Moving Mountains: Bayesian Forecasting As Policy Evaluation*

Patrick T. Brandt
Department of Political Science
University of North Texas
E-mail: brandt@unt.edu

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

April 22, 2002

Abstract

Many policy analysts fail to appreciate the dynamic, complex causal nature of political processes. We advocate a vector autoregression (VAR) based approach to policy analysis that accounts for various multivariate and dynamic elements in policy formulation and for both dynamic and specification uncertainty of parameters. The model we present is based on recent developments in Bayesian VAR modeling and forecasting. We present an example based on work in Goldstein et al. (2001) that illustrates how a full accounting of the dynamics and uncertainty in multivariate data can lead to more precise and instructive results about international mediation in Middle Eastern conflict.

*Prepared for the 2002 Meeting of the Midwest Political Science Association, April 25-28, 2002, Chicago, Illinois. An earlier version of this paper appeared as a research memo for a conference at the Center for Basic Research in the Social Sciences (http://www.cbrss.harvard.edu/programs/humansecurity/papers/mediation.pdf). We thank the participants in the methods workshops at the Universities of Pittsburgh and Texas for comments and criticisms. The authors would like to thank Tim Hellwig for research assistance, Jon Pevehouse and Phil Schrodt for data and guidance, and Mel Hinich and John Williams for comments. Brandt's work has been supported by a University of North Texas Faculty Research Initiation Grant.

Introduction

Political methodology has made tremendous advances in recent years. We now know how to generate and analyze many different types of political data. We have a much better understanding of how to sample from relevant populations, pool panel and time series data, and estimate complex models of political processes. It is not surprising then that, like our counterparts in economics, we increasingly are asked to forecast important events like state failure and war and to evaluate international mediation, the fiscal and monetary policy effects of elections, and other public policies.

However, in comparison to economists and other social scientists, our forecasting and policy evaluation skills are relatively underdeveloped. For example, most of our leading texts do not include any explicit treatment of forecasting and (or) policy evaluation. For many problems, our forecasting and policy evaluation tools are virtually nonexistent. If we are to advise policy makers, we must develop these tools.

This is the purpose of our paper. Drawing on new work in econometrics and building on the work in several branches of political methodology, we present a Bayesian approach to policy evaluation. In particular, we show how policy contingent forecasts can be generated from multivariate Bayesian time series models of political processes. Our approach employs new developments in Bayesian vector autoregression such as the use of reference priors and of conditional forecasting with Gibbs sampling. (Leeper, Sims and Zha 1996, Robertson and Tallman 1998, Roberston and Tallman 1999, Sims and Zha 1998, Sims and Zha 1999, Waggoner and Zha 1999, Zha 1998)

Our discussion is divided into three parts. Part one briefly reviews policy evaluation methods. It argues that these methods suffer from five problems: an inadequate conception and treatment of dynamics; the reliance on rarefied, inaccurate conceptions of causality — a refusal to deal with what are reciprocal (endogenous), multivariate causal relationships; the use of highly cross-sectionally and temporally aggregated measure of political variables; the failure to gauge model uncertainty; and, the failure to incorporate systematically, qualitative-heuristic knowledge — "the wisdom" literature — in the model building process (Schrodt and Gerner 2000, Howell 2001).

We then propose, in part two, a five step method that meets all of the implied desiderata. It entails construction of Bayesian VAR models and forecasts and sampling from the posterior distributions thereof, subject to conditions that represent hard (or soft) policy choices (counterfactuals). At the heart of our approach is the idea of determining whether the policy counterfactuals shift unconditional, multivariate forecast densities for selected policy outcomes to more desirable, policy contingent (conditional) forecast densities or, euphemistically, whether policy counterfactuals are likely to "move mountains."

Next we illustrate our approach with an example. We show how, at the end of 1988, a counterfactual pattern of U.S. mediation might have altered the history of the Israeli-Palestinian conflict. Using daily KEDS data on conflict in the Levant (cf., Goldstein, Pevehouse, Gerner and Telhami (2001), Schrodt, Gerner, Abu-Jabr, Yilmaz and Simpson (2001), Gerner et al. 2002), we demonstrate the potential "real time" value of our approach for U.S. peacemaking.

The conclusion lays out a research agenda for further work on our method and its implications for studying the Israeli-Palestinian and related international conflicts. The main contributions of our paper are demonstrating the feasibility of drawing inferences with a Bayesian model about policy counterfactuals and showing how we can present policy contingent forecasts that account for both model and parameter uncertainty.

Policy Evaluation in Political Science

Political scientists employ both qualitative-heuristic and quantitative-statistical approaches to policy evaluation.[1] The former appears in journals like *Policy Studies, Policy Studies Journal*, and in a plethora of specialized publications such as *The Journal of Health Policy* and *Politics and Law*. Qualitative-heuristic approaches often are thick descriptions of the origin and histories of policies. In many instances they are case studies. But their temporal frame also can be short-term in nature. Evaluation amounts to listing the benefits and costs of policies. Counterfactuals simply are descriptions of what might have occurred historically had policies not been adopted. Out of this approach supposedly emerges "policy wisdom" (Schrodt, 2001) — a set of heuristics or lessons about how policy ought to be formulated and applied.

Quantitative-statistical approaches to policy evaluation appear in the journals mentioned above but also in publications like the *American Political Science Review, American Journal of Political Science, The Journal of Politics, Journal of Conflict Resolution, Public Administration Review*, and *International Studies Quarterly*. These studies can be divided into structural and nonstructural types.[2] Like the Cowles tradition in economics, the structural approach treats theory-as-equations. It posits equations—many times a single equation—as representations of the structure of the respective policy process (Prescott 1991). Policy is

[1] To save space, we will not discuss at length mathematical approaches to policy evaluation. But see fn. 2 below. Interestingly, there is a movement toward more use of quantitative-statistical approaches to policy evaluation in social science generally, the so-called "evidenced based policy movement." This movement places a high value on randomization. See *The Economist*, March 2, 2002: 73-74.

[2] The structural/nonstructural distinction often is used in the literature on economic forecasting; cf., Diebold (1998). The former approach often is associated with the work of the Cowles commission in the 1950s and 1960s. The latter is associated with developments surrounding the rational expectations revolution of the 1970s and 1980s. VAR models sometimes are considered structural both in terms of the contemporaneous relationships that exist between (shocks in) variables and specification of policy rules. See, for example, Stock and Watson (2001). Finally, economists consider dynamic, stochastic, general equilibrium (DSGE) models structural tools for forecasting and policy evaluation. According to Stock and Watson (2001, 113) DSGE models do not perform as well as other models in this regard. Diebold (1998, 185-6) explores efforts econometricians have made to improve the performance of DSGE models. For illustrations of these models in political science-including the analysis of counterfactuals — see Freeman and Houser (1998) and Houser and Freeman (2001).

conceived as an independent, right-hand-side variable in an equation that explains a policy outcome. Cases usually are pooled in an attempt to learn the average effects of the policy. Analysts then ask how a hypothetical (counterfactual) change in the policy variable affects on average, the dependent (policy outcome) variable — that is, if the other independent variables are held constant. The time frames for the structural analyses tend to be more medium and long-term in nature. As a general rule, these investigations do not have the real time value of many qualitative-heuristic studies. Knowledge is the accumulation of robust results about the average effects of policies across time and space.

The nonstructural approach induces the functional form of policy relationships from data. These are often case studies. Consider time series analyses. For example, intervention analysis is used to assess the impact of policies net of what are presumably autonomous sources of change in policy variables. Illustrative are assessments of the impact of clean air policies on what otherwise are seasonal fluctuations in pollution in selected urban areas (Box and Tiao 1975). Policy counterfactuals are studied by manipulating the magnitude and duration of such interventions. When vector autogressions (VARs) are used, policy analysis amounts to simulating the orthogonalized moving average response of the reduced form model to a one standard deviation shock in a policy variable or to a (counterfactual) stylized sequence of such shocks. Monetary and other economic policy counterfactuals are evaluated in this way. This nonstructural genre also includes neural network models of policy. The covariates for such models are induced from data, data that are usually pooled. Policy analysis again amounts to making hypothetical changes in what are presumed to be independent, right-hand-side independent variables and then assessing their impact on the probabilities of realizing certain policy outcomes.

These different approaches to policy analysis can be found in international relations. There is a large body of work that uses qualitative-heuristic approach to learn why international conflicts occur and, more important, how best to mediate and resolve them (Crocker, Hampson and Aall 1999, Greenberg 2000, Hampson 1996, Lederach 1997, Zartman and Rasmussen 1997, Kleiboer 1998, Downs and Steadman 2002). Such studies can be found in major field journals like *Foreign Policy, Foreign Affairs, International Security,* and *Orbis.*

The quantitative-statistical approach to policy evaluation is well established in international relations as well. Structural analyses of the impact of mediation on conflicts like those in southern Europe and the Middle East frequently appear in journals like *International Studies Quarterly* and *The Journal of Conflict Resolution* (for an example and review see Schrodt, 2001). Recently, there has been a burst of work in the nonstructural tradition, work assessing the impact of mediation efforts of Western countries in these two parts of the world. This works stresses the need for triangularity—evidence that mediation by a third party has a causal impact on the relations between pairs of belligerents. It promises real-time evaluation of mediation policies as well as of mediation counterfactuals (Goldstein and Pevehouse 1997, Pevehouse 1999, Goldstein et al. 2001). There also has been a major effort to use logit models to forecast state failure and, in a recent critique and extension of that work, to apply neural network models for the same purpose (King and Zeng 2001*b*). The policy implication of this research is that intervention designed to reduce infant mortality and perhaps to promote democracy will reduce the probability of state failure two years ahead.[3]

[3] Admittedly, much of the new work in the nonstructural quantitative-statistical genre is intended for forecasting, especially, providing early warnings of war and of state failure (Schrodt and Gerner (2000), for instance). The original aim of the state failure task force was of this nature. However, the task force's work clearly has policy implications: by intervening to reduce infant mortality, promote democracy, etc., policy makers presumably can reduce the expected probability of state failure in countries (cf. King and Zeng (2001*b*, 652-653)).

Critique

Policy evaluation methods in political science suffer from at least five major problems.[4] The first is an inadequate treatment of dynamics. Too often these studies are purely cross-sectional analyses that ignore process change (Brunner and Liepelt 1972) or, they wrongly treat units of analysis like state (country)-years as independent events. But the world does not start over at each time unit; macro variables within states and countries are not independent of their previous values. When dynamics are addressed it usually is in the form of a single lag of the dependent variable on the right hand side of equations. For various reasons, this functional form is unlikely to capture the nature of social and political processes. It also creates serious statistical problems, particularly in analyses of count data. (Brandt, Williams, Fordham and Pollins 2000, Brandt and Williams 2001, Meier and Gill 2001)

A failure to come to grips with what are systems of (endogenous) relationships is a second failing of the policy evaluation literature. To begin with, the social systems that policy makers seek to manage are highly interdependent. If qualitative-heuristic studies teach us anything it is that social reality is not characterized by simple one-way, unilinear causal relationships. It follows that policies are not exogenous to outcomes; policy outcomes are sure to affect policy choices if for no other reason than the infallibility of policy makers—policy makers are rarely if, ever able to achieve their goals on the first try. The deeper problem, statistically, is that it is not possible to manipulate a policy variable counterfactually without also affecting the behavior of other variables in a social system. To assume otherwise is, in effect, to claim that the policy variable is strictly exogenous to the other variables in the system or, that the disturbance in the policy equation is uncorrelated with the disturbances in the equations for all the other variables in the system. This usually is an unreasonable

[4]Because it is the genre in which we work, the focus of our critique is the quantitative-statistical literature. But our fifth point is that we should incorporate more systematically, the insights of the qualitative-heuristic literature.

assumption.[5]

Excessive aggregation is a third problem. The penchant for pooling cases reduces the practical value of many policy evaluations. The knowledge it yields amounts to "average effects" (Meier and Gill, 2000). Policy makers guided by such studies know that, if they apply the results for average effects, they run the risk of committing an ecological fallacy. The highly temporally aggregated nature of many studies also diminishes their practical value. Unless one argues that the natural unit of policy making is years, it is difficult to find much "real time" value in many policy evaluations. Temporal aggregation, of course, also is a source of statistical problems (Freeman 1990).[6]

Fourth, neither structural nor nonstructural statistical approaches to policy evaluation adequately address model uncertainty. Policy evaluations ignore a) sampling variability in outcome and other variables and also b) parameter variability. If policy counterfactuals are extrapolations rather than interpolations of existing data, our results can be prone to biases of various kinds (King and Zeng 2001*a*). For example, orthogonalized moving average responses from VARs need standard error bands—measures of how certain we are of that these responses indeed connote meaningful changes in policy outcomes in one direction or

[5] Leeper, Sims, and Zha (1996) write:

> In practice, the usual answer is that simulations of the effects of paths of policy variables or of hypothetical policy rules are conducted under the assumption that such policy changes can be made without producing any change in disturbance terms in other equations, even if the estimated covariance matrix of disturbances shows strong correlations. This is not logically inconsistent, but it amounts to the claim that the true policy disturbance is that part of the reaction function residual that is not correlated with other disturbances in the system. This, in turn, is equivalent to claiming that the true reaction function is a linear combination of what the model labels the reaction function and the other equations in the system whose disturbances are correlated with it. Our view is that if one is going to do this in the end, the assumptions on the model that justify doing so should be explicit from the beginning. (Leeper, Sims and Zha 1996, 9)

See also Freeman, Williams and Lin (1989, 858-859).

[6] In recognition of the problems policy analysts encounter in working with "average effects," Meier and Gill (2000) develop methods (SWAT, GSRLS) for performance isolation and recommendation. They extend their approach to pooled cross-sectional, time series designs. However, they focus more on "wide shallow pools" ($N > T$) than on "deep narrow pools" ($T > N$); they leave time series methods for a future work; cf. Ibid., Chapter 7, fns. 2,4.

the other. The existing methods for creating such bands—Monte Carlo integration—force us to make some strong, often unreasonable assumptions. More important, if they are included in policy analyses, these error bands often show that policy counterfactuals have no impact on outcomes.[7]

Finally, the potential for gains from trade between the qualitative-heuristic and quantitative-statistical approaches has not been realized. The new Bayesianism in political methodology points the way to applying policy "wisdom" in quantitative work. But, despite early attempts to build bridges between the two genres (Jackman and Western, 1994), political methodologists these days are prone to use uninformative priors rather than informed priors.[8]

Recent policy work in international relations avoids some but not all of these problems. The work on state failure is laudable for the way it addresses the fifth issue. From the outset, this body of research, has made a conscious effort to draw on the work of scholars who employ qualitative-heuristic methods. Recent extensions like King and Zeng (2001*b*, 642) draw on qualitative-heuristic work to construct "neural network priors." King and Zeng also make a serious effort to address various sources of model uncertainty. Unfortunately, the state failure research suffers from the first, second, and third problems. It posits a two lag, unilinear (one-way causal) structure for state failure. And it employs a a highly temporally aggregated, pooled set of data. The results it gives us are "average effects." The practical value of these results therefore is a matter of dispute.

The strengths and weaknesses of recent work on international conflict and mediation are the mirror images of those the of the state failure research. The international conflict and

[7]The most common procedure for constructing error bounds is the RATS Monte Carlo integration procedure. As Sims and Zha (1998, 950) point out this imposes the same prior on all the equations in the model; and, this is an uninformed, flat prior. On the use of naive bootstrap and related approaches to constructing error bands for the impulse responses of VAR models see Sims and Zha (1999, 1124-1127). These error bounds will be the benchmark against which we compare the virtue of a reference prior for multivariate, dynamic time series models of conflict in section three of this paper.

[8]To be fair to Jackman, in some of his recent work he does explore the usefulness of informed priors. Cf., Jackman (2000, 17).

mediation literature is attentive to dynamics and causal complexity; indeed the notion of "triangularity" is a based on the idea of a complex causal nexus of relationships between the behaviors of belligerents and mediators. Work on international conflict and mediation is of greater practical use to policy makers than the state failure research. This is because it focuses on particular cases of conflict and it employs temporally disaggregated measures of conflict (*viz.*, events measures often at daily and weekly levels of temporal aggregation). The works' claims of "real time" policy relevance (Pevehouse and Goldstein, 1999) accordingly are quite reasonable.[9]

Unfortunately, both quantitative-statistical research on international conflict and mediation suffers from the fourth and fifth problems. Students of international conflict and mediation make no effort to gauge the degree to which their results stand up in the face of model uncertainty. To our knowledge, no one working in this area has provided any kind of error bands for their policy counterfactuals. Nor do these scholars incorporate systematically in their models the work of foreign policy analysts and historians. In effect, they ignore the new Bayesianism in political methodology.[10]

---

[9] In fact, in comparison to macroeconomics, events data allows us to avoid "vintage" problems—the delay in the publication of certain key economic variables and also, in comparison to financial economics, to work with continuous series with no interruptions for weekends, holidays, etc. For example, Stock and Watson (2001: 112) note that the assumption of the stickiness of output and inflation to monetary shocks might be reasonable for daily data but not, as is usually the case, in the time frame in which macroeconomists are forced, by data availability, to work, namely, months and quarters.

[10] Not one of the conflict and mediation studies mentioned earlier offers error bands for the impulse responses of VAR models. These bands are also missing in Goldstein and Freeman (1990) study of the impact of counterfactual sequences of cooperation of U.S. foreign policy toward the Soviets in the context of the strategic triangle. Williams (1993) includes error bands on his VAR impulse responses, but could not generate error bands for his BVAR impulse responses because the distribution of the posterior for the time-varying BVAR he estimated is unknown.

A New Approach to Policy Evaluation in Political Science

Bayesian Vector Autoregression approach

The previous section reviewed our desiderata for forecasting. In what follows we present a Bayesian time series approach that meets these desiderata.

