 cooperate.² 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 own public's negative reactions, *as well as the Israeli public's negative reactions*. Conversely,

²This only holds if public preferences are negatively or unrelated to a rival's actions (see below).

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.³ 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.⁴ 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.

[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 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

³Note that if accountability does not hold—the public are followers/bystanders, then credibility is irrelevant.

⁴Putnam (1988) described this dynamic as international-to-domestic reverberation. Rivals that cooperate (attack) may signal to the public that cooperation is safe (unsafe) and (not) profitable.

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 rival state 1 (2) toward state 2 (1) is denoted by Policy_{1→2} (Policy_{2→1}), and public opinion in state 1 as Public₁. 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 international politics. Here the public pays little attention to international events. In this view, previous scholarship on reciprocity and triangular relationships is correct in ignoring public opinion in its empirical specifications. High and low politics essentially are unrelated. This thesis predicts that 1) a surprise change in conflict and cooperation either sent or received by one state 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 internationally and the public evaluates 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 the public changes support of conflict or cooperation toward a rival, its government’s policies reflects

this. This third perspective contends that the public is paying attention not only to what their government does (state 1), but also what the rival (state 2) may be doing. Thus there should be a bi-directional relationship between conflict/cooperation and public opinion.

The final model incorporates the belief that if accountability relationships exist within one country, its rival is likely to use that information to formulate its own foreign policy toward that country. In this *credibility model* public preferences in state 1 affect policy in state 2 toward state 1. If the public has relevant ratification power, as suggested by the accountability model, this information is unlikely to be ignored by the rival.

[Table 1 about here.]

Another relevant process is external third-party intervention (state Z in Table 1). Following previous work on triangularity, localized conflicts are likely to react to great power interventions that encourage or discourage 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.

3 Research Design

3.1 The Case

The Israeli-Palestinian 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. Schrodtt 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.⁵

These studies do not consider the possibilities of accountability or credibility. This is despite 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. Local elections and general elections were held recently in Palestine.

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).⁶ The events are scaled with Goldstein scores and aggregated into monthly averages.⁷ 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

⁵Sylvan, 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.

⁶In the future, we would like to collect and analyze data collected from alternative sources. On the importance of employing multiple sources see Shellman, Reeves and Stewart (2006).

⁷Monthly 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 that demonstrates similar results about reciprocity and triangularity see Brandt and Freeman (2006a).

directed by the U.S. toward the Palestinians at time t . Our sample is monthly from April 1996 to March 2005. Thus, our analysis stops before the recent Israeli and Palestinian elections, the election victory of Kadima, and the election of Hamas.

The public opinion measure is from polls conducted by the Tami Steinmetz Center for Peace Research (TSC). The TSC is a multidisciplinary academic enterprise of faculty 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 that Rabin commissioned. But it provides a continuous sounding of the Israeli public's evaluation of their government's effort to create peace and of the prospects for peace.⁸ Finally, a few peace index observations are missing—out of the 126 months 8 have missing values in 2003. The missing values were inputted via an ARIMA model.⁹ The Jewish peace index, denoted as JPI_t , and the six monthly average event data series are depicted 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

⁸The 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 are: 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.

⁹The imputation is via 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.

(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.¹⁰ Finally, two dummy variables for changes in the stochastic trends in the mean level of conflict are included. 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).¹¹

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*a,b*). The B-SVAR model has seven equations—six for the dyadic interactions for the Israelis, Palestinians, and Americans, and one for the Jewish opinion about the peace process. Test statistics support using two lagged values for each of the seven endogenous variables.¹² Each equation in the B-SVAR includes the nine exogenous variables described earlier.

¹⁰The 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.

¹¹The 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.

¹²F-statistics support using 2 lags while AIC, BIC and Hannan-Quinn criteria support using 1 lag. We err on the side of inefficiency by including 2 lags in the model, since using too few lags produces biased estimates. We also looked

The next sections describe elements of the model that allow us to capture the essence of the four 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 Palestinians, 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 (zero) 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. All four imply a number of zero restrictions that make the matrix of contemporaneous relationships non-recursive and over-identified.¹³

Consider first the bystander model. Like the others, this model allows for contemporaneous reciprocal and triangular relationships in directed behaviors. The first seven rows of Table 2 represents the contemporaneous relationships among the variables (the A_0 matrix discussed in the Appendix). The rows in the bystander block of Table 2 correspond to the equations for each directed-dyad or Jewish public opinion equation. The columns are the contemporaneous shocks that enter each 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 shock and a given row equation. The bystander

at the posterior distribution summaries for models with more lags and found that they fit worse than 2 lag models.

¹³In this seven equation (variable) model given the formulation of the B-SVAR, there are at most 28 free parameters in this matrix of contemporaneous 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.

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 in public opinion. The absence of these contemporaneous relationships are represented as blanks and are the zero order restrictions imposed by this first theory.¹⁴

[Table 2 about here.]