Our approach builds on recent work of Christopher Sims, Tao Zha and researchers at the Atlanta Federal Reserve Bank. They have worked to extend earlier work on Bayesian VAR models (e.g., Litterman (1986) and Doan, Litterman and Sims (1984)) to account for recent advances in computation and estimation for Bayesian models in general. Their work advances the state of the art in Bayesian VAR (BVAR) models in two ways. The first is showing how a simple class of prior distributions can give rise to a tractable posterior distribution for model parameters and forecasts, that offer significant improvements over classical or non-Bayesian VAR forecasting models. The gains came through the use of meaningful or informative priors that embody widely held economic beliefs about the properties of macroeconomic time series. Recent work improves on the specification of this prior and creates a conjugate form with a known posterior distribution. Leeper et al. (1996), and Sims and Zha (1999) propose the use of a "reference prior." This prior embodies many of the widely held beliefs about macroeconomic time series and allows for the computation of a tractable posterior distribution for VAR model parameters and forecasts. This in in contrast to the approach of Litterman (1986) and Kadiyala and Karlsson (1997) where the asymmetry of the prior across variables and equations in the VAR model meant that the treatment of "own versus other lags" invalidated the use of a normal distribution for the posterior of the model. Sims and Zha's reformulation means that a tractable theory can be built to sample and draw inference about the probability distribution of VAR parameters and forecasts.

The second contribution of this recent research has been to expand the inference and forecasting techniques for VAR models. Sims and Zha (1998) present a likelihood based technique for assessing the shape and scale of impulse responses — in contrast to the widely used Monte Carlo procedure that generates symmetric pointwise confidence intervals for impulse responses.

These developments of Bayesian time series models, particularly the ability to generate tractable posterior distributions for such highly parameterized models (in contrast to earlier Bayesian VAR models) have produced major advances in forecasting and policy analysis. The work of Sims and Zha (1998, 1999) and Waggoner and Zha (1999) provide a framework for using BVAR models for policy analysis and inference. They show how informative priors can be used to generate both the parameter and forecast error distributions for BVAR models. This means that we can account for *both parameter and forecast uncertainty* in our conditional forecasts and generate error bands and posterior intervals around all of our quantities of interest.[11]

Our desiderata are met by this new Bayesian time series approach. First, these models are dynamic and multivariate by construction. The BVAR model allows us to induce the dynamic processes and interrelationships in the endogenous variables without imposing a structural model on the data.

Second, the BVAR models allow us to explicitly model and test for complex causal relationships. This allows us to investigate a wide variety of both contemporaneous and lagged causal effects. In addition, the BVAR models can be constructed to account for explicit theoretical structural models using simultaneous equations and structural VAR (SVAR) techniques (Leeper, Sims and Zha 1996, Amisano and Giannini 1997).

Third, BVAR models easily allow us to gauge the effects of temporal and cross-sectional

[11] Political scientists have also contributed a number of Bayesian models of political phenomena that can serve as a guide to our work. Williams (1993) presents a time-varying Bayesian VAR model that replicates work from Goldstein (1988). However, because of the intractability of the earlier Litterman prior, no closed form distribution can be simulated to generate error bands for Williams' BVAR impulse responses. This limitation has been removed by the recent advances cited earlier.

aggregation. We can easily rerun similar models for different groups of countries / nations / states / actors to see the effects of aggregation. The advent of new data sources and event data generation tools such as CAMEO, KEDS, TABARI, for international conflict analysis means that data aggregation and disaggregation issues will likely be more important.

Further, there are well developed methods for capturing and accounting for model and forecast uncertainty with these methods. Waggoner and Zha (1999) outline a Gibbs sampling algorithm that allows us to account for both sources of uncertainty.

Finally, the Bayesian time series approach allows us to include prior information about the data. In general, this is what we see as its main advantage. It allows us to bring together insights from the "wisdom literature" that offer guidance about the nature of international conflict, its dynamics, its persistence, and its responses to shocks of various kinds. These ideas can be incorporated as an informative prior for a BVAR model and thus enhance its forecasting performance and usefulness in policy analysis.

BVAR forecast design

To address our desiderata for Bayesian forecasting, we now propose a five step research design. The goal is to offer forecast summaries that account for analysts' uncertainty about model specification and about the forecasts themselves. The five components of our approach are

1. Case selection and data choices
2. Initial VAR specification with testing for lag length and causality
3. Specification of the prior distribution of VAR coefficients to incorporate qualitative knowledge about the subject political process.
4. Evaluation of policy counterfactuals and conditional forecasting.
5. Comparison and evaluation of forecasts and policy contingent (conditional) forecast densities.

The first two steps of our approach amount to existing practices for building multivariate time series models such as VARs. We advocate the selection of specific cases and data instead of pooling data. This is contrary to some of the accepted wisdom and practice in political science, particularly in international relations and political economy where large cross-national time series models using time-series-cross-sectional (TSCS) methods have become the norm. However, it recognizes the point made by Meier and Gill (2000) to account for more than "average effects" in a policy analysis.

As regards the specification of a VAR model, we first establish a proper lag specification for the model using information criteria tests (Lutkepohl 1993). We also establish causal linkages using Granger causality tests. This is important, particularly for policy based forecasts and counterfactuals because we must be able to establish a certain causal structure among the variables in our model. Absent any statistically significant causal relationships between our policy instruments and policy outcome variables, we will be unable to determine the forecasted effects of policy changes in our variables of interest.

Third, it is well known that VAR models are highly parameterized, and therefore their forecasts trade off bias and efficiency. Their high degree of parameterization means they tend to generate unbiased forecasts, but these forecasts tend to be relatively inefficient (Litterman 1986). One way that can be used to improve the quality of inferences and forecasts from such a model is to use a prior and estimate a Bayesian VAR.

The Doan, Litterman and Sims (1984) prior or "Litterman prior" advocated by Litterman (1986) is based on the idea that the best predictor of a vector of time series is the vector of its past values. Therefore, for stationary data, the sample mean is the best predictor; for non-stationary data, a random walk is the best model. Litterman's framework allows us to attach a prior belief to these intuitions by specifying a mean and scaling the standard deviation around autoregressive parameters in a VAR. The tighter the belief, the smaller the standard error around the autoregressive coefficients in the VAR. This prior belief is set equation-by-equation for the VAR and allows for the lags of the dependent variable in the equation ("own lags") to have a smaller standard deviations than other variables ("other lags") in the same equation. This is a belief that each variable is better explained by its past values than other values. This asymmetry in the prior, however, makes the posterior distribution of the models parameters intractable.

The reference prior used by Sims and Zha (1999) builds on Litterman (1986) and Doan, Litterman and Sims (1984). Unlike the Litterman version of the BVAR prior, the Sims-Zha reference prior treats lags of other variables in the $i'th$ equation the same as the $i'th$ equation left hand side variable. This makes the prior tractable and allows for a posterior that can be computed, unlike the earlier versions (See Kadiyala and Karlsson (1997), Sims and Zha (1998), and Robertson and Tallman (1998) for a discussion and comparison of these two priors).

To fix ideas, consider the following structural VAR model (matrix dimensions indicated below matrices).[12]

$$\sum_{l=0}^{p} \underset{1\times m}{y_{t-l}} \underset{m\times m}{A_l} = \underset{1\times m}{d} + \underset{1\times m}{\epsilon_t}, \quad t = 1, 2, \ldots, T. \tag{1}$$

This is an m-dimensional VAR for a sample of size T with y_t a vector of observations at time t, A_l the coefficient matrix for the l^th lag, p the maximum number of lags, d a vector of constants, and ϵ_t a vector of i.i.d. normal *structural shocks* such that

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

and

$$E[\epsilon_t | y_{t-s}, s > 0] = \underset{1\times m}{0}.$$

From this point forward, A_0, the contemporaneous coefficient matrix is assumed to be non-singular and subject only to linear restrictions.

[12]Here we use the term "structural," in a manner consistent with the VAR literature, to denote a model that is a dynamic simultaneous system of equations. The model is structural in that its interpretation and estimation require us to make an assumption about the structure of A_0, the decomposition of the error covariance matrix. In what follows, we assume that this is a Cholesky triangular decomposition of the covariance matrix of the residuals. See Leeper et al. (1996) for a discussion of alternatives.

Before proceeding, define the following compact form for the VAR coefficients in equation (1):

$$a = \begin{pmatrix} a_0 \\ a_+ \end{pmatrix}, \quad \text{where} \quad a_0 = vec(A_0), \quad \text{and} \quad a_+ = vec \begin{pmatrix} -A_1 \\ \vdots \\ -A_p \\ d \end{pmatrix} \tag{2}$$

Under general conditions, using either a flat or reference prior, Sims and Zha (1998) show that conditional on a_0, the posterior distribution of these VAR coefficients is of a multivariate normal form. Thus, forecasts can be generated by exploiting the multivariate normality of the posterior distribution of the coefficients and the innovations.

The Sims-Zha prior is specified by positing a mean for $A_1 = I$, allowing $A_2 \dots A_p$ to decay, and scaling the variance of the VAR coefficients by setting the seven parameters in Table 1. The scale of the prior covariance on the VAR coefficients is determined by a series of AR(p) regressions for each variable. These scale factors are then adjusted by these seven hyperparameters. The resulting prior assumes that the VAR regression parameters are multivariate normal with the means given by the coefficients on $A_1, \dots, A_p$ and a covariance matrix for Σ that follows a Wishart distribution. Since the prior applies to the entire system of VAR equations, estimation of the BVAR posterior coefficients is done for the entire system, not equation-by-equation as in the case of the maximum-likelihood estimator for a VAR model. Estimation is done using Theil's mixed estimation method for a seeminly unrelated regression (SUR) model (Theil and Goldberger 1961, Theil 1963).

The prior incorporates our beliefs about the parameters in the BVAR model. For λ_0 we set the overall scale of the prior covariance matrix relative to the estimated covariance scale estimated from the univariate AR OLS models. The parameter λ_1 establishes the standard deviation around the prior for A_1, the AR(1) coefficients in the VAR model. The parameter λ_2 sets the weight of own versus other lags. This is set to 1 in the Sims-Zha prior, to keep the model tractable. Relaxing this assumption produces the Litterman prior. The lag decay

Parameter	Values	Interpretation
λ_0	[0,1]	Overall scale of the error covariance matrix Σ
λ_1	> 0	Standard deviation around A_1
λ_2	$= 1$	Weight of own lag versus other lags
λ_3	> 0	Lag decay (harmonic)
λ_4	≥ 0	Scale of standard deviation around constant term
μ_5	≥ 0	Sum of coefficients / Cointegration
μ_6	≥ 0	Initial observations / dummy observation
ν_0	> 0	Prior degrees of freedom, $m + 1$

Table 1: Hyperparameters of Sims-Zha BVAR reference prior

is set by λ_3. A value equal to one indicates harmonic lag decay in the VAR — larger values indicate more rapid decay.[13] The parameters μ_5 and μ_6 allow us to provide weights to models with unit roots and cointegration without imposing exact cointegrating relationships.

With the Sims-Zha prior, we can utilize expert or delphic information about the dynamics and causal relationships in the subject political or policy process. For instance, suppose we believe that our policy output variable is highly persistent and subject to large random shocks. Then we would place a tight prior on the first order autoregressive coefficients by setting λ_1 to be small. At the same time we would set λ_4 to be large to reflect the variability in the mean of the process. Similarly, if we think that policy shocks die out rapidly, we would speed the lag decay by choosing larger values of λ_3.

The fourth step in the policy evaluation is the construction of relevant policy rules and forecasts. Doan, Litterman and Sims (1984) note that we may know the path of one endogenous variable in a dynamic system of simultaneous equations before we see another (such as unemployment which is measured monthly, while GNP remains unobserved until the end of the quarter). This is not however, the only form of conditional or contingent forecast we might construct. We also could hypothesize alternative paths for a policy variable such as the level of U.S. mediation or trade sanctions in an international conflict and then look at

[13] The parameter λ_3 is used to scale the variance of the $j'th$ AR coefficients by a factor of $\frac{1}{j^{\lambda_3}}$. So for $\lambda_3 = 1$, we get harmonic decay. Larger values indicate faster decay.

the resulting forecasts of the conflict. In both cases, we are placing a set of constraints on the forecasts we can make, because the estimated error covariance sets the correlation between the forecasts for the variables in the VAR forecasts. This idea led Doan, Litterman and Sims to derive the set of linear conditions on forecast innovations implied by the simulated path of policy variables.

Waggoner and Zha (1999) extend this idea. They derive the mean and variance of these constrained forecasts. Further, they demonstrate how to use information about the forecasts' innovations subject to constraints on the forecast of one or more endogenous variables in a VAR to generate conditional forecast distributions that account for both parameter uncertainty and forecast uncertainty. Waggoner and Zha do this by using Gibbs sampling with data augmentation to generate a sequence of model estimates and forecasts that summarize the conditional forecasts and their associated uncertainty.

There are two ways we can proceed to construct a such policy counterfactual. The first uses a *hard condition* to specify the path of a given endogenous variable. A hard condition sets the value of an endogenous variable to a fixed value or path of values. Alternatively, we could use a *soft condition* and posit a range of values for this policy variable. For instance a hard condition for an international conflict model assumes that the level of mediation remains at a fixed level for some time into the future. A soft condition assumes that mediation take on one of a range of values over some future horizon.[14]

Formally, consider the forecast equations for the reduced form VAR model:

$$y_t = c + \sum_{l=1}^{p} y_{t-l} B_l + \epsilon_t A_0^{-1} \tag{3}$$

[14]The canonical example here is the examination of monetary policy where the Federal Reserve Funds rate (FFR) is either fixed at a given value as part of a Federal Reserve policy rule (hard condition) or a range of values greater than some level is examined (a soft condition). In both cases, the forecast paths of the different policy rules are traced out to see the effects on GNP and the economy at large. See Waggoner and Zha (1999).

As in Waggoner and Zha (1999) the reduced form forecast in (3) is related to the structural model in (1) by

$$c = dA_0^{-1} \quad B_l = A_l A_0^{-1}, \quad l = 1, 2, \ldots, p, \quad \text{and} \quad \Sigma = A_0^{-1} A_0^{-1'}$$

Given these forecast functions, the h-step ahead forecasts can be written as:

$$y_{T+h} = cK_{h-1} + \sum_{l=1}^{p} y_{T+1-l} N_l(h) + \sum_{j=1}^{h} \epsilon_{T+j} M_{h-j}, \quad h = 1, 2, \ldots \tag{4}$$

where

$$\begin{aligned}
K_0 = I, \qquad K_i &= I + \sum_{j=1}^{i} K_{i-j} B_j, \quad i = 1, 2, \ldots; \\
N_l(1) &= B_l, \quad l = 1, 2, \ldots, p; \\
N_l(h) &= \sum_{j=1}^{h-1} N_l(h-j) B_j + B_{h+l-1}, \quad l = 1, 2, \ldots, p, \ h = 2, 3, \ldots; \\
M_0 = A_0^{-1}, \quad M_l &= \sum_{j=1}^{i} M_{i-j} B_j i = 1, 2, \ldots;
\end{aligned}$$

where we use the convention that $B_j = 0$ for $j > p$.

This h-step forecast equation (4) gives the dynamic forecasts with structural shocks. It shows how these forecasts can be decomposed into the components with and without shocks. The first two terms in the sum are the effects of the past lagged values of the series and the constant or trends. The key point in conditional forecasting is that setting the path of one variable, say y_{1t}, constrains the possible innovations in the forecasts of $y_{2t} \ldots y_{mt}$. To see this, consider the following formulation for a hard condition on a VAR forecast. Suppose that the value of the $j'th$ variable forecast is constrained to be $y(j)^*_{T+h}$. Then from equation (4) it must be that

$$y(j)^*_{T+h} - cK(j)_{h-1} - \sum_{l=1}^{p} y_{T+1-l} N_l(h)(j) = \sum_{j=1}^{h} \epsilon_{T+j} M_{h-j} \tag{5}$$

where the notation (j) refers to the $j'th$ column matrix.

The left hand side of (5) implies that the innovations on the right hand side are constrained. That is, there is a restricted parameter space of innovations that are consistent with the hypothesized conditional forecast. These constraints can be expressed as a set of encompassing conditions. These hard conditions take the form of linear constraints:

$$\underset{q\times k}{R(a)'}\underset{k\times 1}{\epsilon} = \underset{q\times 1}{r(a)}, \quad q \le k = mh \tag{6}$$

The elements of these matrices correspond to the forecast constraints. The notation assumes that there are q constraints, and there can be no more constraints than the number of future forecasts for all the variables, $k = mh$. In any case, the elements of R and r may depend on estimated parameter of the reduced form, denoted by a (the vectorized coefficients in equation (2)).[15]

It is this final fact that leads us to use a Gibbs sampling technique to generate the distribution of the conditional forecasts. It allows us to account for the path of the conditional shocks, and the possible uncertainty surrounding the parameters used to generate these conditional forecasts. We start by estimating a BVAR model and generating a conditional forecast from this model. We then use this conditional forecast to augment the data and resample the parameters. This procedure accounts for both sources of uncertainty in the forecasts.[16] We outline this algorithm and its notation in an Appendix.

[15]For a soft condition, $r(a)$ is not a vector, but a set that contains the admissible forecast values for the forecast condition on the $j'th$ variable. See Waggoner and Zha (1999) for a discussion.

[16]As Waggoner and Zha (1999, 642-643) note, the sampling of the model parameters is

> ... a crucial step for obtaining the correct finite-sample variation in parameters subject to a set of hard conditions in constraints .... Because the distribution of parameters is simulated from the posterior density function, the prior plays an important role in determining the location of the parameters in finite samples. Under the flat prior, the posterior density is simply proportional to the likelihood function, which, in a typical VAR system, is often flat around the peak in small samples. Moreover, maximum-likelihood estimates tend to attribute a large amount of variation to deterministic components (Sims and Zha 1998). Such a bias, prevalent in dynamic multivariate models like VARs, is the other side of the well known bias toward stationarity of least-squares estimates. These problems can have substantial effects on the distribution of conditional forecasts ...

The final component of our policy evaluation approach is a comparison of conditional forecast densities to unconditional forecast densities. This allows us to see how changes in policy induce different policy outcomes. If our policy counterfactuals are meaningful, we should see forecast densities that are preferable to the expected paths of our variables. It is this that we refer to as "moving mountains" because relevant changes to policy should move the conditional forecast density to a region that is preferred to the unconditional forecast density.

Discussion

There are a number of issues that arise in using BVAR models to evaluate policy changes. Here we offer a brief review of some of the issues and concerns that may emerge in applying our method.

First, the specification of policy counterfactuals is not always obvious. Our approach requires us to have a clear understanding of the possible rules that should be used to generate conditional and unconditional forecasts. This is necessary to specify the conditions for the Gibbs sampling algorithm outlined in the Appendix.[17]

Second, the BVAR conditional forecasting model proposed above provides a method for generating error bands for forecasts. This is because we can generate the density for the forecasts and model parameters through Gibbs sampling. This Gibbs sampler, though, raises interesting convergence issues. The number of forecast points is the number of variables (m) by the number of forecast points (h). This means we need to check the convergence of an $m \times h$ matrix for convergence. For this, we use a conservative criteria, checking the convergence of each variable at each forecast point using a Geweke diagnostic. An alternative would be to evaluate the matrix of forecasts and assess convergence of the matrix across the Gibbs iterations.

[17] Here we focus on the evaluation of hard conditional forecasts. For the soft condition case, see Waggoner and Zha (1999, 643-4).

Third, comparisons of conditional and unconditional forecasts raise issues of forecast density evaluation. As noted in Granger (1999, Chapter 3), our policy recommendations may hinge on the cost functions we use. In some cases, we know that the costs of one outcome is significantly more costly than an alternative — such as increased conflict versus cooperation in international relations, or budget deficits versus budget surpluses for fiscal policy data. Thus, we must be aware that cost functions may be asymmetric when evaluating density forecasts (King and Zeng 2001*b*).

Fourth, our approach still must deal with specification issues. These issues include variable selection, model scale, lag specification, and the selection of the forecast horizon. The model scale is determined in part by the selection of variables. The risk is that we may have omitted a variable that is Granger causal or that affects the Granger causality relationships in our model. Thus, our forecast performance would suffer.[18]

The selection of the forecast horizon also raises some issues. For instance, in the analysis of Palestinian-Israeli conflict we present below, the selection of the forecast period and the sample of data used can have significant effects on our inferences and forecasts because there are different breakpoints or epochs in the data (post-Camp David, the intifada, Oslo Accords, etc.) Therefore, we must be sensitive to the fact that we may have different models that describe different epochs.

Fifth, are computational concerns. The estimation of the BVAR models is relatively cheap. The models are simply a large regression model with the number of regression parameters equal to $m^2p + m$, where m is the number of endogenous variables and p is the number of lags (for a model with no exogenous variables and only an intercept). The problem is drawing from the distribution of the model parameters (for the normal-Wishart prior this is $m^2p + m$ regression coefficients and an $m \times m$ Wishart covariance matrix) *and* the

[18] Litterman and Weiss (1985, 137-8) note that the inclusion of variables has only a knife-edged effect on findings of exogeneity because adding new variables can reverse findings of non-causality only in highly special circumstances.

constrained forecast innovations for the hard condition. This is a matrix of $m \times h$ structural shocks. The speed of the Gibbs sampler simulation is fixed by the ratio of the draws of the regression parameters to the structural innovations. For the simulations presented below, with a burnin of 500 draws and a final sample of 1000 forecasts from a 6 variable, six lag BVAR model with 10 forecast periods takes about 35 minutes.

Sixth, we face the problem of translating the qualitative information of experts into the quantitative prior we use in the BVAR. At present, we advocate the use of expert policy histories to assess the properties of data, the dynamics of the forecasts, and the information about the series of interest. This part of the process is where political scientists have the most work to do since unlike economists, we do not have the twenty-plus years of experience with these models and data (cf., Sims and Zha (1998, fn. 7)).

Finally, our approach is subject to but offers some response to the Lucas critique of econometric and policy analysis (Lucas 1976). Lucas notes that if policy makers have rational expectations about policy choices, they would not use the same econometric model to predict policy outcomes. Thus, rational expectations about policy undermine econometric modeling because analysts do not account for the possible changes policies spawn in the decisions of economic agents. We believe however, that the Gibbs sampling approach to VAR models that we use in our conditional forecasts allows us to look at the set of *plausible* parameter values. Since a policy maker has some belief about the range of parameters that generate the policy outcome, in drawing from this distribution we can start to account for the possible changes in parameters and incorporate this information into our forecasts (on this point see Zha 1998, esp. p.19).

Illustration

> "... the Middle East will burn with American interests and even lives going up in flames unless the U.S. intervenes swiftly and much more neutrally in the conflict."
> *The Economist*, February 2, 2002, p. 14.

International conflicts are enduring feature of our Post-Cold War world. From southern Europe to the Middle East and Africa to Asia and Latin America, armed struggles continue to rage. The United States increasingly is asked to serve as a mediator in these conflicts. The United States is asked to use its resources to end the violence and bloodshed and, if possible, create the conditions that can produce lasting peace. The conflict between the Israelis and the Palestinians is illustrative. For decades, these two peoples have battled one another. Since the end of World War II the U.S. has been more and more involved in this conflict. Today, as this conflict reaches a new phase, people once again ask the U.S. to intervene. For the U.S. and other potential mediators, solving the Israeli-Palestinian conflict seems tantamount to "moving mountains."

Political scientists have studied the Israeli-Palestinian conflict for many years. Among recent quantitative statistical investigations of it are Schrodt et al. (2001) and Goldstein et al. (2001). Both of these studies employ the Kansas Events Data System; each use WEIS codes. Schrodt et al. (2001) is a collection of exploratory analyses of the impacts of mediation on the behavior of the belligerents in the period April 1979-September 1999. They use simple time series regression and cross-correlation methods to analyze the Israeli-Palestinian conflict. They find evidence that U.S. mediation is motivated by and has a salutary impact on this conflict. Full-blown, multivariate time series methods are used in Goldstein et al. (2001). These researchers find evidence of "triangularity" between U.S. behavior toward Israel and the Palestinians, Israeli behavior toward the Palestinians and Palestinian behavior toward Israel. As Goldstein et al. (2001: 612) discuss: "Israeli and Palestinian (behaviors) were reciprocal, indicating that cooperation or conflict received from the United States was 'passed along' in kind to the neighbor." They argue that this triangularity provides the basis for the evolution of cooperation between the Israelis and Palestinians or, that there is potential for effective U.S. mediation of this conflict.[19]

[19]Schrodt et al. (2001) define mediation as cooperative behavior sent to both belligerents in a seven day period. They work at the monthly level of temporal aggregation and find much more evidence that U.S. is

Unfortunately, neither study is methodologically sound. Schrodt et al. (2001) do not meet our second, fourth and fifth desiderata — their investigation does not allow for complex, multivariate causal relationships between behaviors; they do not gauge the impact of model uncertainty; and they make no effort to systematically incorporate the insights of the "wisdom literature" on international mediation into their analysis. The fourth and fifth desiderata also are not met by Goldstein et al. (2001). They rely solely on point estimates in assessing the impact of U.S. mediation, for example.[20] In addition, both studies also suffer from technical problems, most important, the use of time series that are improperly indexed (for leap year days). See Brandt (Forthcoming) for a discussion of these issues.

Conditional Forecast Design for Levant Data

As an initial application of our method, we analyze a mediation counterfactual at the end of 1988. In particular, we examine the hypothetical impact of U.S. mediation in the aftermath of Yasser Arafat's meeting U.S. demands (1975-1988) to renunciate all forms of terrorism and accept U.N. Resolutions 242 and 338 (Gerner (1994, 142-4) and Morris (2001, 608-610)). We make a "hard forecast" of Israeli-Palestinian conflict under the assumption that immediately after Arafat's capitulation on December 14, 1988, the U.S. counterfactually behaved for the next ten days in a markedly cooperative way toward Israel. To be more specific, our model was based on data with the same origination date as Schrodt et al. (2001) and Goldstein et al. (2001), beginning April 15, 1979. But our series are terminated on December 15, 1988 — the date on which Arafat met U.S. demands; the period of our contingent forecast was December 16, 1988 to December 25, 1988. Note that our estimation period was after the

effective in the Levant than in the Serbian-Bosnian conflict. Goldstein et al. (2001) focus on U.S. cooperative behavior toward the Israelis and Palestinians. They use weekly aggregates for some of their specification tests, but focus on daily data for their analysis of reciprocity and triangularity. The VAR model they employ has 24 variables with 28 (1979-1990) or 6 (1991-1995) lags on each variable. Both studies find high levels of contemporaneous correlations of variables (residuals).

[20] To be fair, Schrodt et al. (2001) describe their analyses as "plausibility probes." They do not propose, let alone defend a (non)structural model for mediation. Goldstein et al. (2001) essentially employ the statistical multivariate time series technology of the early 1990s.

Camp David Accords and before the Madrid conference, Oslo Accords, and Gulf War. This period also is one in which there were unity governments in Israel and the PLO was much more unified (arguably) than it is today. The U.S. government was, at least in comparison to the Nixon administration, more unified as well.[21]

To achieve further comparability to Schrodt et al. (2001) and Goldstein et al. (2001) and because it represents one of the most complete data sets for the Levant, we used the the same KEDS series in our analysis. We analyzed the daily aggregates of the WEIS coded events. However, in contrast to Schrodt et al and Goldstein et al, we indexed our series for leap days (see Brandt (Forthcoming)). Because this is an illustration of method, we focused on the causal nexus between six variables: American behavior toward (from) Israel, AI, IA; Israeli behavior toward (from) the Palestinians, IP, PI; and American behavior toward (from) the Palestinians, AP, PA. We ordered the system AI, AP, IA, PA, IP, PI. In fact, there is little reason to believe that findings of exogeneity between these six variables are affected by omitted variable bias (Litterman and Weiss 1985, 137-8). Also, the Bayesian vector autoregressive analysis is not affected by an orthonormal transformation of the variables (Waggoner and Zha 1999, 641)

The results of the model building phase were as follows. The AIC statistic indicated that six lags of the variables were sufficient ($p = 6$). According to the causality (F-tests), the basic causal nexus for the conflict can be depicted as in Figure 1. Like Goldstein et al. (2001), we find reciprocity in the system: IP and PI are strongly related. This is consistent with the historical literature with its emphasis on cycles of violence and revenge (Smith 2001, 142,197, 438, 467), as well as with the announced policies of "reciprocity" of the actors involved (e.g., see Smith (2001, 479) and Morris (2001, 645). Like Goldstein et al. (2001) and Schrodt (2001) et al., we too find of triangularity — there is evidence that AI and AP have a causal

[21]We thank Phil Schrodt for suggesting this mediation counterfactual as a good starting point for our research. Most historical accounts agree that Arafat's capitulation to U.S. demands—the expressed policy of the American government toward the PLO for almost 14 years—was a major watershed in the conflict.

impact on the IP and PI. The potential for mediation therefore exists in our sample period.

Finally, our forecast experiments revealed that in terms of RMSE and MAE, a specification with the following hyperparameters performed best: $\lambda_0 = 0.6$, $\lambda_1 = 0.1$, $\lambda_2 = 1$, $\lambda_3 = 2$, $\lambda_4 = 0.1$, $\mu_5 = \mu_6 = 0$, and $\nu = 7$.[22] Table 2 compares other specifications of the prior that embody different beliefs. These comparisons are generated by estimating a BVAR model with the specification in the row for the sample period of our analysis (April 4, 1979 to December 15, 1988). We then use these estimates to generate an unconditional forecast for the subsequent ten days. We compared these forecasts to the actual series for these ten days to generate the RMSE and MAE measures. The priors in this table are then ranked by their RMSEs; models with flat priors are listed in the first two rows of the table.

Searching over priors in this manner is necessary because we have little experience with the exact numerical parameters that will translate our qualitative information about the nature of Middle East conflict into a quantitative prior for a BVAR model. This search over the prior parameter space is loosely based on similar work in the early applications of BVAR models such as Doan, Litterman and Sims (1984) and Litterman (1986) where various specifications of the prior are evaluated to determine their suitability for economic forecasting and policy evaluation. It is our expectation that as we analyze more regional conflicts and interact with foreign policy makers, we will be able to develop a better "reference prior" for such data.

Once we had determined a set of plausible parameters for the BVAR prior, we turned to conditional forecasting. In order to demonstrate the workings and usefulness of the Bayesian approach to contingent forecasting, we generated mediation contingent forecasts for the

[22] With the help of Timothy Hellwig we replicated the results in Goldstein et al. (2001) for the data that are not indexed for leap days. When the corrected data and AIC rather than the Likelihood Ratio test are employed, only six lags are needed to describe the data. Also, some of causal relationships in the Goldstein et al. (2001) paper are altered. These reanalyses are available from the authors.

λ_0	λ_1	λ_3	λ_4	p	RMSE	MAE
Flat	Prior			6	5.0295	3.3091
Flat	Prior			12	5.0938	3.4337
0.6	0.1	2	0.1	6	4.4548	2.8386
0.6	0.1	2	0.1	12	4.4575	2.8413
0.8	0.1	2	0.1	6	4.4678	2.8554
0.6	0.1	2	0.25	6	4.4688	2.8562
0.6	0.1	2	0.5	6	4.4711	2.8590
0.8	0.1	2	0.1	12	4.4711	2.8588
0.6	0.1	2	0.25	12	4.4715	2.8588
0.6	0.1	2	0.5	12	4.4737	2.8615
0.8	0.1	2	0.25	6	4.4762	2.8654
0.8	0.1	2	0.5	6	4.4775	2.8668
0.8	0.1	2	0.25	12	4.4795	2.8686
0.8	0.1	2	0.5	12	4.4808	2.8701
0.6	0.25	2	0.1	6	4.5072	2.8945
0.6	0.25	2	0.1	12	4.5132	2.9003
0.6	0.25	2	0.25	6	4.5208	2.9089
0.6	0.25	2	0.5	6	4.5230	2.9111
0.6	0.25	2	0.25	12	4.5268	2.9144
0.6	0.25	2	0.5	12	4.5289	2.9165
0.8	0.25	2	0.1	6	4.5532	2.9347
0.8	0.25	2	0.1	12	4.5606	2.9419
0.8	0.25	2	0.25	6	4.5608	2.9422
0.8	0.25	2	0.5	6	4.5620	2.9434
0.8	0.25	2	0.25	12	4.5682	2.9500
0.8	0.25	2	0.5	12	4.5694	2.9514
0.6	0.1	1	0.1	6	4.6416	3.0172
0.6	0.1	1	0.25	6	4.6528	3.0301
0.6	0.1	1	0.5	6	4.6546	3.0321
0.6	0.1	1	0.1	12	4.6783	3.0613
0.6	0.1	1	0.25	12	4.6888	3.0725
0.6	0.1	1	0.5	12	4.6905	3.0742
0.8	0.1	1	0.1	6	4.7178	3.0844
0.8	0.1	1	0.25	6	4.7234	3.0904
0.8	0.1	1	0.5	6	4.7243	3.0914
0.8	0.1	1	0.1	12	4.7597	3.1342
0.8	0.1	1	0.25	12	4.7651	3.1395
0.8	0.1	1	0.5	12	4.7659	3.1403
0.6	0.25	1	0.1	6	4.8724	3.2014
0.6	0.1	0.5	0.1	6	4.8771	3.2071
0.6	0.25	1	0.25	6	4.8786	3.2076
0.6	0.25	1	0.5	6	4.8796	3.2086
0.6	0.1	0.5	0.25	6	4.8835	3.2135
0.6	0.1	0.5	0.5	6	4.8846	3.2146
0.6	0.25	1	0.1	12	4.9150	3.2707
0.6	0.25	1	0.25	12	4.9227	3.2769
0.6	0.25	1	0.5	12	4.9239	3.2779
0.8	0.25	1	0.1	6	4.9289	3.2416
0.8	0.1	0.5	0.1	6	4.9305	3.2442
0.8	0.25	1	0.25	6	4.9320	3.2446
0.8	0.25	1	0.5	6	4.9325	3.2451
0.8	0.1	0.5	0.25	6	4.9336	3.2473
0.8	0.1	0.5	0.5	6	4.9341	3.2478
0.6	0.1	0.5	0.1	12	4.9417	3.3028
0.6	0.1	0.5	0.25	12	4.9495	3.3092
0.6	0.1	0.5	0.5	12	4.9507	3.3102
0.8	0.25	1	0.1	12	4.9731	3.3231
0.8	0.25	1	0.25	12	4.9777	3.3267
0.8	0.25	1	0.5	12	4.9784	3.3272
0.6	0.25	0.5	0.1	6	4.9924	3.2834
0.8	0.1	0.5	0.1	12	4.9946	3.3487
0.6	0.25	0.5	0.25	6	4.9965	3.2875
0.6	0.25	0.5	0.5	6	4.9971	3.2881
0.8	0.1	0.5	0.25	12	4.9993	3.3522
0.8	0.1	0.5	0.5	12	5.0000	3.3528
0.8	0.25	0.5	0.1	6	5.0080	3.2943
0.8	0.25	0.5	0.25	6	5.0103	3.2966
0.8	0.25	0.5	0.5	6	5.0107	3.2970
0.6	0.25	0.5	0.1	12	5.0522	3.3990
0.6	0.25	0.5	0.25	12	5.0603	3.4052
0.6	0.25	0.5	0.5	12	5.0617	3.4062
0.8	0.25	0.5	0.1	12	5.0694	3.4132
0.8	0.25	0.5	0.25	12	5.0743	3.4170
0.8	0.25	0.5	0.5	12	5.0751	3.4175

Table 2: Levant BVAR Priors and Forecast Performance. $\mu_5 = \mu_6 = 0$ for all these analyses

following subset of our BVAR prior parameters:

$$\lambda_0 = 0.6, \qquad \lambda_1 = \{0.1, 0.25\},$$
$$\lambda_2 = 1, \qquad \lambda_3 = 2,$$
$$\lambda_4 = \{0.1, 0.25, 0.5\}, \qquad \nu = 7.$$

We evaluated these priors to see what impact they have on our conditional forecasts, and as a check on the robustness of our results. These priors allow us to evaluate the impact of tighter or looser beliefs about the persistence of the conflict dynamics; where $\lambda_1 = 0.1$ is a tight belief and $\lambda_1 = 0.25$ is a loose belief. Also, we vary the prior on the constant's standard deviation to account for different beliefs about the variance of the series.