The other theories allow for additional contemporaneous relationships in the structural time series 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 the Israeli, Palestinian, and American governments monitor and react immediately to shifts in Jewish opinion and 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.

¹⁴Of 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.

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 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.¹⁵

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 produce useful forecasts.¹⁶ 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 non-zero coefficients of the lagged effects, especially those at the most distant (largest) lags. 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

¹⁵If 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 $B = A_+ A_0^{-1}$.

¹⁶The 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 than 5000 coefficients. Because of such scale, meaningful causal inferences about the dynamic responses of their system will be difficult to make. For this reason investigators usually don't even try to analyze dynamics, conduct innovation accounting or produce forecasts. There is good reason to believe that most of the coefficients on the distant lags 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.

degrees of persistence in behavior.¹⁷ Details are in Sims and Zha (1998) and Brandt and Freeman (2006a,b).¹⁸

3.3.3 Model evaluation

Model scale also creates challenges for model evaluation. Complex models are bound to overfit the data making conventional (frequentist) fit statistics less useful. The current time series literature and Bayesian statistics employ two measures to summarize posterior model fit (Sargent, Williams and Zha, forthcoming; Sims and Zha, forthcoming; Brandt and Freeman, 2006b). The first is the Bayesian information criteria (BIC). The BIC is a measure of model fit based on the log posterior density adjusted for the degrees of freedom and penalized by the number of parameters in the model.¹⁹ The second is the log marginal data density (log MDD) which 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. Differences in log MDDs across models are 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).²⁰

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

¹⁷This 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.

¹⁸The 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 (see Brandt and Freeman, 2006a). The hyperparameter 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.

¹⁹The log posterior density (LPD) is the log likelihood plus the log prior. It is proportionate to the posterior probability for the model.

²⁰It is important when using Bayes factors to assess the sensitivity of the results to the prior specification for each model. We address this issue in the Appendix.

bands for these responses. For example, if the bystander model receives strong support, its BIC value should be smaller than the BIC values for the other three models. And if 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 for the responses of I2P and P2I should include zero.

For the best fitting model, we produce a set of forecasts for April 2005–March 2007, and assess their plausibility. 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.²¹

4 Results

4.1 Model fit and selection

Table 3 presents the posterior fit measures for the four different structural models—bystander, follower, accountability, and credibility. Moving down the rows of this table, each model contains more contemporaneous parameters (in the its A_0 matrices). The BICs are listed in the first column of Table 3. The BIC measures show that best fitting model is the credibility model with a BIC of 508 which is lower than the other models.²²

[Table 3 about here.]

The best fitting model, with the largest log MDD, is the credibility model. The final three columns of Table 3 are the Bayes factors. There is little difference between the bystander, follower,

²¹Measure 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).

²²The BICs values presented here are computed using the LPDs. Typically the BIC is computed using the log likelihood. But, as argued by Sims and Zha (forthcoming), the LPD provides a better measure of the uncertainty of the model fit. Using the log likelihood to compute the BICs produces a qualitatively similar result.

and accountability models (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 three models, with values between 10.98 and 12.74. This is very strong evidence for the credibility model over the other contemporaneous specifications.²³ 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.

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 in the estimated structural models. The full set of impulse responses for the systems estimated here include 196 responses (49 responses per model \times 4 models). Rather than present all these responses, we focus on those of greatest substantive interest: the responses of I2P and P2I to Jewish public opinion shocks and the responses of Jewish public opinion to conflict shocks.

Given the high volatility and correlated uncertainty about the dynamics of these models, we compute error bands using an eigenvector decomposition method (Brandt and Freeman, 2006a).²⁴ The reported error bands are 90% (posterior) confidence regions around the median estimates. They provide a summary of the general direction and shape of the responses. The impulse re-

²³We 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.

²⁴The 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 for the correlation in the responses. The variation explained by the first eigenvector of each response is between 45% to 100% of the variation in the impulse responses. (for details see Brandt and Freeman, 2006a). Error bands computed from the empirical percentiles of the posterior responses generate similar inferences.

sponses presented here are interpreted differently from those typically seen in the literature. In standard reduced form VAR models with a recursive identification of the contemporaneous error covariance, one analyzes the responses of *positive shocks* to each equation in the system. Such a normalization of the shocks is not possible in non-recursive B-SVAR models like ours (Waggoner and Zha, 2003b). Since the contemporaneous shocks to a set of equations may be (negatively) correlated with each other in the non-recursive system, shocks or innovations of different signs may enter each equation. Thus positive shocks in one equation may imply negative shocks in other equations. This is most important for interpreting the responses for reciprocity and triangularity in the model. This sign normalization difference is particularly acute in the models presented here because there is no consistent pattern of signed shocks entering the equations for the models.

Figure 3 shows the responses of the I2P and P2I equations to shocks in the I2P and P2I variables. The rows in this figure are the responses for these equations to shocks in the column variables.²⁵ The bystander (credibility) model responses and error bands are shown with solid (dashed) lines. In the first row, the bystander model I2P responses are for positive shocks in the column variables, while those in the credibility model are for negative shocks. The same “flipped” sign pattern is seen in the responses of P2I to shocks in I2P and P2I where the shocks enter the equations with different signs across the two models.²⁶ 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 negative. Here the response in P2I is briefly positive but the error bands soon span zero. This indicates, implausibly, that the Palestinian response to Israeli conflict innovations is very short-lived to nonexistent.

[Figure 3 about here.]

²⁵These are one standard deviation shocks of the residuals from the respective equation in the B-SVAR model.

²⁶One could flip the signs of the shocks / responses in each equation and generate the same pattern of responses for the I2P equation. The consequence is that the response signs for the P2I equation would not be invariant to the sign flip, since they are endogenously related to the I2P responses.

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 negative and the response again is reciprocal or negative in I2P. But row two of Figure 3 shows that for this model a positive innovation in I2P produces a clear, sustained inverse response in P2I. This finding that the only the more democratic member of the conflict dyad reciprocates cooperation (and conflict) is consistent with the work of Guisinger and Smith (2002) and McGilliray and Smith (2000) who suggest that democratic institutions create incentives for leaders to use reciprocal strategies. Conversely, their models imply 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. 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.²⁷

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

²⁷The contemporaneous correlation of shocks in the credibility model generate a plausible rivalry scenario. Below we map the signs of the diagonal of A_0 for the credibility model into a vector and interpret the response pattern to a contemporaneous surprise surge in conflict across all of the equations:

$$\begin{pmatrix} - \\ + \\ - \\ - \\ - \\ + \\ - \end{pmatrix} = \begin{pmatrix} \text{Increased Israeli conflict toward the Palestinians in response} \\ \text{Up-tick Palestinian's verbal communication/cooperation} \\ \text{American threats to Israel to avoid conflict spiral} \\ \text{American threats to Palestinian leadership to avoid conflict spiral} \\ \text{Israeli warnings to the U.S. that it will retaliate} \\ \text{Palestinian verbal communication with the U.S.} \\ \text{Decrease in the Jewish public's support for peace} \end{pmatrix}$$

(JPI) to unexpected surges in Israeli conflict toward the Palestinians. The solid (dashed) lines are the responses from the follower (credibility) model. Both these responses are for negative one standard deviation changes in I2P. For the follower model, there is a weak positive response—indicating an inverse public reaction to increased conflict. In the credibility model response in Figure 4, the median response in JPI to a negative 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 endogeneity. In the credibility model which allows endogeneity between the Jewish support for the peace process and conflict, the presence of accountability and credibility generates a strong public opinion response to changes in the level of conflict. The direction of 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 and P2I equations for innovations in Jewish opinion. The main difference between of the credibility model and the others are the impacts of 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 column is change in JPI while the rows are the responses of the I2P and P2I equations. JPI shocks are positive one standard deviation innovations for both equations in the accountability model. JPI are negative (positive) one standard deviation innovations in the I2P (P2I) equation for the credibility model. 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 always include zero for the two models.²⁸ But for the accountability model, increases in Jewish support for the peace process produce a sustained, pacific Palestinian response toward the Israelis. In contrast, for the credibility model, an innovation in P2I initially produces a negative response quickly falls within error bands that span zero. 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 overestimated. On the other hand, in the credibility model—where the instantaneous feedback between opinion and events is explicitly modeled—a cross-level (domestic to international) pattern of asymmetric interaction is evident. The Palestinians immediately react to public support for peace with conflict and opposition to peace with cooperation. This response, like the asymmetric reciprocity finding discussed above (Figure 3), is that Palestinian actions counteract rather than accelerate changes in public opinion.

4.3 Forecasts

For our best fitting credibility model, a set of monthly forecasts for April 2005 to March 2007 were estimated. The forecasts, based on the posterior sample of the coefficients, account for 1) parameter uncertainty and 2) uncertainty about structural shocks. The forecasts assume no structural changes in the model. They are based on the state of the world in March 2005 and do not include events or information about events like Ariel Sharon's split from Likud and the formation of Kadima,

²⁸The evidence that Israeli does not respond to innovations in opinion does not necessarily invalidate hypotheses about accountability. Leaders could anticipate changes in their public's opinion. It is important that Israeli policy responds negatively (positively) to a negative (positive) shock in opinion, as expected from an accountability perspective. The upper bound of the confidence region for this credibility model response is very close to zero.

his illness and removal from office, or the outcomes of the recent Palestinian and Israeli elections. These are what an analyst using the credibility model might have forecast *ex ante* in March 2005.