On the basis of discussions with Phil Schrodt and others, we kept the hyperparameters for unit roots and cointegration and also for the influence of the initial condition ($\mu 5$, $\mu 6$) at zero. In this way, we explicitly incorporate experts' opinions that the Israeli-Palestinian conflict is stationary. Operationally, our mediation counterfactual amounted to the assumption that the U.S. behaved cooperatively toward Israel at daily aggregate levels of 5 and 10 for each of ten days immediately after Arafat met the U.S. demands for negotiation — so AI is fixed at 5 or 10 over the ten days we forecast. This amounts to the counterfactual assumption that on each of days in the forecast the U.S. government took one or more actions toward Israel the (combined) impact(s) of which would have been coded (together) at 5 or 10, e.g., on one of the days between 12:15:88 and 12:25:88, the U.S. promised policy support (Goldstein scale value of [4.5]) and (or) material support (Goldstein scale value of [5.2]) .[23]

[23] We thus define mediation like Goldstein et al. (2001). Future work on the model will employ the definition of cooperative behavior toward both belligerents employed by Schrodt et al. (2001). We return to the issues of model scale and variable order below. Goldstein et al. (2001, fn. 35) indicate that they checked for unit roots in their data and found none. Schrodt et al. (2001) use a deterministic time trend in some of their structural models. But, in personal communications, Schrodt indicates that he has no clear justification for this practice. And he is quite sure that for the period we are studying the Israeli-Palestinian conflict is stationary. For the period for which we fit the model, 1979:4:15-12:15:88, the mean of AI was .35 with standard deviation 2.52. Thus the mediation counterfactuals of 5 and 10 are about 2 and 4 standard deviations larger than the mean, respectively. The question of whether this is a "factual counterfactual" (King and Zeng 2001*a*) is an open question for future research.

The forecasts are generated as follows: Using the algorithm outlined in the Appendix, we estimated a BVAR model with one of the priors specified above, subject to the relevant mediation counterfactuals (constraints on the future values of AI). We then used 500 burn-in draws for the Gibbs sampler, followed by a final 1000 draws. These final 1000 draws are the basis for forecast summaries we present below. At each iteration of the Gibbs sampler, we used the previous iterations forecast to augment the data and reestimate BVAR parameter densities.

What do we expect to find in our analysis? In terms of our contingent forecasts, methodologically, we expect that the posterior means for the flat prior or MLE forecasts will be much more irregular than those for the BVAR forecasts; the former also ought to have much wider error bands than the latter. Also on days later in the forecast period, the conditional densities for the mediation contingent (conditional) BVAR forecasts ought to be more leptokurtic than the densities for the (unconditional) MLE forecasts. Substantively, if the mediation counterfactual is likely to be successful, we should find (a) the contingent, BVAR forecasts show an increase in the posterior means of both IP and PI — ideally, increases that connote cooperation where the lower error band is above zero, (b), a marked shift in the cooperative direction in the mediation contingent (conditional) BVAR forecast densities for IP and PI and (c), an equivalent movement in the joint, conditional BVAR forecast densities for the two series. Our mediation contingent forecasts are alternative futures that might obtain if the U.S., Palestinians and Israel behaved in such a way that the counterfactual value of AI was repeatedly obtained. If policy could be altered to this counterfactual path, the forecast densities we generate are our predictions about the path of conflict between these actors.

These expectations are illustrated in Figure 2. In this figure, the left graph shows an example of the forecasts we expect to see — namely unconditional forecasts based on an classical or flat prior VAR model that are less than those conditional forecasts generated by a BVAR model with a mediation contingency. Further, in the right graph, when we look at the distribution of a daily forecast, we expect that the mediation contingent forecasts of IP and PI will indicate less conflict and have densities with smaller variance. The mediation contingent forecast densities for IP and PI will be centered over positive values connoting cooperation and will be more leptokurtic than the unconditional MLE forecast densities.

Results

The forecasts and a subset of forecast densities are presented in Figures 3-17. Each of these figures corresponds to a different specification of the BVAR prior. The forecasts in each row are for different mediation counterfactuals ($AI = 5$, $AI = 10$, or $AI = \mu + \sigma$ — its mean plus one standard deviation). The forecasts in red are the conditional BVAR forecasts; those in black are the unconditional VAR (flat prior) forecasts. For each forecast, we also present 90% probability intervals based on the Gibbs sampled forecasts.[24] The horizontal dashed line in each graph indicates the value of the series in the respective column on December 15, 1988, the last day of our sample estimation period.

For convenience and because they are the main focus of our analysis, we will concentrate on the behavior of IP and PI (the fourth and sixth columns of Figures 3-7. To begin with, the posterior means for the mediation contingent (conditional) BVAR forecasts are substantially smoother than those of the (unconditional) MLE forecasts. All five specifications of the Sims-Zha prior imply the counterfactual pattern of U.S. cooperation toward Israel causes less conflictual behavior on the part of Israel toward the Palestinians. At the same time, the mediation counterfactual causes an immediate shift to less hostility in Palestinians behavior toward the Israelis, a shift that is sustained for the entire ten days of counterfactual mediation.

In these figures, we present forecasts based on six different specifications of the prior.

[24]We checked the convergence of each of the forecasts pointwise (so we computed 6 variables $\times$ 10 days = 60 Geweke convergence statistics), one for each variable-forecast point we generated. According to this test, the convergence of the forecasts appears to be good.

Some of these priors reflect beliefs that the dynamics of the VAR model are highly persistent (those with $\lambda_1 = 0.1$) and rapidly mean reverting (those with $\lambda_4 = 0.1$). We also present conditional forecasts based on "looser" priors, where we allow more variance in the dynamics and the constant ($\lambda_1 = \lambda_4 = 0.25$). What we see across these figures is that the selection of a given prior has a large impact on the forecasts relative to the flat prior models. However, in terms of their implication for the Israeli-Palestinian conflict, which prior we choose matters little — they all generate qualitatively similar predictions. The real gain then derives from the use of an informative prior, embodying some qualitative beliefs about the nature of the data and conflict in the Levant; the specifics of the prior matter less in this case.

Figures 8-17 help us gauge the magnitude of this shift. Each of these figures is made up of four graphs showing the density of the IP and PI forecasts for December 25, 1988 with different priors. The black densities in the left column show the unconditional forecast densities for the MLE forecasts, while the red lines show the conditional forecast results for the specified BVAR model. The vertical line in these graphs show the value of the indicated series on December 15, 1988. The graphs in the right column of these figures show joint density of IP and PI for the forecast date. The bottom right graph shows a three-dimensional version of the densities with the BVAR conditional forecast density in red and the MLE unconditional forecast density in black. The top right graph shows a contour plot (or top down view) of the densities, with the same color scheme.

These density plots show that U.S. cooperation toward Israel probably would not have caused a major shift in IP from its December 15, 1988 value of -3.0. However, the U.S. mediation counterfactual would have created a major change in Palestinian behavior toward Israel: the BVAR mediation contingent forecast density shifts substantially to the right connoting much less hostility. This also can be seen in the right panels of figures 8-17. The solid circle in the lower right corner of these displays indicate the levels of behavior of the IP and PI on December 15, 1988. The stars in these panels indicate the mean levels of behavior of IP and PI for the sample (April 1979 to December 1998). The wider set of black concentric "circles" are a depiction of the unconditional MLE forecast for December 25, 1988 whereas the more compact set of red "circles" represent the mediation contingent, BVAR forecasts for this day. The latter are directly above the solid circles indicating the minor shift in Israeli behavior toward the Palestinians. But note that, in terms of PI, the joint, conditional forecast densities all suggest a shift to more cooperative behavior. U.S. mediation — whether two or four standard deviations above AI — would have diffused some Palestinian hostility toward Israel. This shift in PI moves the conditional forecast much closer to the long run mean of IP and PI (the star which is in the red contours) than the unconditional MLE forecast. Thus, regardless of our policy evaluation baseline (the value of conflict in December 15, 1988 indicated by the solid dot, or the long run mean level of conflict indicated by the star), the conclusion we draw is that the mediation would more quickly bring the level of conflict back to its long run mean than if there had been no mediation at all after December 15, 1988. This indicates that by either policy baseline we see improvement, and more rapid improvement based on the mediation contingent BVAR forecast.

Issues

As noted earlier, our approach poses some challenges for political methodologists and policy makers. For political methodologists, the challenges include handling model scale. There are reasons to believe that the behavior of regional and global actors influenced and continue to influence the Israeli-Palestinian conflict. For example, historians are convinced that Israeli-Syrian and U.S.-Soviet relations played an important role in the conflict in the first half of our time period, more specifically, during the Israeli-Lebanese war of 1982-1985. How many of these additional behavior linkages can be included in the model and analysis before it becomes uninterpretable or intractable is unclear. Closely related is the issue of breakpoints. It could be that the dynamics and causal relationships were different in the period following the Lebanese war (the last half of the 1980s). Therefore our model ought to be fit only for 1986-1988. It could also be that elections and other (ir)regular breakpoint events affect the models parameters.[25] Exactly where such breakpoints are located and how they should be determined is a matter of much research. Third, is the Lucas critique. Just as in economics, there is much evidence that the belligerents, on both sides, not only anticipate but purposely try to subvert mediation efforts. Some historians call such practices strategic revenge (one belligerent commits an atrocity in an attempt to force the other to stop talking to the mediator and engage in some form of equally aggressive retaliation; cf., for instance, Smith, 2001: 357-8 and Down and Steadman 2002: 81-2). Econometricians are not clear about the extent to which their approach accounts for this problem; the suggestion is that such strategic behavior is captured by the representation of forecast uncertainty — the error bands, for instance.[26] But whether this is true is unclear.

For policy makers, the challenges are numerous. Above all, there is the need to evaluate and hopefully refine the model specifications. We believe that policy makers' intuitions about international conflicts can improve our forecasts. Here we have only begun to translate those intuitions into hyperparameters. Much more work along these lines — perhaps even a new methodology for constructing such "informed priors" — needs to be done. Especially vital in this regard are the choices of λ_4, the parameter for the scale of the constant, and perhaps for μ_5, μ_6. Our exercise is quite short term in nature. But it incorporates long-term forces in our constant and, concomitantly, the assumption of stationarity. Substantively, we are

[25]There is evidence that domestic electoral politics affects both the propensity of countries to act as mediators (belligerents) and the receptivity of countries to mediation. See, for instance, Gaubatz (1991), Gowa (1998) and Fearon (1994).

[26]Zha (1998, 19) alludes to this point when discussing the application of dynamic multivariate models for forecasting the economy for the Federal Reserve:

> . . . dynamic multivariate modeling has complex structures in the sense that it allows both contemporaneous and dynamic interactions among the macroeconomic variables . . . [c]onsequently, both the Federal Reserve's complex behavior and the public's expectations about future policy actions are implicitly embedded in dynamic multivariate models.

building in a propensity to revert to a fixed mean level of conflict. Whether this is reasonable and what it means for medium and long term mediation remains to be determined.

Appendix: Gibbs Sampling Algorithm for Conditional Forecasts

Here we describe the algorithm for calculating conditional forecasts under hard policy counterfactuals. This parallels the discussion in Waggoner and Zha (1999), but with slightly more detail about the steps and the computations for BVAR models with the Sims-Zha prior.

Waggoner and Zha (1999) show that conditional on (6) in the text, and the parameter vector of the VAR ($a = (a_0 a_+)$), the joint conditional h-step forecast distribution is Gaussian with

$$p(y_{T+n}|a, \mathbf{Y}_{T+n-1}) = \phi\left(c + \sum_{l=1}^{p} y_{T+n-l}B_l + \mathbf{M}(\epsilon_{T+n})A_0^{-1}\,;\, A_0^{-1\prime}\mathbf{V}(\epsilon_{T+n})A_0^{-1}\right) \tag{7}$$

where $\mathbf{Y}_{T+n-1}$ is the data matrix up to $T+n-1$. $\mathbf{M}(\epsilon_{T+n})$ and $\mathbf{V}(\epsilon_{T+n})$ are the mean and variance of the constrained innovations under the conditional forecast:

$$p(\epsilon_t|a, R(a)'\epsilon = r) = \phi(R(a)(R(a)'R(a))^{-1}r(a)\,;\, I - R(a)(R(a)'R(a))^{-1}R(a)') \tag{8}$$

With these distributions, the Gibbs sampling algorithm of Waggoner and Zha (1999) becomes:

Let N_1 be the number of burn-in draws, and N_2 be the number of Gibbs samples after the burn-in. Then,

1. Initialize the values of a_0 and a_+ for the VAR. This can be done using either a BVAR or other estimator. These values should come from the peak of $p(a|\mathbf{Y_T})$
2. Generate $y_{T+1}\ldots y_{T+h}$ based on the draw of a_0 and a_+. This includes first sampling the constrained innovations sequence from (8) and using these structural innovations in the computation of the forecasts using the reduced form representation in (3), in the text. Note that at each innovation one must recompute the value of the mean of ϵ, which depends on r, which in turn depends on a, which is sampled in the Gibbs iterations.
3. Repeat the previous steps until the sequence

   $$\{a^1, y^1_{T+1}, \ldots, y^1_{T+h}, \ldots, a^{N_1+N_2}, y^{N_1+N_2}_{T+1}, \ldots, y^{N_1+N_2}_{T+h}\}$$

   is simulated.
4. Keep the last N_2 draws.

As Waggoner and Zha note, the crucial part of the computation is step (2), where we account for both the *parameter uncertainty* and the *structural shocks* which are constrained for a conditional forecast.

Most existing forecasting inference and forecasting procedures (particularly those that are non-Bayesian), ignore step (2) and therefore take the innovations as the only source of uncertainty.

References

Amisano, Gianni and Carlo Giannini. 1997. *Topics in Structural VAR Econometrics*. 2nd ed. Berlin: Springer.

Box, George E.P. and G. C. Tiao. 1975. "Intervention Analysis with Applications to Economic and Environmental Problems." *Journal of the American Statistical Association* 70:70–79.

Brandt, Patrick T. Forthcoming. "Using the Right Tools for Time Series Data Analysis." *The Political Methodologist* .

Brandt, Patrick T. and John T. Williams. 2001. "A Linear Poisson Autoregressive Model: The Poisson AR(p) Model." *Political Analysis* 9(2):164–184.

Brandt, Patrick T., John T. Williams, Benjamin O. Fordham and Brian Pollins. 2000. "Dynamic Modeling for Persistent Event Count Time Series." *American Journal of Political Science* 44(4):823–843.

Brunner, Ronald D. and Klaus Liepelt. 1972. "Data Analysis. Process Analysis and System Change." *Midwest Journal of Political Science* 17:538–569.

Crocker, Chester, Fen Osler Hampson and Pamela Aall. 1999. *Herding Cats: Multiparty Meditation in a Complex World*. Washington, D.C.: U.S. Institute of Peace Press.

Diebold, Francis X. 1998. "The Past, Present and Future of Macroeconomic Forecasting." *Journal of Economic Perspectives* 12(2):175–192.

Doan, Thomas, Robert Litterman and Christopher Sims. 1984. "Forecasting and Conditional Projectsion Using Realistic Prior Distributions." *Econometric Reviews* 3:1–100.

Downs, George and Stephen John Steadman. 2002. Evaluation Issues in Peace Implementation. In *Ending Civil Wars: The Implementation of Peace Agreements*, ed. Stephen John Steadman, Donald Rothschield and Elizabeth Cousins. Lynne Reiner Publishers.

Fearon, James. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–592.

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

Freeman, John R. 1990. Systematic Sampling. Temporal Aggregation and the Study of Political Relationships. In *Political Analysis*, ed. James Stimson. Vol. 1 Ann Arbor: University of Michigan Press pp. 61–98.

Freeman, John R. and Daniel Houser. 1998. "A Computable Equilibrium Model for the Study of Political Economy." *American Journal of Political Science* 42(2):628–660.

Gaubatz, Kurt. 1991. "Electoral Cycles and War." *Journal of Conflict Resolution* 35(2):212–244.

Gerner, Deborah J. 1994. *One Land, Two Peoples: The Conflict Over Palestine*. Boulder. Colorado: Westview Press.

Goldstein, Joshua and Jon C. Pevehouse. 1997. "Reciprocity, Bullying and International Cooperation: Time Series Analysis of the Bosnian Conflict." *American Political Science Review* 91:515–524.

Goldstein, Joshua S. 1988. *Long Cycles: Propserity and War in the Modern Age*. New Haven, CT: Yale University Press.

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

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.

Gowa, Joanne. 1998. "Parties, Voters, and the Use of Force Abroad." *International Organization* 52(3):123–154.

Granger, Clive W.J. 1999. *Empirical Modeling in Economics: Specification and Evaluation*. Cambridge: Cambridge University Press.

Greenberg, John H. Barton [and Margaret E. McGuinness], ed. 2000. *Words Over War Meditation and Arbitration to Prevent Deadly Conflict*. Boulder, Colorado: Rowman & Littlefield Publishers.