Figure 6 show the 24 month forecasts from April 2005. These are median forecasts with 68% error bands (approximately ± 1 standard deviation) that provide an assessment of the risk or uncertainty of the various paths of the variables over time.²⁹ 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. 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. The changes in Israeli and Palestinian actions towards the Americans change very little over the forecast period.

[Figure 6 about here.]

Finally, the forecast for Jewish public opinion is for a steady increase over the 24 months. By March 2007, the forecasted median level of JPI is 83 which is well above its historical (sample) maximum of 67 and well above its sample mean of 57 (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. It is not that the violence escalates as support for peace drops. As the forecasts show, the level of violence and support for peace move in opposite directions—consistent with the idea that the public constrains rather than follows international policy.

5 Conclusion

The theoretical value of this case study should not be underestimated. The Tami Steinmetz Center polls provide one of the only opportunities to test the dynamic implications of four influential

²⁹The smaller 68% posterior regions are used for the forecasts to summarize the central tendency or likelihood and location of the forecasts, not the posterior inference of the model (see Sims and Zha, 1999).

models of two-level interactions. To promote peace and understand how cooperation is built, it is of paramount importance to understand the role domestic publics play in the Levant.

The results support the credibility model which produces significant, reasonable *ex ante* forecasts for 2005-2007. 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. These 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 produces 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 international relations and the foreign policy process. (see Colaresi, Forthcoming, 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 high and low politics seriously and are inconsistent with alternative models of the foreign policy process that treat public opinion as irrelevant to international relations.

Additional cases of conflict will be studied in future research. These extensions will include analyses of Balkan and India-Pakistani conflicts. They will tell us if credibility is a unique feature of the Levant or, as theory predicts, a common characteristic of international conflict. Bringing the power of Bayesian time series analysis to bear on these questions will produce deeper insights into 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. This knowledge enhances our ability to provide *ex ante* forecasts (early warnings) to policy makers.

$1, \dots, p$:

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

where $\pi(a_0)$ is the prior on the A_0 elements (Waggoner and Zha, 2003a, see), 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):

$$Pr(A_0, A_\ell, \ell = 1, \dots, p) \propto \phi(a_+ a_0 | Y) \phi(\tilde{a}_+, \Psi) \pi(a_0) \quad (3)$$

The Bayesian posterior estimates are obtained as detailed in Brandt and Freeman (2006b) and Waggoner and Zha (2003a). 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 5000 iterations and 60000 iterations for the final sample 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, 2006a). The responses here are based on 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,

$$y_t = Z_t C + y_{t-1} B_1 + \dots + y_{t-p} B_p + u_t, \quad t = 1, 2, \dots, T, \quad (4)$$

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 ℓ^{th} 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

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

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 $N = 60000$ posterior forecasts are then used to compute the pointwise error bands for 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. 2006a. "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. 2006b. "Modeling Macro Political Dynamics." working paper.
- 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, Siddhartha. 1995. "Marginal Likelihood from the Gibbs Output." *Journal of the American Statistical Association* 90(432):1313–1321.
- Colaresi, Michael. Forthcoming. "The Benefit of the Doubt: Testing an Informational Theory of the Rally." *International Organization* .
- 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. 1998. "Bargaining, Enforcement, and International Cooperation." *International Organization* 52(2):269–306.
- Fearon, James D. 1994. "Domestic political audiences and the escalation of international disputes." *American Political Science Review* 88(3):577–593.
- 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 Schrodtt, Ray Whitmer and Adam Shanko. 2006. *A New Kind of Social Science: Analyzing Israeli-Palestinian Event Data Using Reverse Wolfram Models*. International Studies Association San Diego, CA: .
- Kass, Robert E. and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90(430):773–795.
- Lippmann, Walter. 1922. *Public Opinion*. New York: Free Press.
- Martin, Lisa. 2000. *Democratic Commitments*. Princeton, NJ: Princeton University Press.
- McGilliray, Fiona and Alastair Smith. 2000. "Turst and Cooperation Through Agent-Specific Punishments." *International Organization* (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.

- Sargent, Thomas, Noah Williams and Tao Zha. forthcoming. "Shocks and Government Beliefs: The Rise and Fall of American Inflation." *American Economic Review* .
- 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, Oemur 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. International Studies Association San Diego, CA: .
- 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. 1999. "Error Bands for Impulse Responses." *Econometrica* 67(5):1113–1156.
- Sims, Christopher A. and Tao A. Zha. forthcoming. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* .
- 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. 2003a. "A Gibbs sampler for structural vector autoregressions." *Journal of Economic Dynamics & Control* 28:349–366.
- Waggoner, Daniel F. and Tao A. Zha. 2003b. "Likelihood Preserving Normalization in Multiple Equation Models." *Journal of Econometrics* 114:329–347.
- 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.

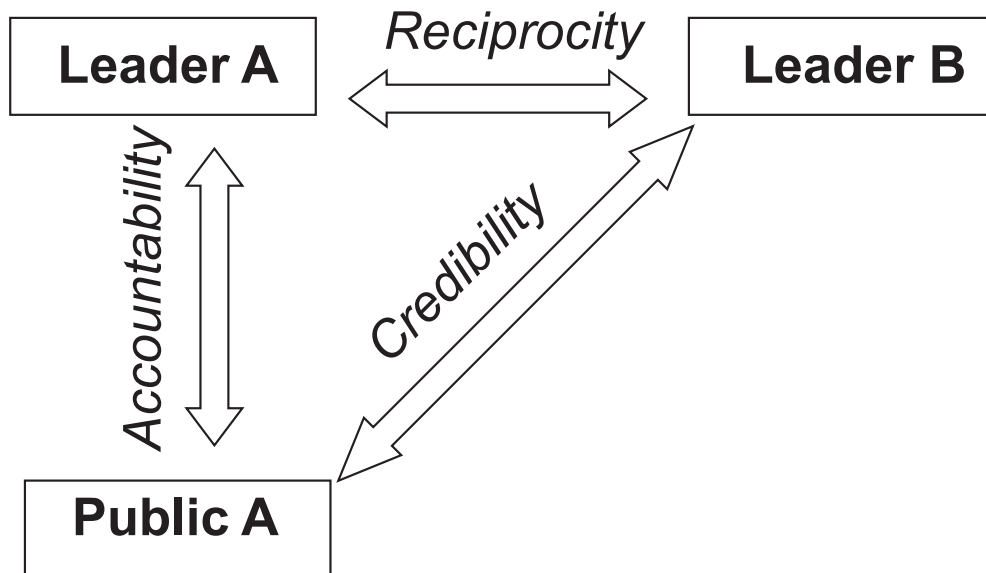


Figure 1: A Road Map of Two-way Streets: Reciprocity, Accountability, and Credibility

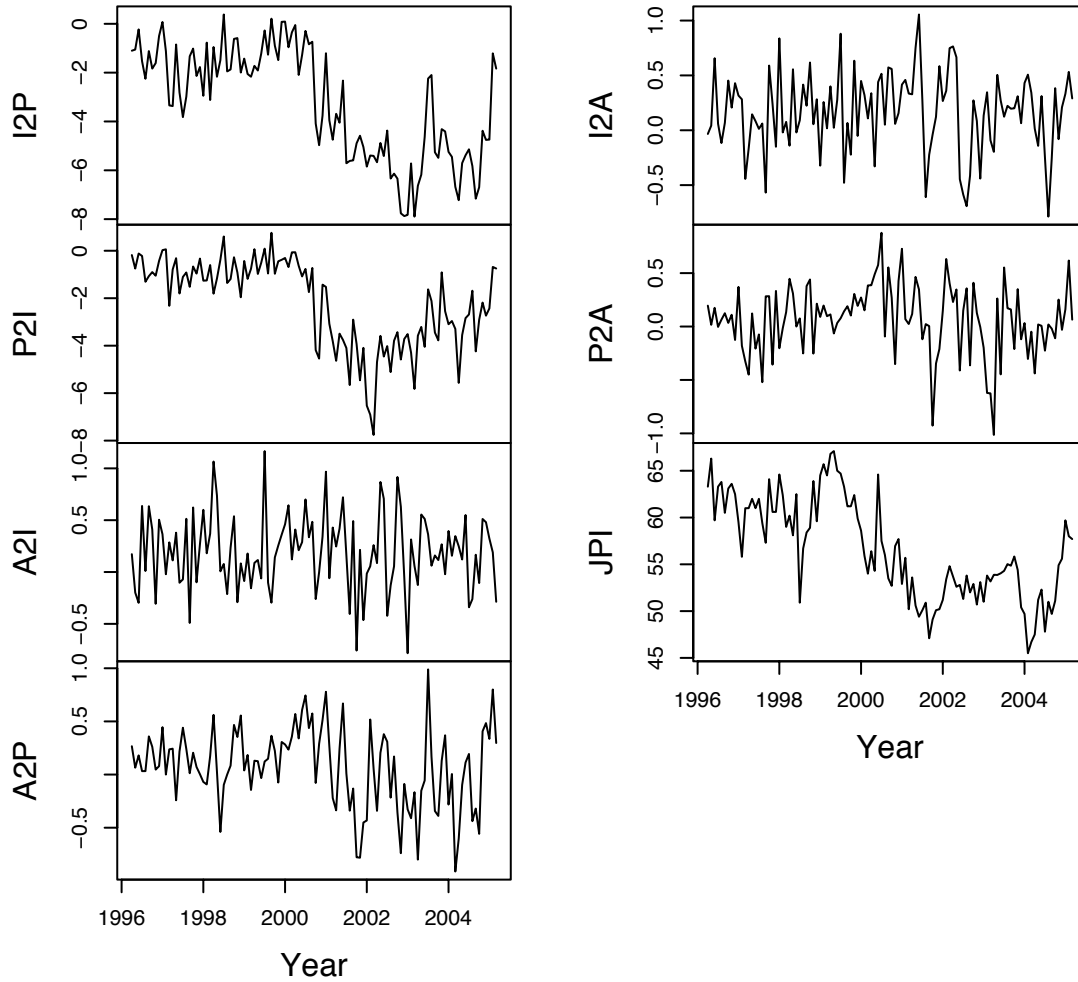


Figure 2: Average Monthly Relations Between Israel, Palestine, and U.S. and Jewish Public Opinion Data, April 1996-March 2005