Hampson, Fen Osler. 1996. *Nurturing Peace: Why Peace Settlements Succeed or Fail*. Washington, D.C.: U.S. Institute of Peace Press.

Houser, Daniel and John Freeman. 2001. "Economic Consequences of Political Approval Management in Comparative Perspective." *Journal of Comparative Economics* 29:692–721.

Howell, Llewellyn D. 2001. *The Handbook of Country and Political Risk Analysis*. 3rd ed. Syracuse: The PRS Group.

Jackman, Simon. 2000. "Estimation and Inference Via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo." *American Journal of Political Science* 44(2):375–405.

Kadiyala, K. Rao and Sune Karlsson. 1997. "Numerical Methods For Estimation and INference in Bayesian VAR-Model." *Journal of Applied Econometrics* 12:99–132.

King, Gary and Langche Zeng. 2001*a*. "How Factual Is Your Counterfactual." Manuscript Cambridge, MA: Harvard University.

King, Gary and Langche Zeng. 2001*b*. "Improving Forecasts of State Failure." *World Politics* 53:623–658.

Kleiboer, Marieke. 1998. *The Multiple Realities of International Meditation*. Boulder, Colorado: Lynne Rienner, Publishers.

Lederach, John Paul. 1997. *Building Peace: Sustainable Reconciliation in Divided Societies*. Washington, D.C.: U.S. Institute of Peace Press.

Leeper, Eric M., Christopher A. Sims and Tao Zha. 1996. "What Does Monetary Policy Do?" *Brookings Papers on Economic Activity* 1996(2):1–63.

Litterman, Robert B. 1986. "Forecasting with Bayesian Vector Autoregressions — Five Years of Experience." *Journal of Business, Economics and Statistics* 4:25–38.

Litterman, Robert B. and Laurence Weiss. 1985. "Money, Real Interest Rates, and Output: A Reinterpretation of Postwar U.S. Data." *Econometrica* 53(1):129–156.

Lucas, Robert. 1976. "Econometric Policy Evaluation: A Critique." *Journal of Monetary Economics* 1 Supplementary Series:19–46.

Lutkepohl, Helmut. 1993. *Introduction of Multiple Time Series Analysis.* Berlin: Springer-Verlag.

Meier, Kenneth and Jeff Gill. 2001. *What Works: A New Approach to Program and Policy Analysis.* Boulder, Colorado: Westview Press.

Morris, Benny. 2001. *Righteous Victims: A History of the Zionist-Arab Conflict 1881- 2001.* New York: Vintage Books.

Pevehouse, Jon C. [and Joshua S. Goldstein]. 1999. "Serbian Compliance or Defiance in Kosovo? Statistical Analysis and Real-Time Predictions." *Journal Of Conflict Resolution* 43(4):538–546.

Prescott, Edward C. 1991. "Real Business Cycle Theory: What Have We Learned?" *Revisita de Analisis Economico* 6:3–19.

Roberston, John C. and Ellis W. Tallman. 1999. "Vector Autoregressions: Forecasting and Reality." *Economic Review (Federal Reserve Bank of Atlanta)* First Quarter:4–18.

Robertson, John C. and Ellis W. Tallman. 1998. "Data Vintages and Measuring Forecast Model Performance." *Economic Review, Atlanta: Federal Reserve Bank of Atlanta* Fourth Quarter:4–20.

Schrodt, Philip A. and Deborah 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.

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

Sims, Christopher and Tao Zha. 1999. "Error Bands for Impulse Responses." *Econometrica* 67(5):1113–1156.

Smith, Charles D. 2001. *Palestine and The Arab-Israeli Conflict: A History with Documents.* 4th ed. Boston, Massachusetts: Bedford-St. Martins.

Stock, James H. and Mark W. Watson. 2001. "Vector Autoregressions." *Journal of Economic Perspectives* 15(4):101–115.

Theil, Henri. 1963. "On the Use of Incomplete Prior Information in Regression Analysis." *Journal of the American Statistical Association* 58(302):401–414.

Theil, Henri and Arthur S. Goldberger. 1961. "On Pure and Mixed Statistical Estimation in Economics." *International Economic Review* 2(1):65–78.

Waggoner, Daniel F. and Tao Zha. 1999. "Conditional Forecasts in Dynamic Multivariate Models." *Review of Economics and Statistics* 81(4):639–651.

Williams, John T. 1993. "Dynamic Change, Specification Uncertainty, and Bayesian Vector Autoregression Analysis." *Political Analysis* pp. 97–125.

Zartman, I. William and J. Lewis Rasmussen, eds. 1997. *Peacemaking in International Conflict: Methods and Techniques.* Washington, D.C.: Institute of Peace Press.

Zha, Tao. 1998. "A Dynamic Multivariate Model for the Use of Formulating Policy." *Economic Review (Federal Reserve Bank of Atlanta)* First Quarter:16–29.

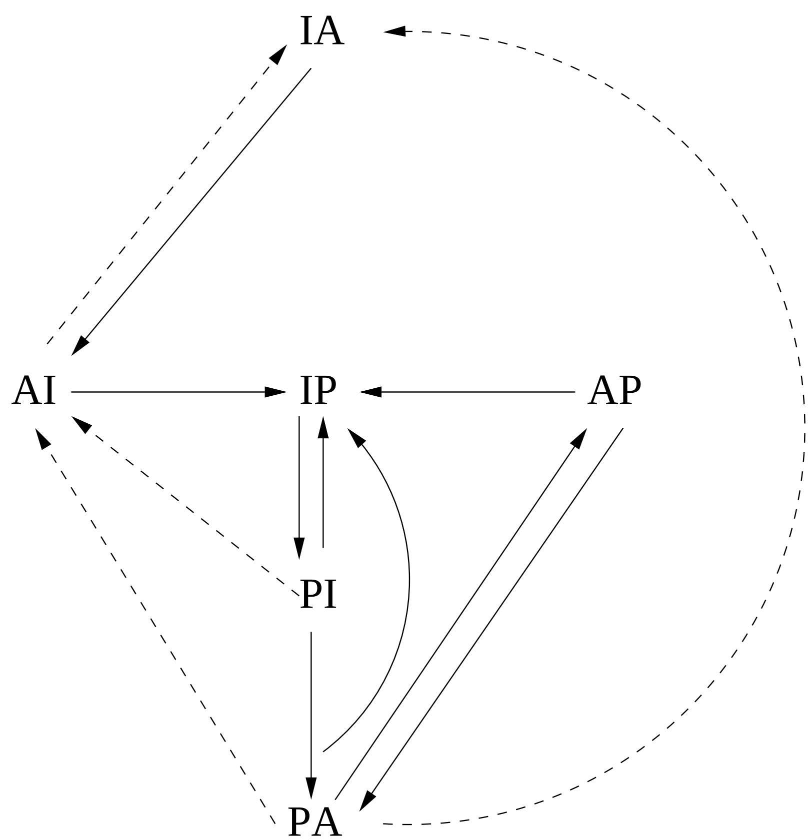

Figure 1: Levant VAR Causal Nexus. Causal relationships found in VAR(6) model for daily data from April 15, 1979 to December 15, 1988

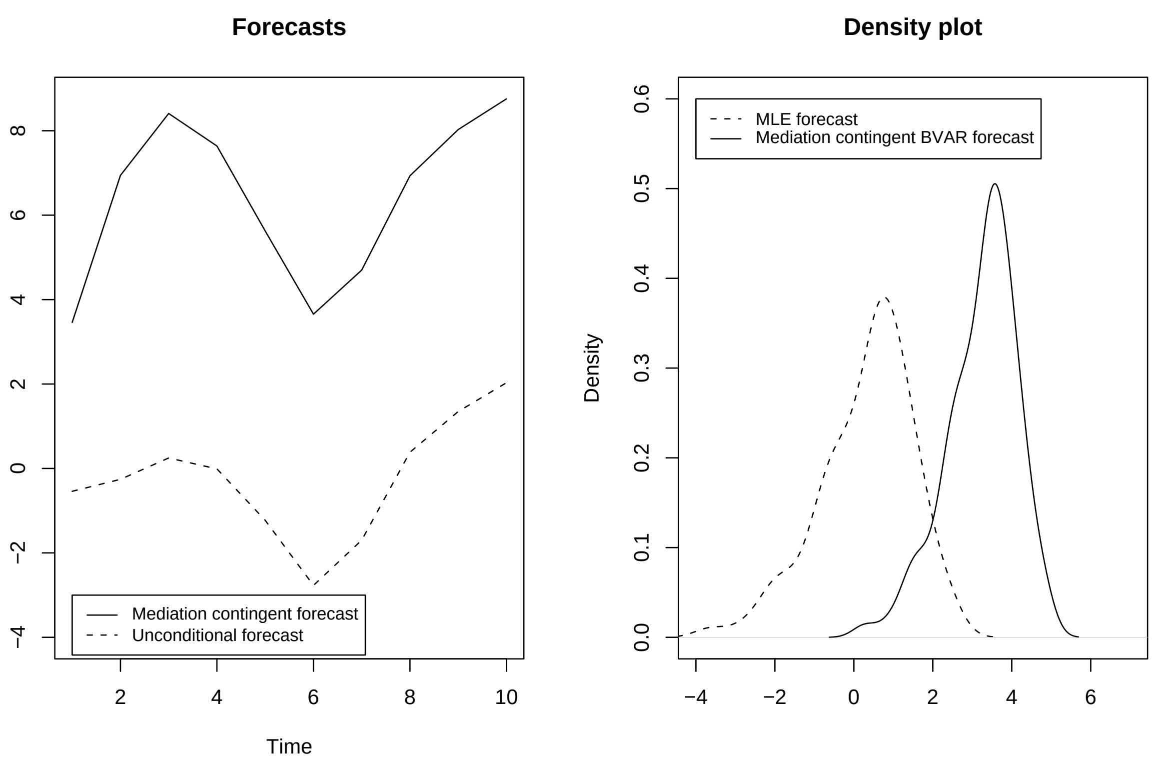

Figure 2: Expected BVAR results

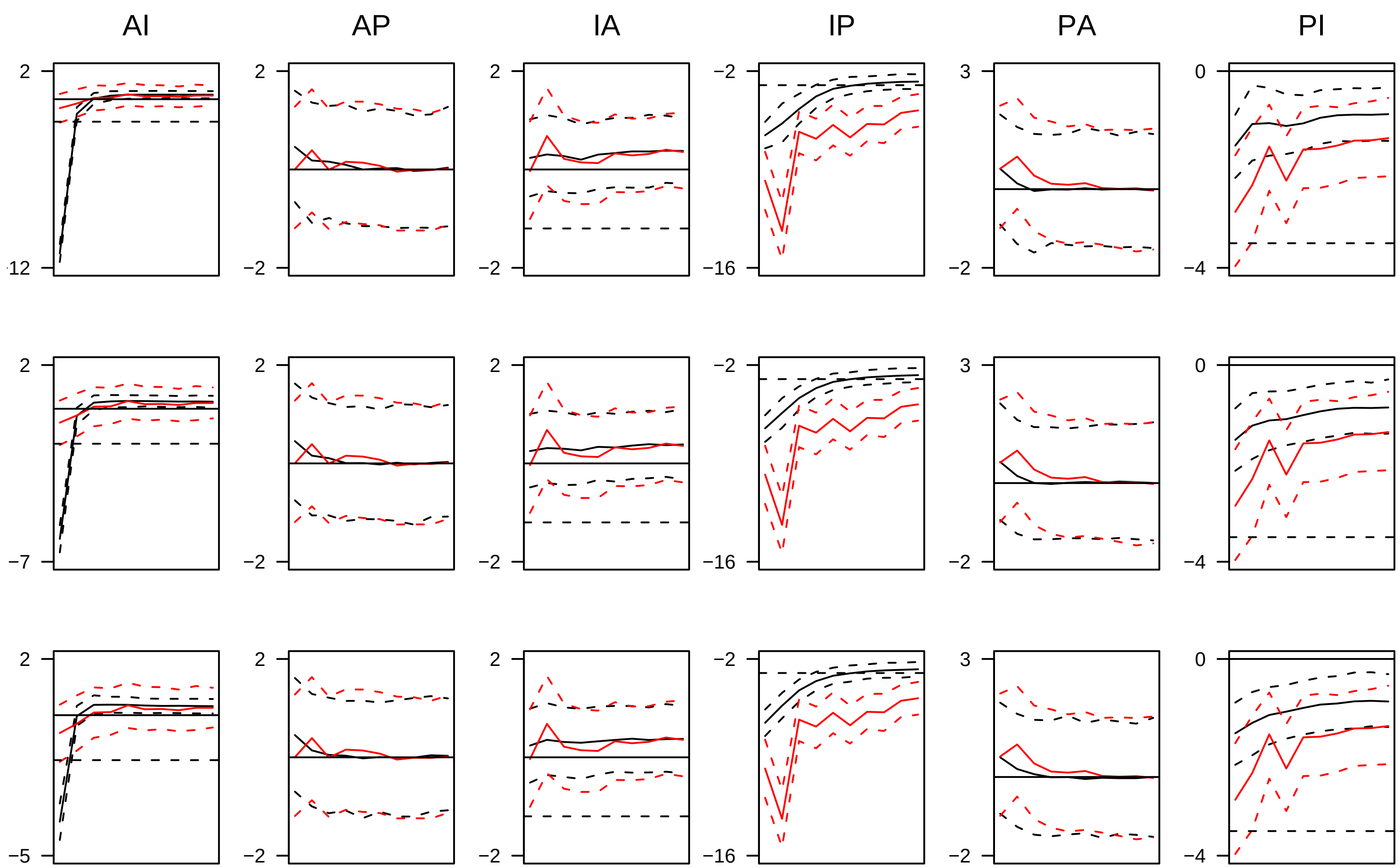

Figure 3: Sims-Zha Bayesian VAR Conditional Forecasts with 0.90 Probability Bands

Row 1: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = 10$

Row 2: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = 5$

Row 3: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = \mu + \sigma$

Forecasts in red are based on a flat prior.

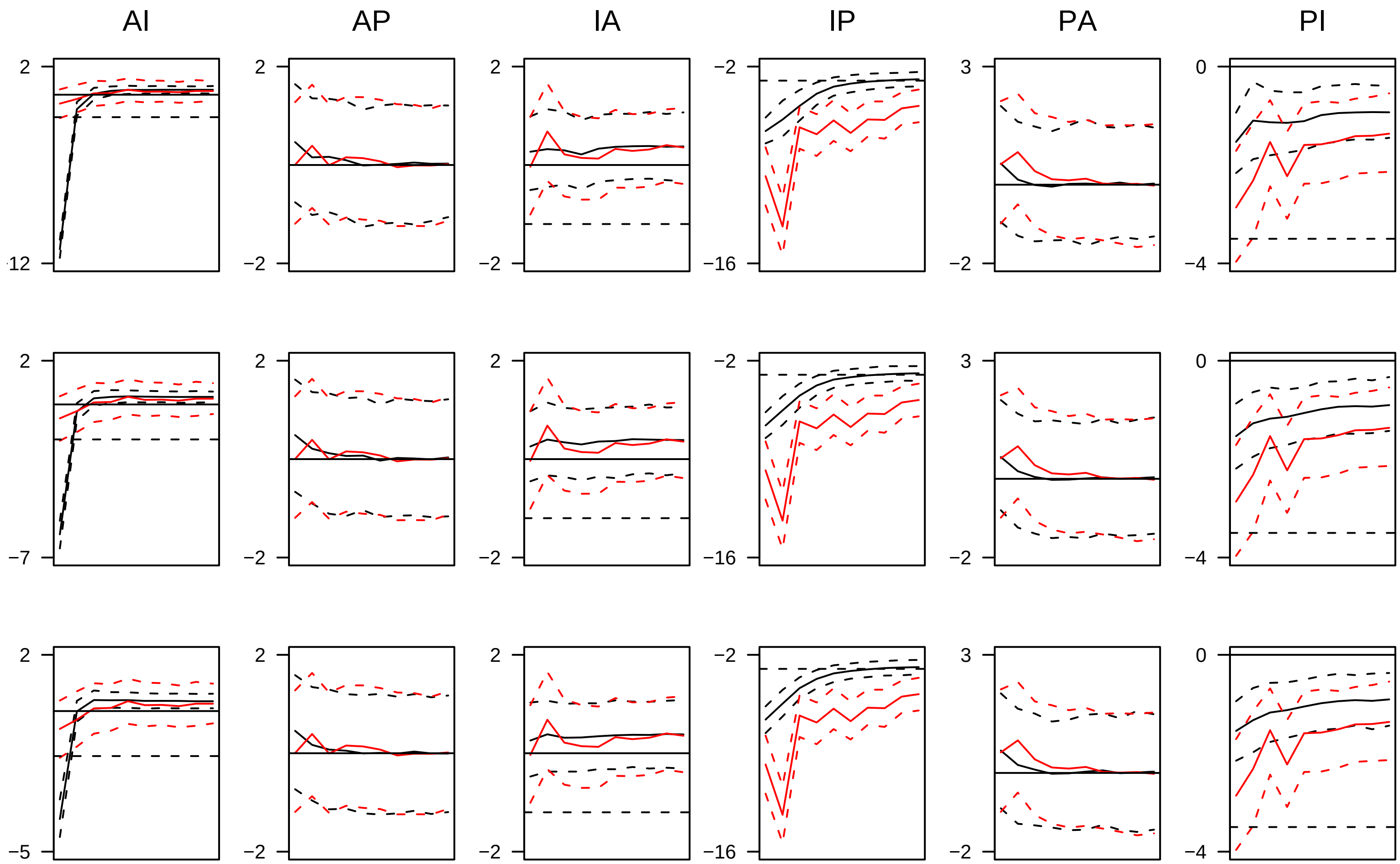

Figure 4: Sims-Zha Bayesian VAR Conditional Forecasts with 0.90 Probability Bands
Row 1: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.25$ $p = 6$ $AI = 10$
Row 2: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.25$ $p = 6$ $AI = 5$
Row 3: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.25$ $p = 6$ $AI = \mu + \sigma$
Forecasts in red are based on a flat prior.