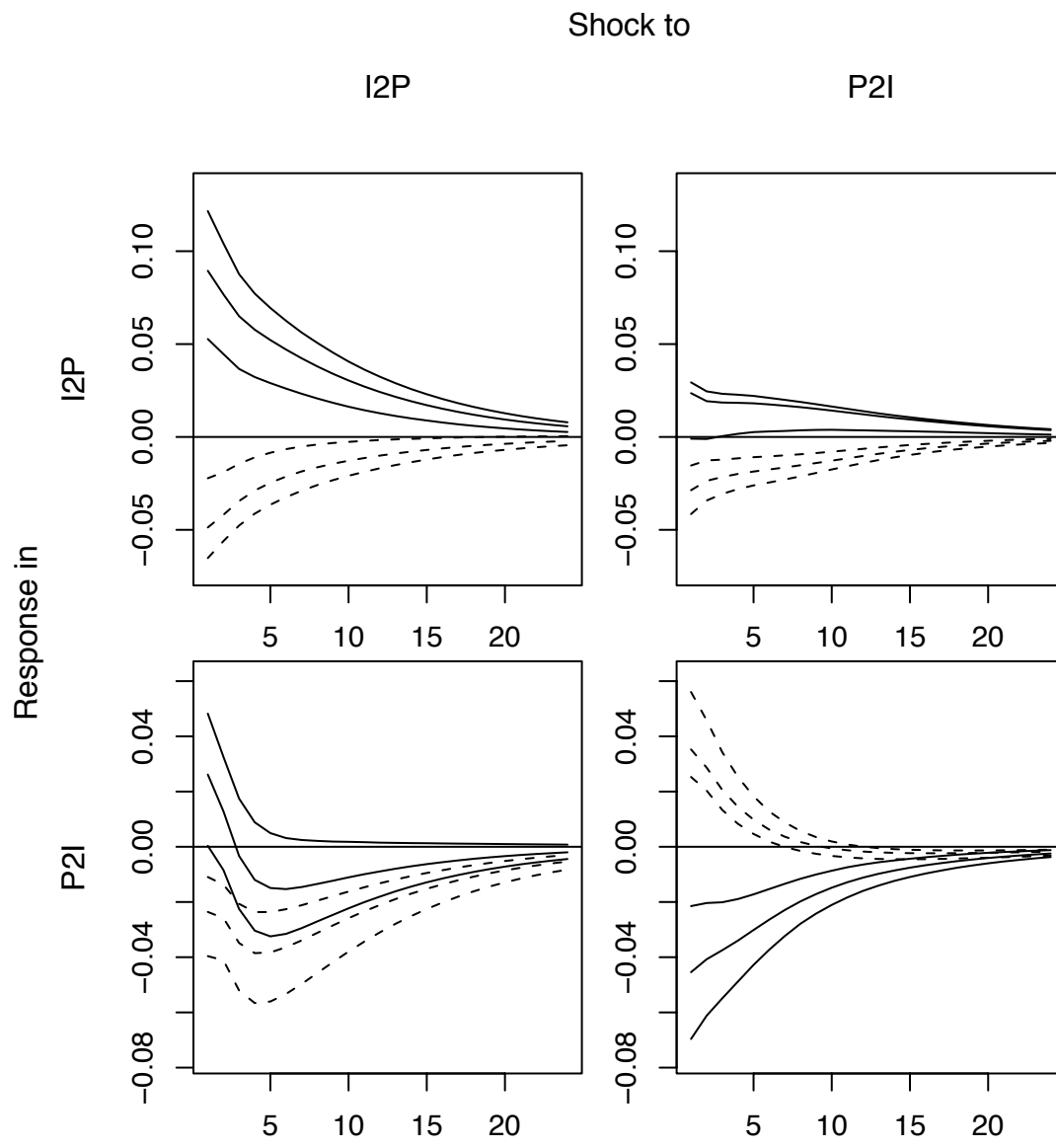


Figure 3: Reciprocity responses for the Bystander and Credibility models. Bystander model results are shown with solid lines and credibility model results are shown with dashed lines. Responses are median estimates with 90% error bands computed by eigenvector decomposition method over 24 months. Shocks to the I2P equation (row) are positive (negative) one standard deviation innovations in the bystander (credibility) model. Shocks to the P2I equation (row) are negative (positive) one standard deviation shocks in the bystander (credibility) model. Own shocks across the two equations are negatively related in the bystander and credibility models.

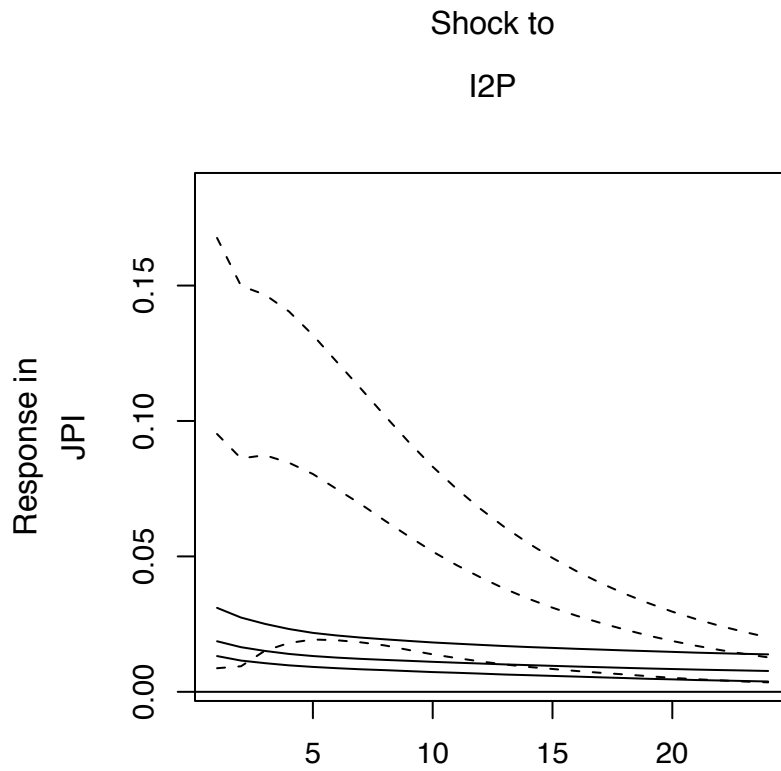


Figure 4: Responses of the Jewish Peace Index to innovations in Israeli actions towards the Palestinians. Follower model results are shown with solid lines and Credibility model results are shown with dashed lines. Responses are median estimates with 90% error bands computed by eigenvector decomposition method over 24 months. Shocks to the JPI equation are negative one standard deviation innovations in I2P for both models.

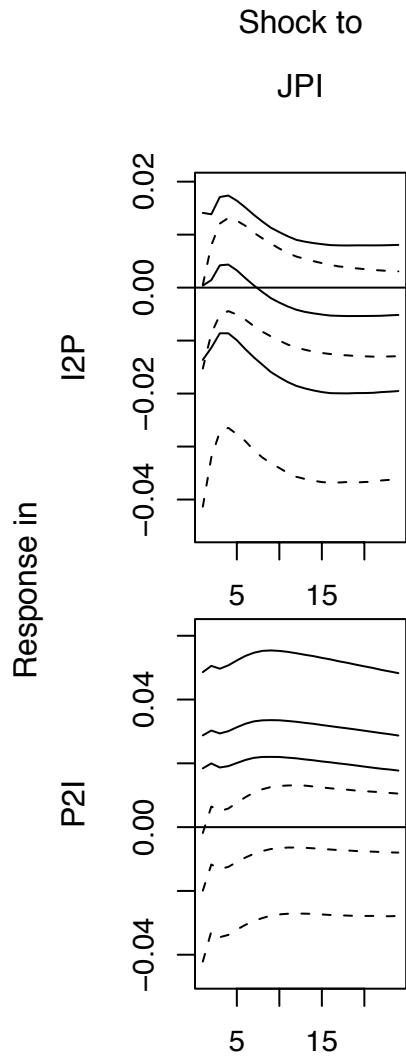


Figure 5: Israeli and Palestinian responses to innovations in the Jewish Peace Index. Accountability model responses are shown with solid lines and Credibility model results are shown with dashed lines. Responses are median estimates with 90% error bands computed by eigenvector decomposition method over 24 months. JPI shocks are positive one standard deviation changes for both equations in the accountability model. JPI are negative (positive) one standard deviation changes in the I2P (P2I) equation for the credibility model.