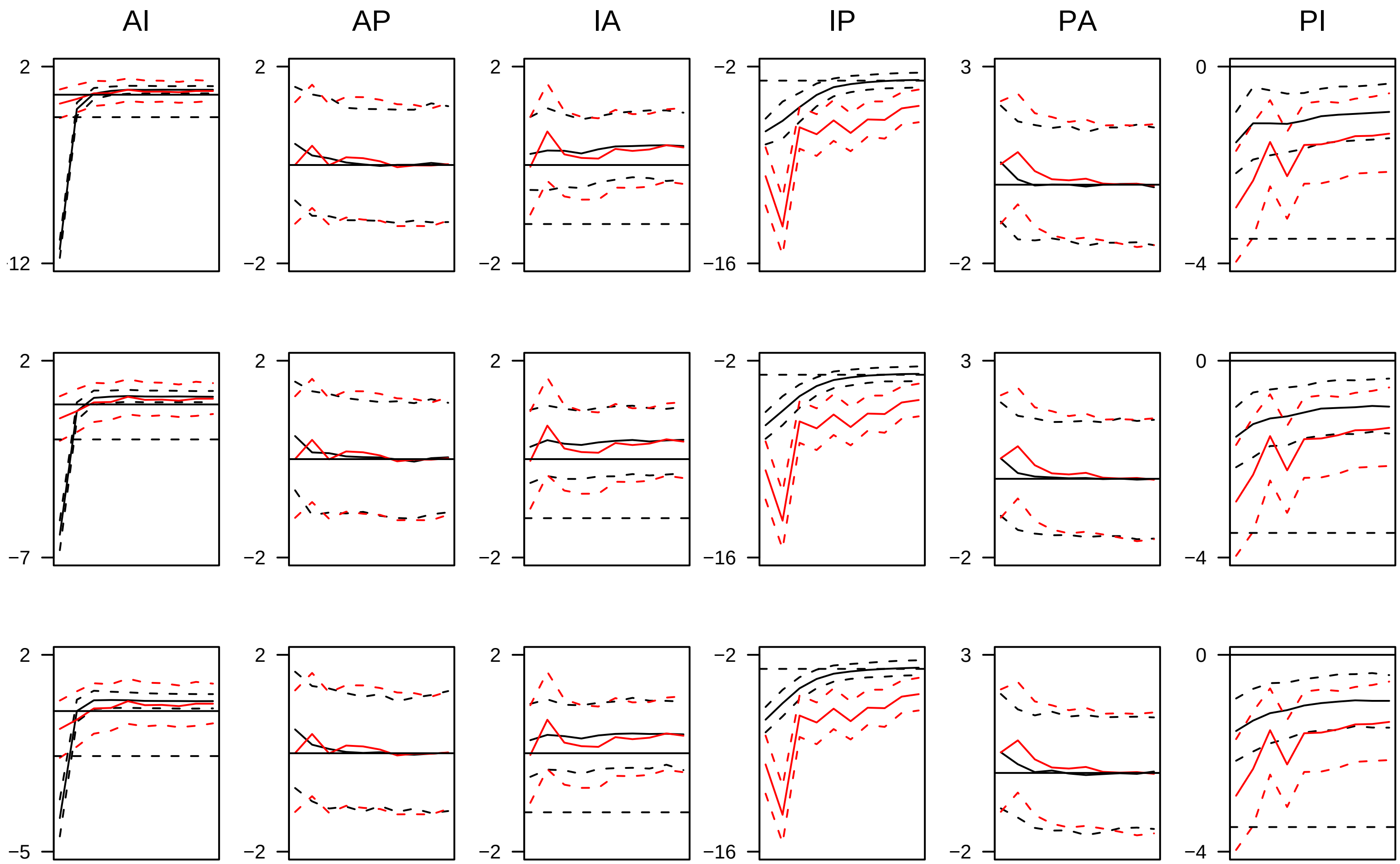

Figure 5: Sims-Zha Bayesian VAR Conditional Forecasts with 0.90 Probability Bands
Row 1: $\lambda_0 = 0.6 \quad \lambda_1 = 0.1 \quad \lambda_3 = 2 \quad \lambda_4 = 0.5 \quad p = 6 \quad AI = 10$
Row 2: $\lambda_0 = 0.6 \quad \lambda_1 = 0.1 \quad \lambda_3 = 2 \quad \lambda_4 = 0.5 \quad p = 6 \quad AI = 5$
Row 3: $\lambda_0 = 0.6 \quad \lambda_1 = 0.1 \quad \lambda_3 = 2 \quad \lambda_4 = 0.5 \quad p = 6 \quad AI = \mu + \sigma$
Forecasts in red are based on a flat prior.

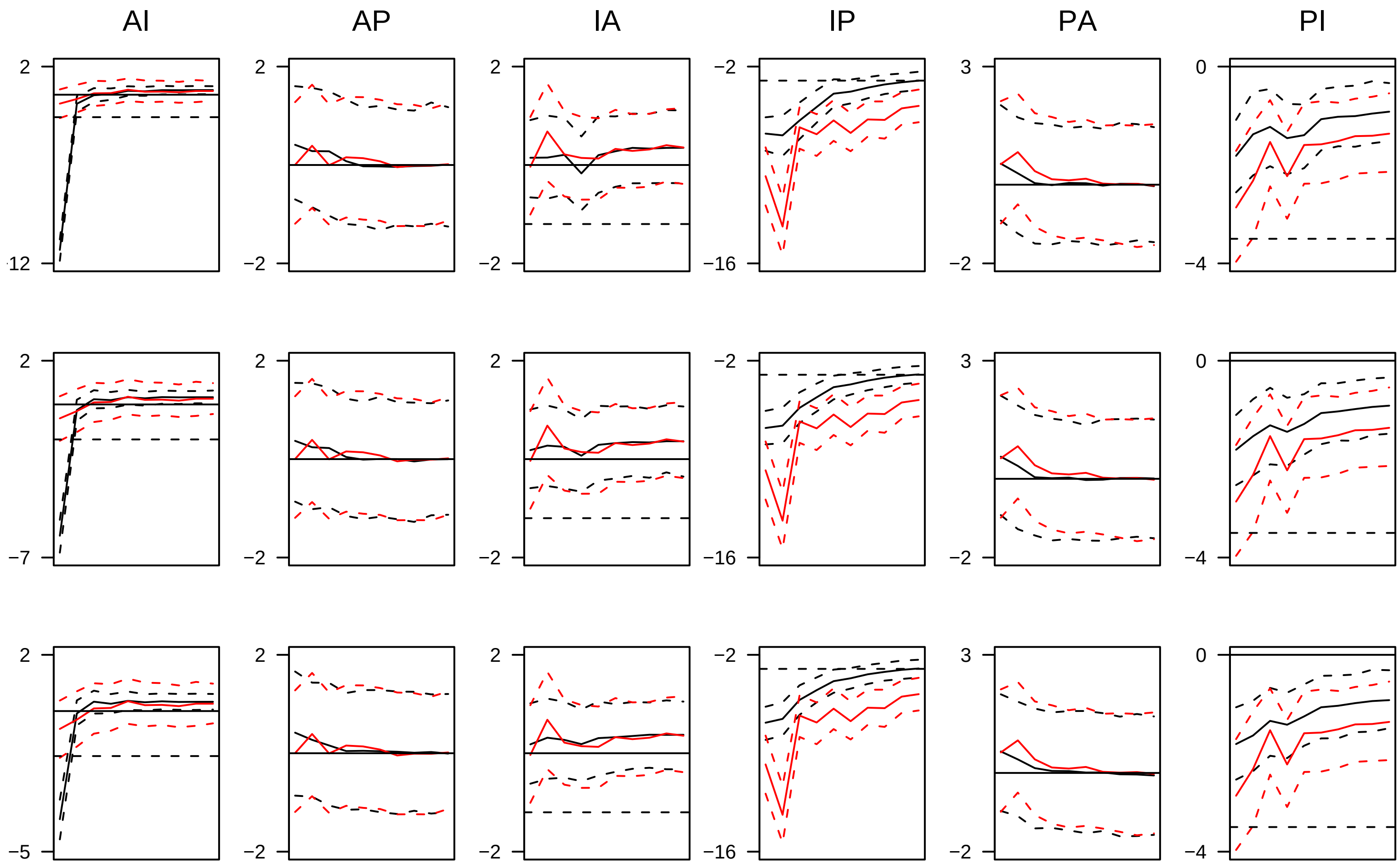

Figure 6: Sims-Zha Bayesian VAR Conditional Forecasts with 0.90 Probability Bands
Row 1: $\lambda_0 = 0.6$ $\lambda_1 = 0.25$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = 10$
Row 2: $\lambda_0 = 0.6$ $\lambda_1 = 0.25$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = 5$
Row 3: $\lambda_0 = 0.6$ $\lambda_1 = 0.25$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = \mu + \sigma$
Forecasts in red are based on a flat prior.

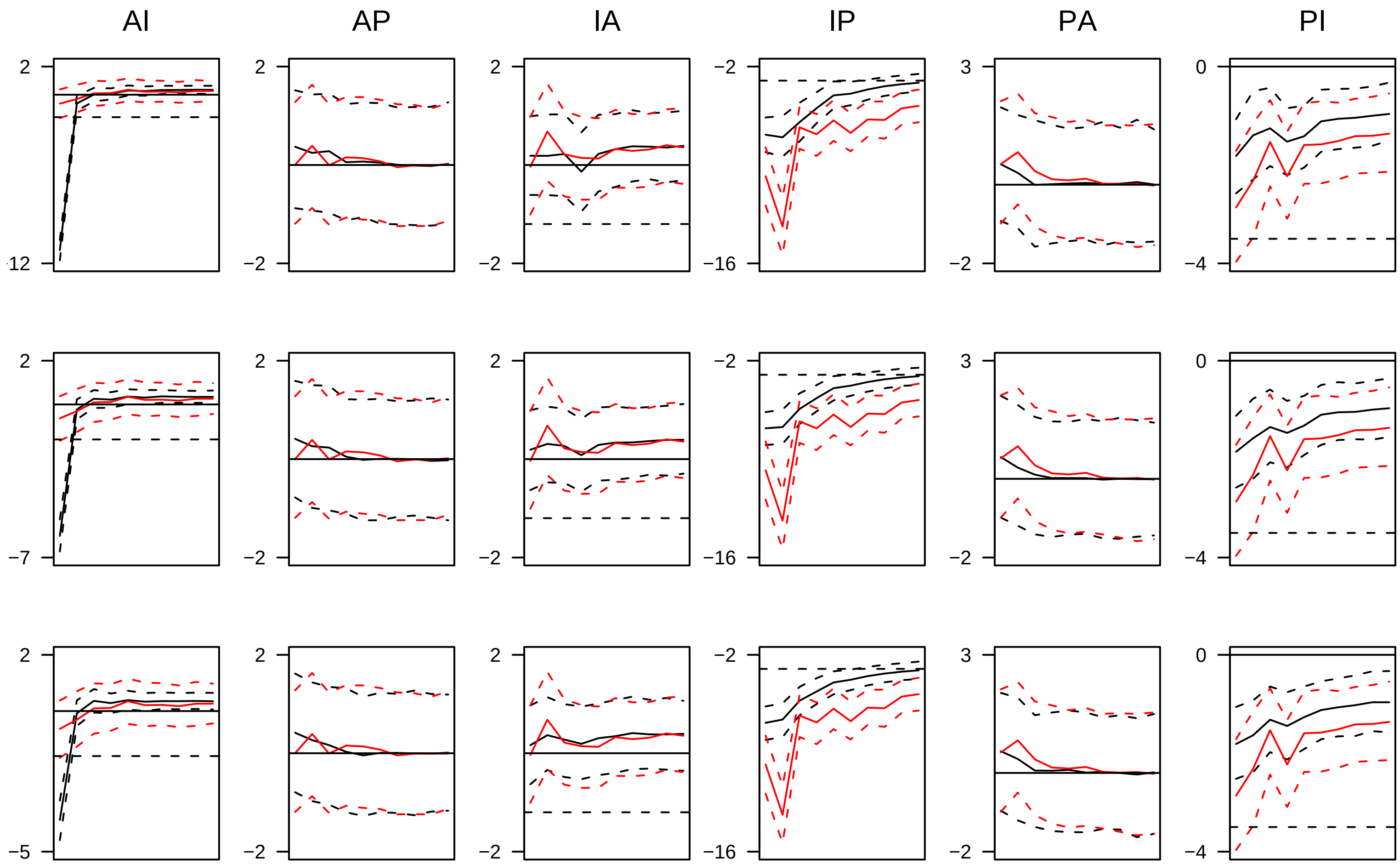

Figure 7: Sims-Zha Bayesian VAR Conditional Forecasts with 0.90 Probability Bands

Row 1: $\lambda_0 = 0.6 \quad \lambda_1 = 0.25 \quad \lambda_3 = 2 \quad \lambda_4 = 0.25 \quad p = 6 \quad AI = 10$

Row 2: $\lambda_0 = 0.6 \quad \lambda_1 = 0.25 \quad \lambda_3 = 2 \quad \lambda_4 = 0.25 \quad p = 6 \quad AI = 5$

Row 3: $\lambda_0 = 0.6 \quad \lambda_1 = 0.25 \quad \lambda_3 = 2 \quad \lambda_4 = 0.25 \quad p = 6 \quad AI = \mu + \sigma$

Forecasts in red are based on a flat prior.

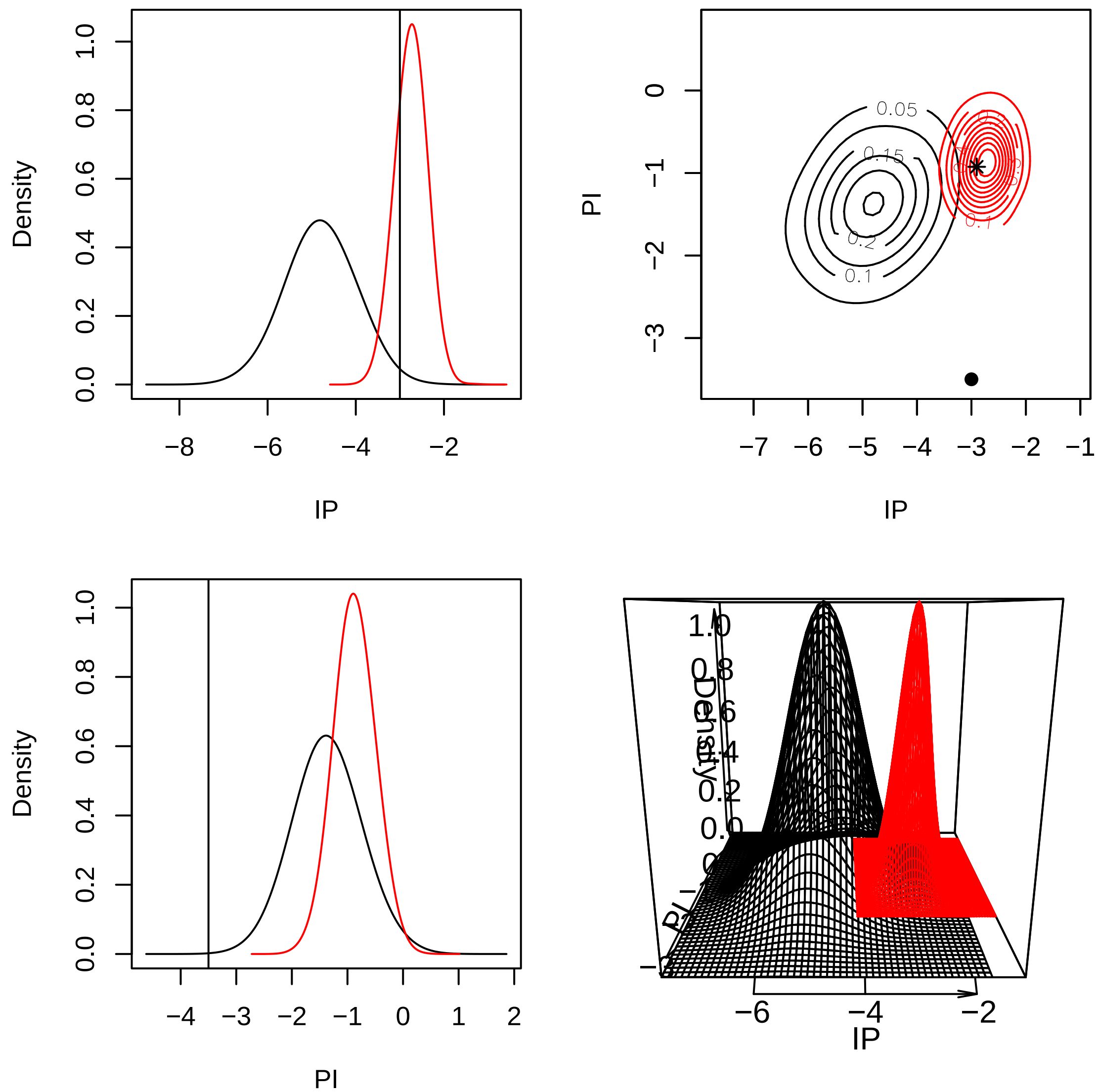

Figure 8: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988
Red: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = 10$
Black: Flat Prior Unconditional

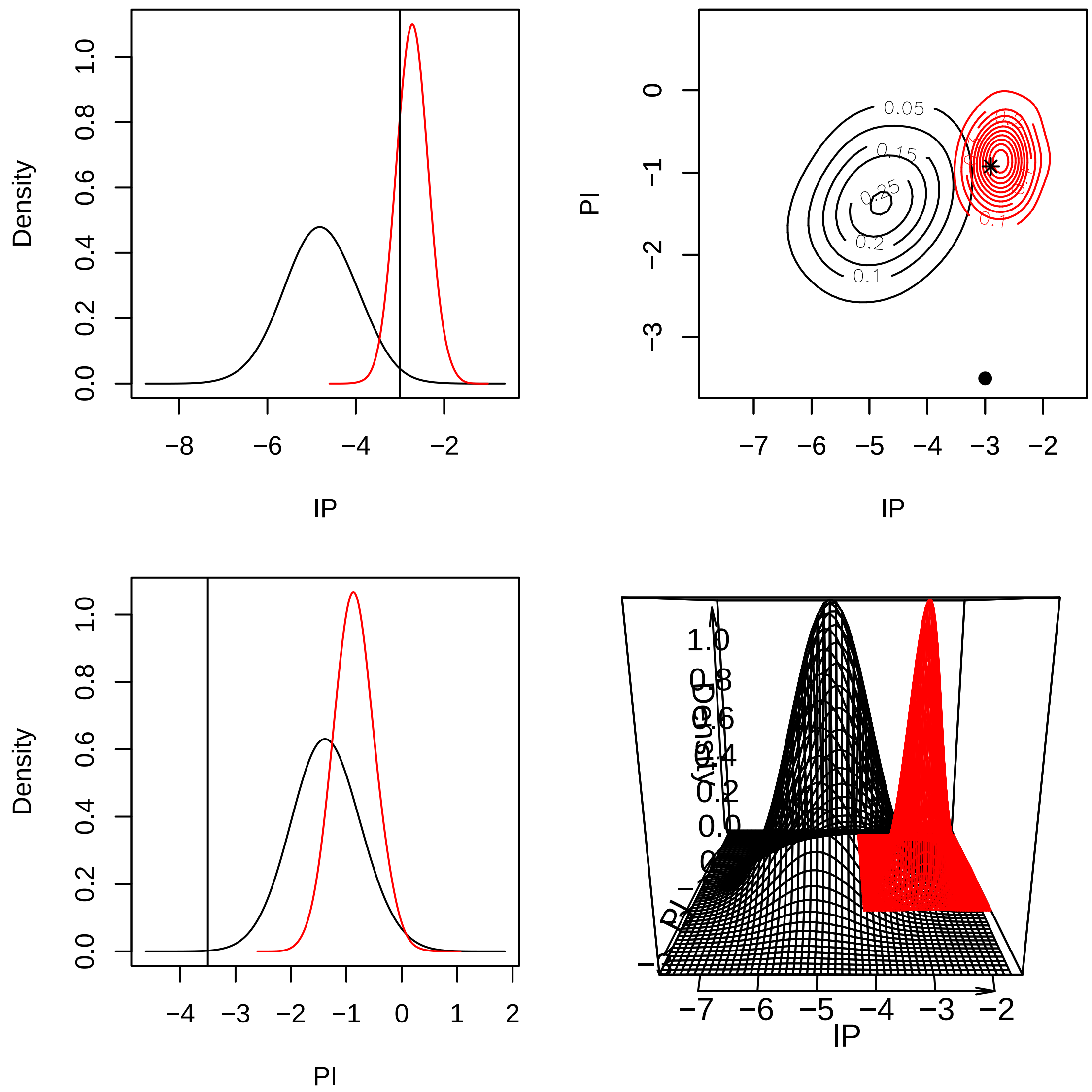

Figure 9: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988

Red: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = 5$

Black: Flat Prior Unconditional

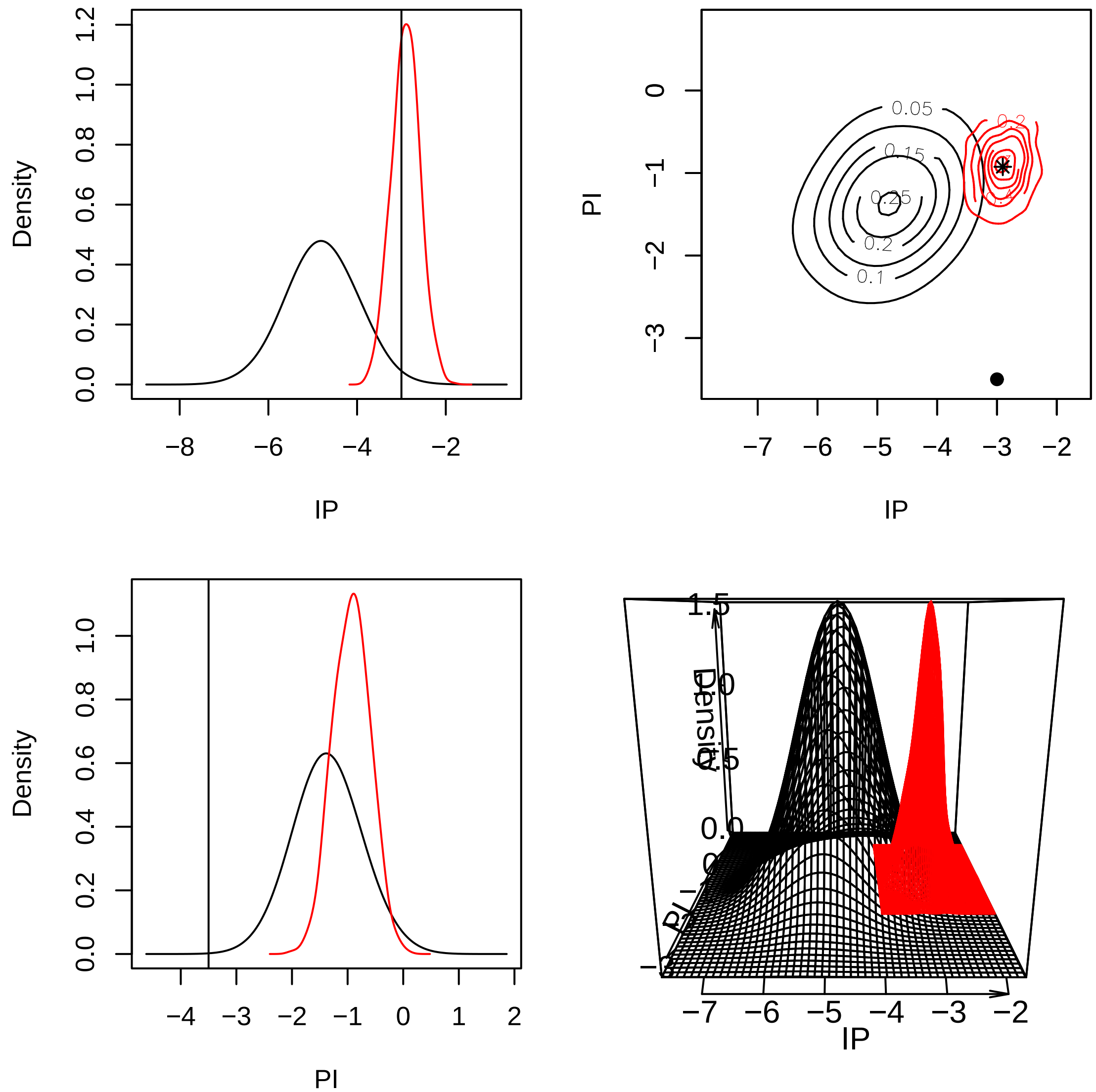

Figure 10: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988

Red: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.25$ $p = 6$ $AI = 10$

Black: Flat Prior Unconditional

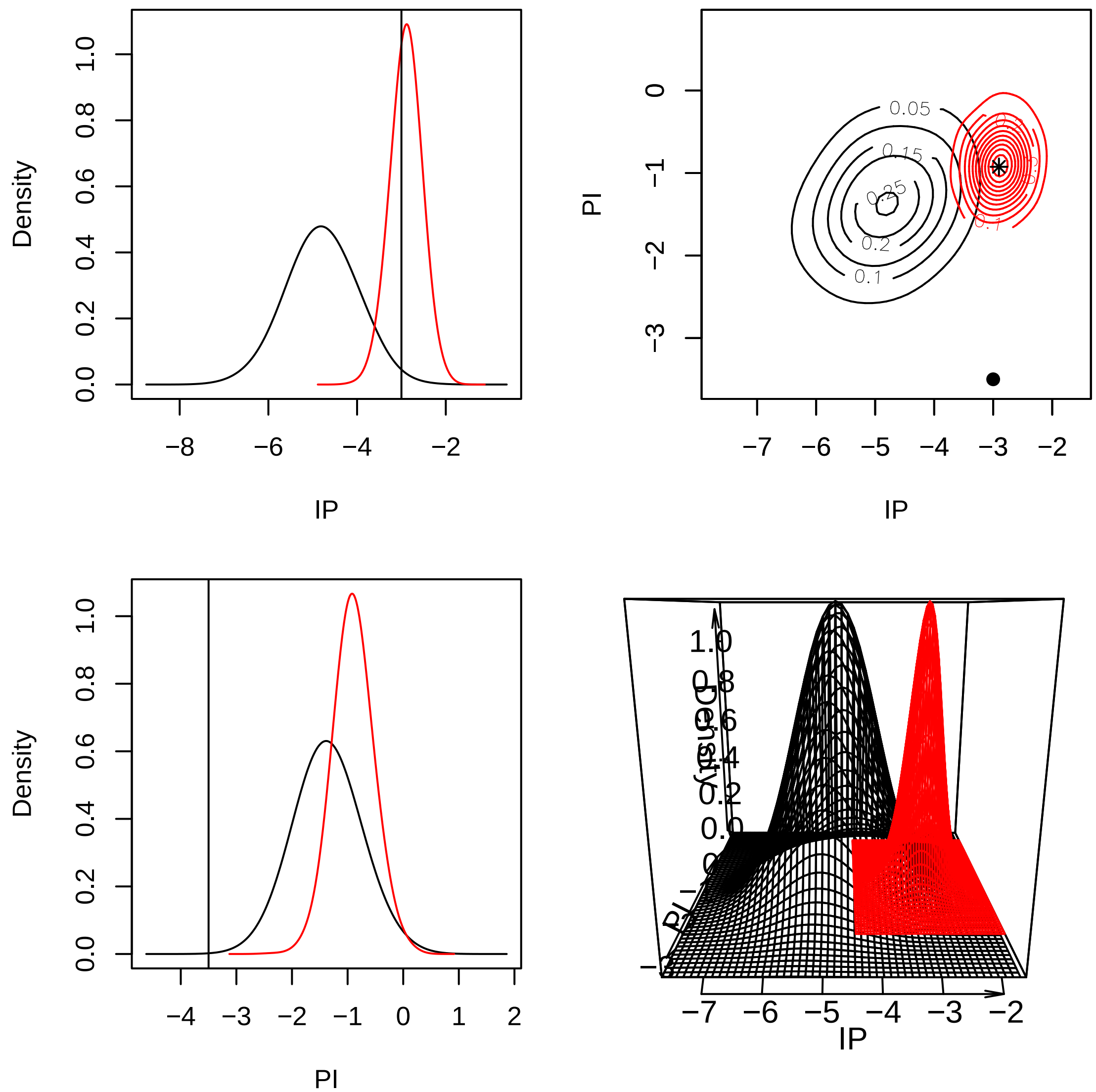

Figure 11: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988
Red: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.25$ $p = 6$ $AI = 5$
Black: Flat Prior Unconditional

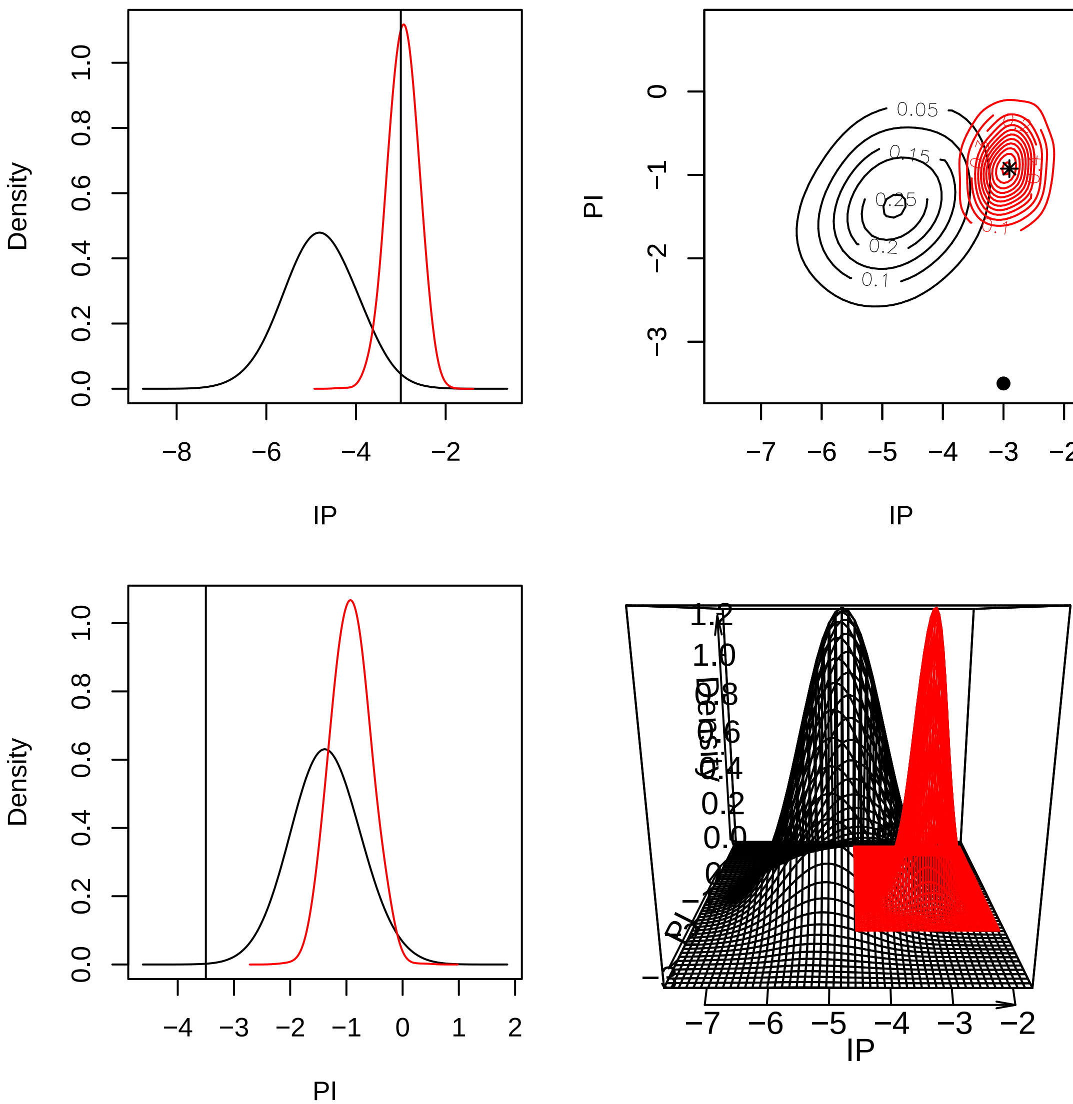

Figure 12: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988

Red:	$\lambda_0 = 0.6$	$\lambda_1 = 0.1$	$\lambda_3 = 2$	$\lambda_4 = 0.5$	$p = 6$	$AI = 10$
Black:	Flat Prior					Unconditional

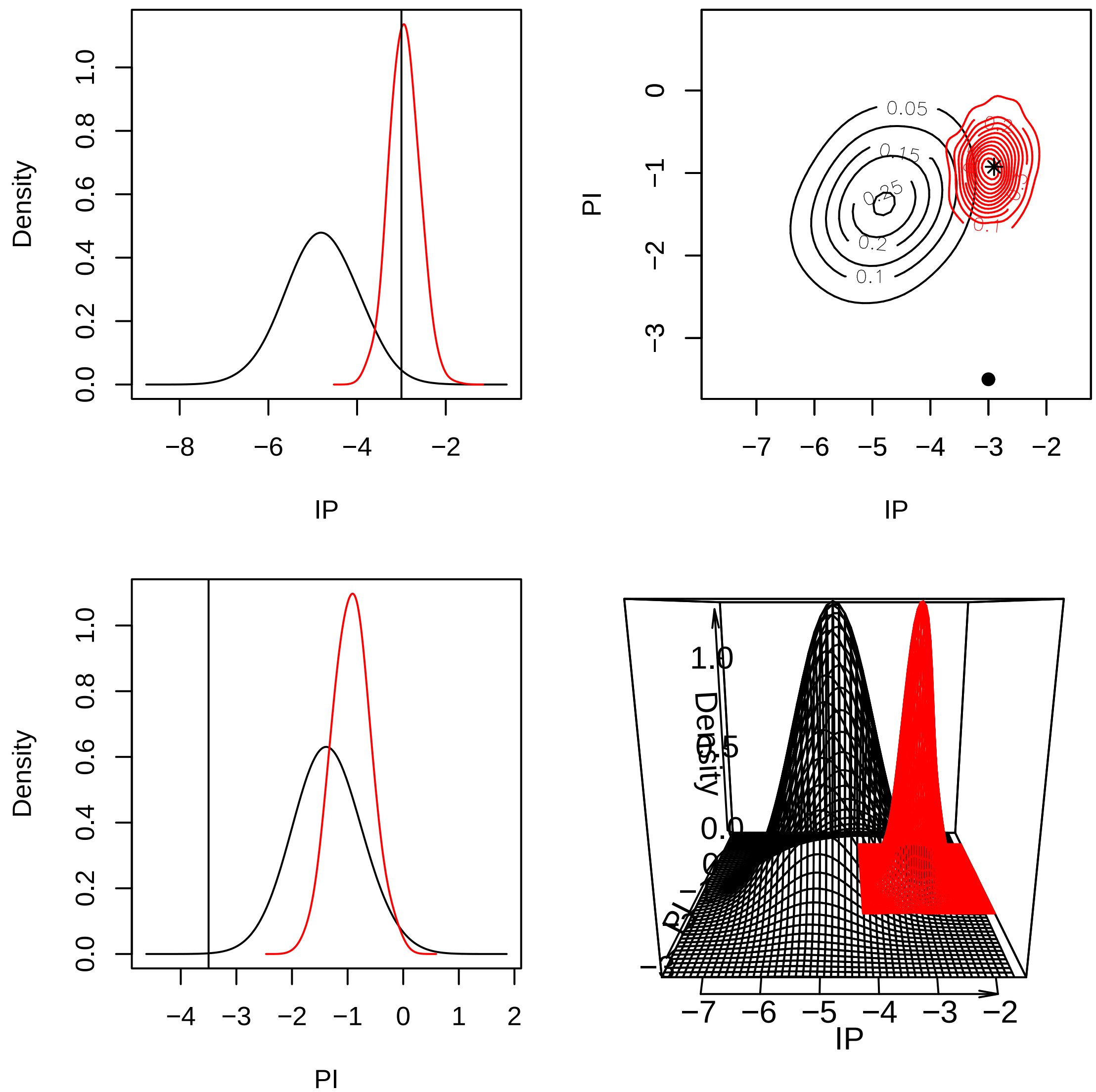

Figure 13: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988

Red: $\lambda_0 = 0.6$ $\lambda_1 = 0.1$ $\lambda_3 = 2$ $\lambda_4 = 0.5$ $p = 6$ $AI = 5$