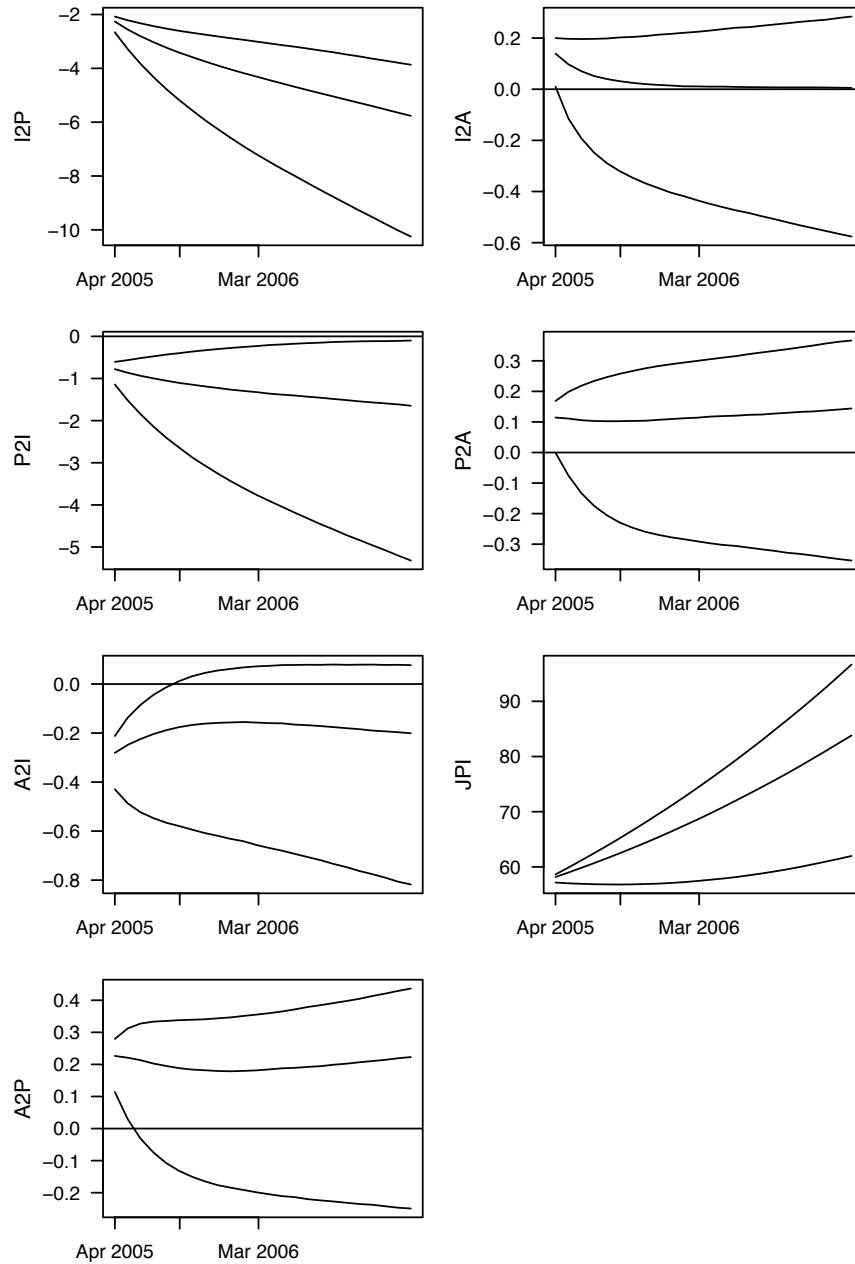


Figure 6: Credibility model forecasts, April 2005–March 2007. Forecast are Bayesian median estimates with 68% pointwise posterior confidence regions (approximately ± 1 standard deviation).

<i>Model</i>	<i>Response in</i>	<i>Shock in</i>			
		<i>Policy_{1→2}</i>	<i>Policy_{2→1}</i>	<i>Policy_{Z→i}</i>	<i>Public₁</i>
Bystander	Policy _{1→2}	-	Yes	Yes	No
	Policy _{2→1}	Yes	-	Yes	No
	Public ₁	No	No	No	-
Follower	Policy _{1→2}	-	Yes	Yes	No
	Policy _{2→1}	Yes	-	Yes	No
	Public ₁	Yes	No	No	-
Accountability	Policy _{1→2}	-	Yes	Yes	Yes
	Policy _{2→1}	Yes	-	Yes	No
	Public ₁	Yes	Yes	No	-
Credibility	Policy _{1→2}	-	Yes	Yes	Yes
	Policy _{2→1}	Yes	-	Yes	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
Bystander	$I2P_t$	X	X	X				
	$P2I_t$	X	X		X			
	$A2I_t$			X		X		
	$A2P_t$				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		
	$A2P_t$				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		
	$A2P_t$				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		C
	$A2P_t$				X		X	C
	$I2A_t$			X		X		
	$P2A_t$				X		X	
	JPI_t	C	C					X

Table 2: Contemporaneous Relationships for 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	BIC	Log MDD	Bayes Factor versus		
			Bystander Model	Follower Model	Accountability Model
Bystander	513.41	409.87			
Follower	514.40	411.63	1.76		
Accountability	513.89	411.48	1.61	-0.15	
Credibility	508.25	422.61	12.74	10.98	11.13

Table 3: Posterior model summaries