Black: Flat Prior Unconditional

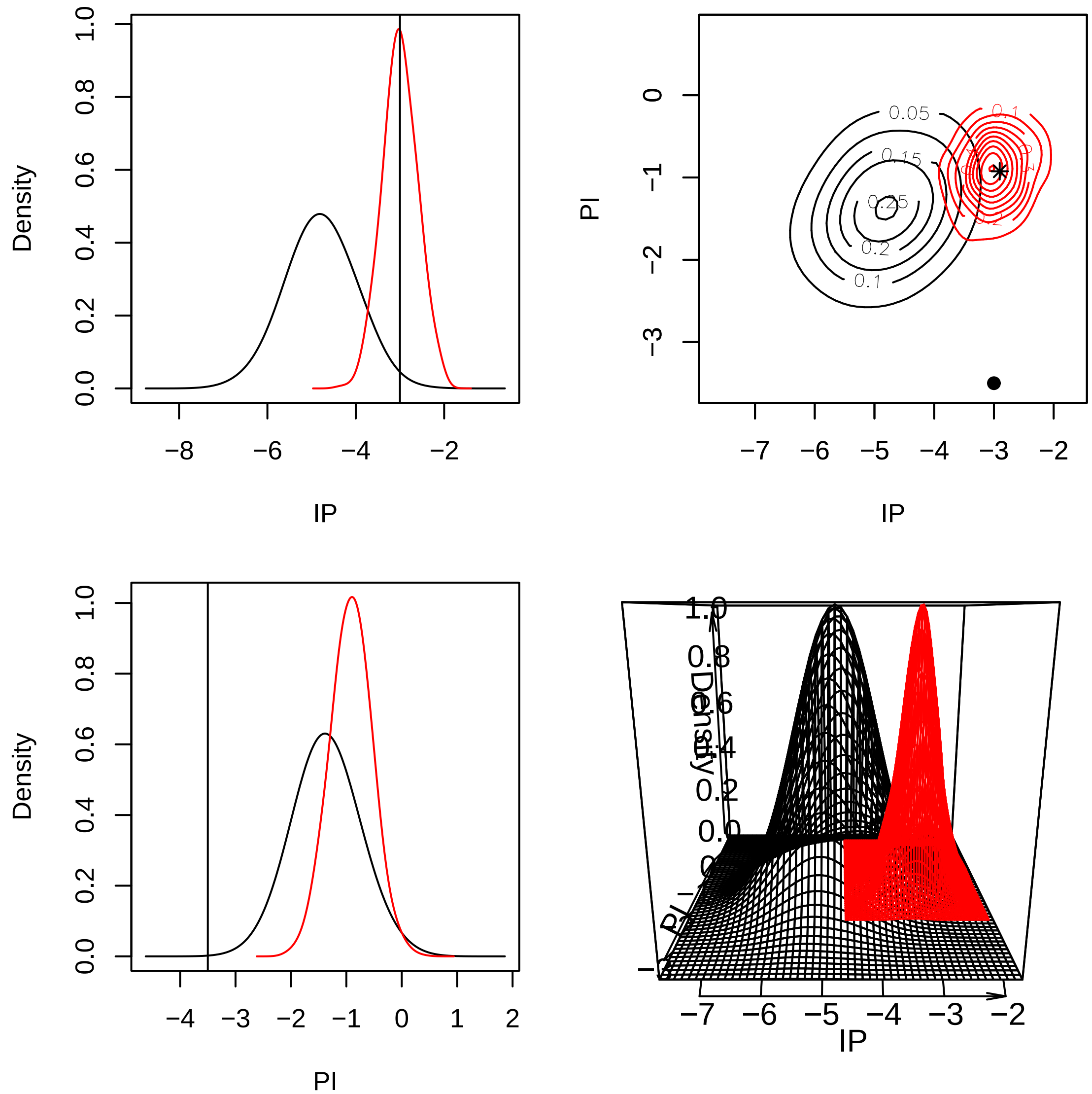

Figure 14: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988

Red:	$\lambda_0 = 0.6$	$\lambda_1 = 0.25$	$\lambda_3 = 2$	$\lambda_4 = 0.1$	$p = 6$	$AI = 10$
Black:	Flat Prior					Unconditional

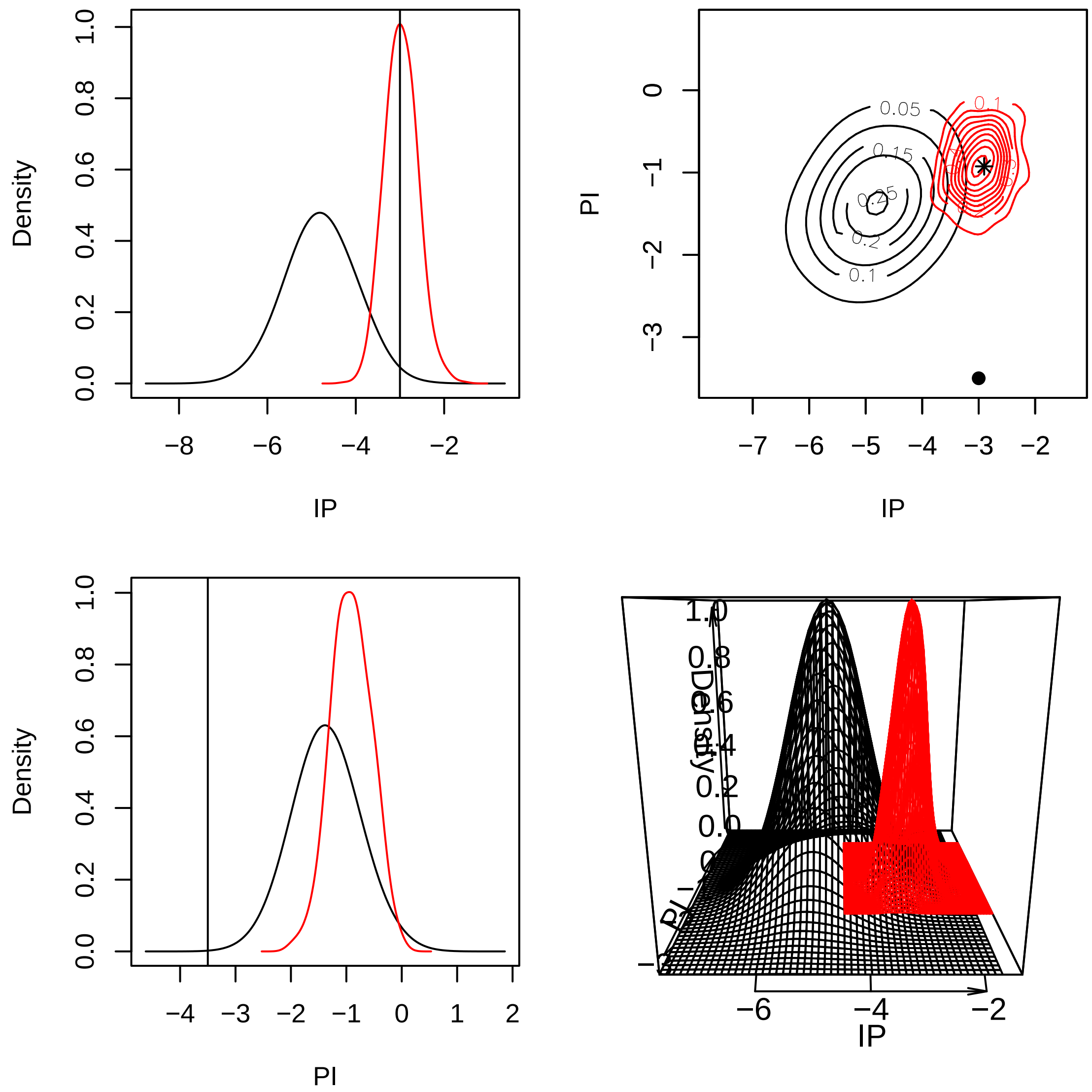

Figure 15: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988

Red: $\lambda_0 = 0.6$ $\lambda_1 = 0.25$ $\lambda_3 = 2$ $\lambda_4 = 0.1$ $p = 6$ $AI = 5$

Black: Flat Prior Unconditional

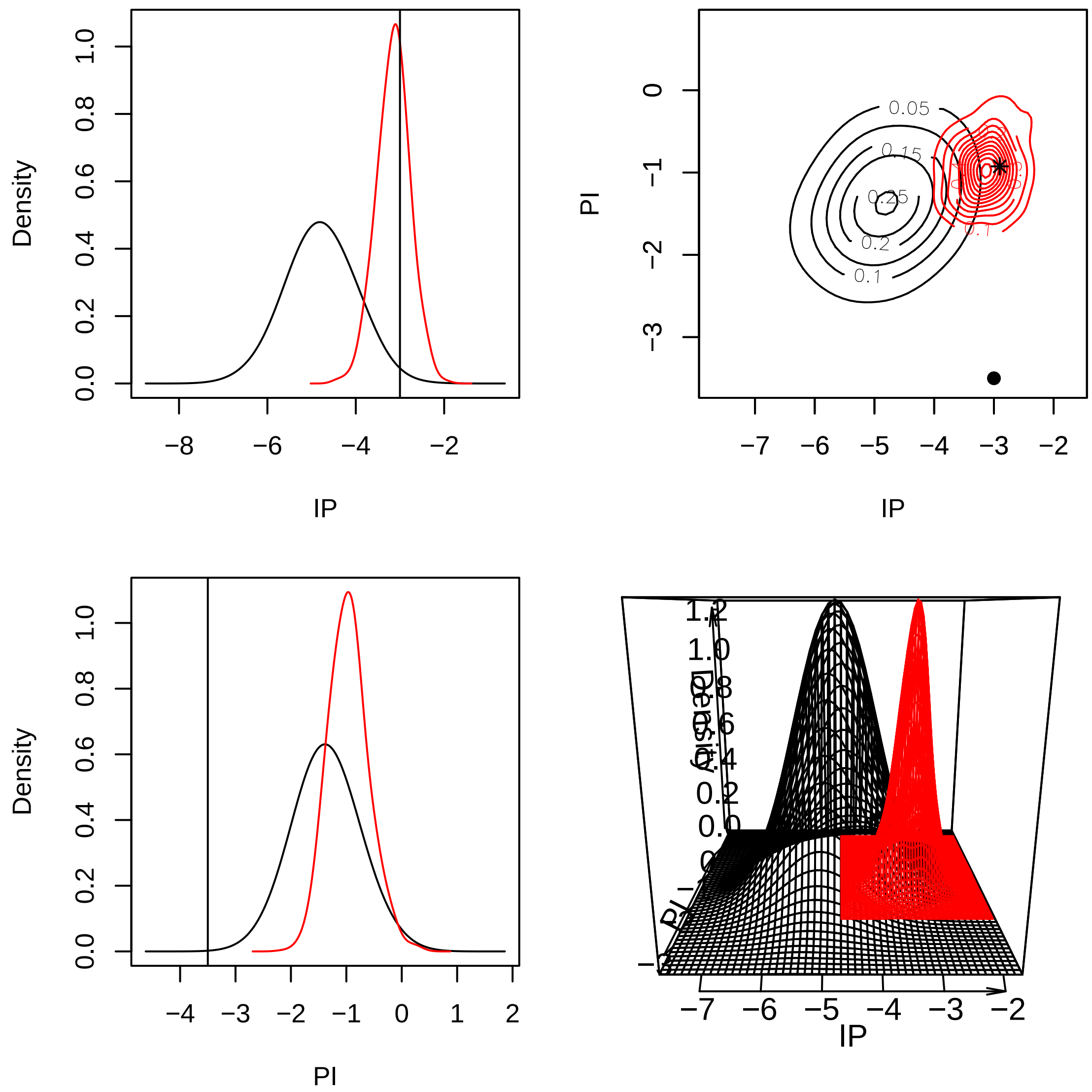

Figure 16: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988
Red: $\lambda_0 = 0.6$ $\lambda_1 = 0.25$ $\lambda_3 = 2$ $\lambda_4 = 0.25$ $p = 6$ $AI = 10$
Black: Flat Prior Unconditional

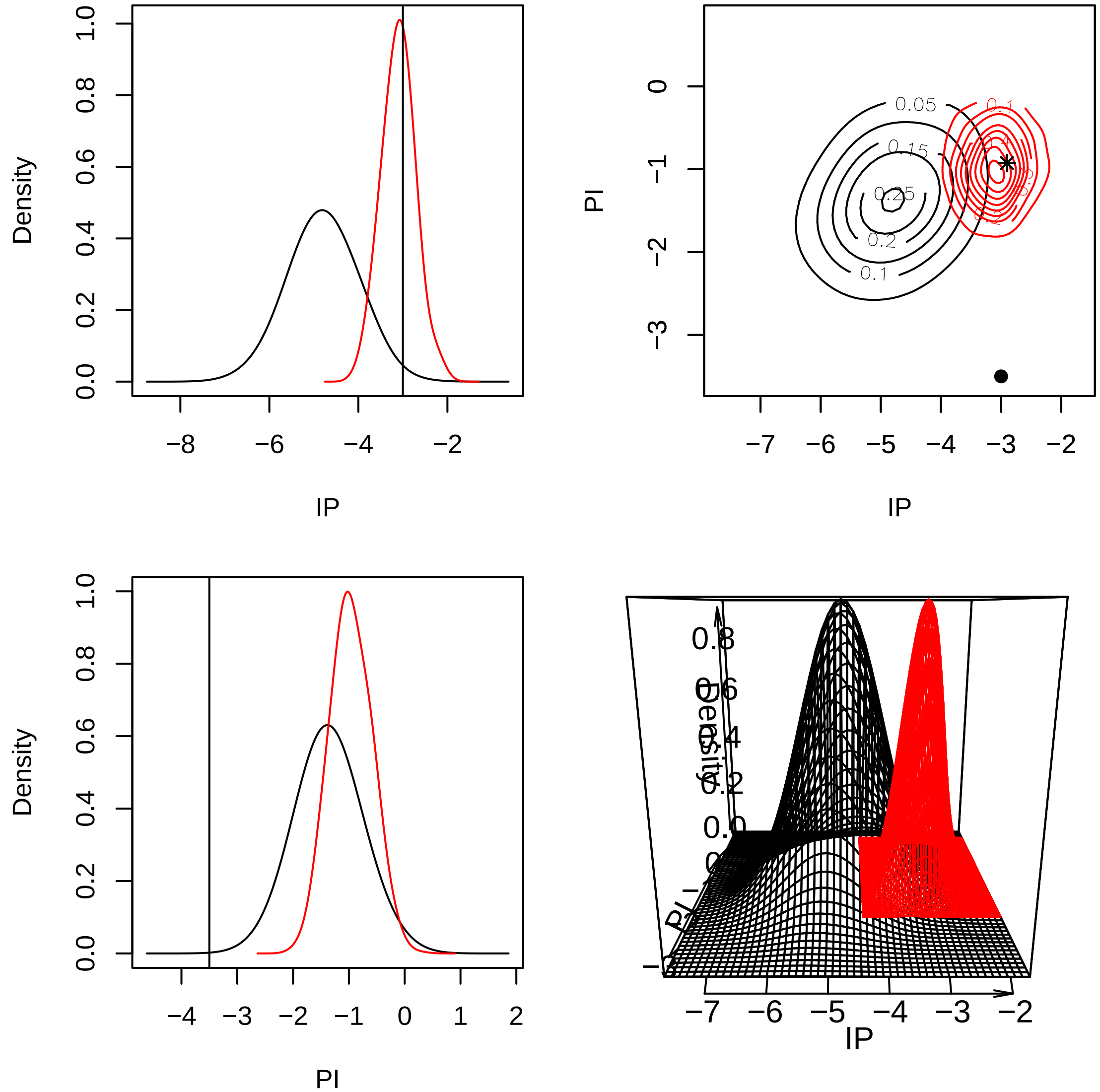

Figure 17: Forecast Density Comparisons for Israel-Palestine and Palestine-Israel, December 25, 1988
Red: $\lambda_0 = 0.6$ $\lambda_1 = 0.25$ $\lambda_3 = 2$ $\lambda_4 = 0.25$ $p = 6$ $AI = 5$
Black: Flat Prior Unconditional