Web Appendix for "The Dynamics of Reciprocity, Accountability, and Credibility" *

Patrick T. Brandt
School of Economic, Political and Policy Sciences
University of Texas at Dallas
E-mail: pbrandt@utdallas.edu
Phone: 972-883-4923

Michael Colaresi
Department of Political Science
Michigan State University
E-mail: colaresi@msu.edu
Phone: 517-775-4745

John R. Freeman
Department of Political Science
University of Minnesota
E-mail: freeman@polisci.umn.edu
Phone: 612-624-6018

December 20, 2007

*Freeman and Brandt's research is supported by the National Science Foundation under grant numbers SES-0351179, SES-0351205, SES-0540816.

Web Appendix

Model

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

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

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

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

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

The prior for the A_0 and A_+ parameters is specified for (column major) vectorized $a_0 = vec(A_0)$ and $a_+ = vec(A_+)$ where A_+ is a column major stacking of the parameters A_ℓ, $\ell = 1, \ldots, p$:

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

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

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

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

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

Impulse responses and forecasts

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

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

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

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

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

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

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

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

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

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

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

4. Repeat steps 1–3 N times.

The N posterior forecasts are then used to compute the error bands for the forecasts. A posterior sample of 100000 draws is used for computing the forecasts.

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

Model comparisons / Posterior probabilities of A_0

To compute the probabilities of each of the models, one first must find the log marginal data density. This is the probability that the sample data was produced by the model. The log marginal

Model	$log(P(A_0, A_+))$	$log(p(Y\|A_0, A_+))$	$log(\pi(A_+\|A_0))$	$\sum log(\pi(A_0(i)))$	$log(m(y))$
Recursive	−207.83	−64.37	869.51	−22.29	−1119.41
Accountability	−182.25	−71.00	869.51	−2.49	−1120.27
Bystander	−180.96	−71.65	869.51	−2.29	−1119.83
Credibility	−179.62	−70.72	869.51	−0.56	−1119.28
Follower	−181.93	−71.60	869.51	−3.84	−1119.20

Table 1: Log probabilities of various quantities of interest for the B-SVAR models reported in the paper.

data densities are computed using the following representation of the Bayesian Marginal Identity (BMI) from Chib (1995):

$$log(m(Y)) = log(p(A_0, A_+)) + log(p(Y|A_0, A_+)) - \sum_{i=1}^{m} log(\pi(A_0(i))) - log(\pi(A_+|A_0)), \quad (6)$$

where $log(m(Y))$ is the log marginal data density, $log(p(A_+, A_0))$ is the log prior, $log(p(Y|A_+, A_0))$ is the log-likelihood, $\sum_{i=1}^{m} log(\pi(A_0(i)))$ is the log posterior probability of the A_0 matrix and $log(\pi(A_+|A_0))$ is the log posterior of the A_+ terms, conditional on A_0. All of the quantities on the right-hand side of the equation are trivially computed from the Gibbs sampler output.

Note that equation 6 is just the logarithm of the following:

$$m(y) = \frac{p(Y|A_0, A_+) \cdot p(A_0, A_+)}{\left[\prod_{i=1}^{m} \pi(A_0(i))\right] \pi(A_+|A_0)}. \quad (7)$$

A little rearranging of this gets us back to the standard posterior or Bayes Theorem form (so the above has just reverse engineered the standard presentation of the log MDD derivation):

$$\left[\prod_{i=1}^{m} \pi(A_0(i))\right] \pi(A_+|A_0) = \frac{p(Y|A_0, A_+) \cdot p(A_0, A_+)}{m(y)}. \quad (8)$$

Thus, the quantities produced in computing the log marginal data density can be used to compute the log probability of A_0 for each model.

Table 1 shows the log marginal data density and its component quantities computed from the Gibbs sampler output for each of the models. The items in the third column match those in the paper.

One question here is whether the log marginal data density is the quantity of interest for the model. Gelman et al. (2004) argue that in many cases the Bayes factors are *the* quantity of interest, rather than we care about some marginal probability statement. So in the present context we prefer a probability statement about the structure of the model, not the fit of the model to the data. In this case, the quantity of interest should be the model with the largest posterior probability, $\pi(A_0)$ or $\pi(A_0, A_+|Y)$? The rational is as follows: suppose we estimate two different models that differ only in their A_0 specifications. Then the likelihoods, and the estimates of other quantities above are going to be basically the same. So we should then focus on the posterior probability of A_0 as the point of comparison, since this is really the only specification difference across the models. Note that under this criteria, the $log(\pi(A_0))$ is largest for the credibility model.

References

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

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

Gelman, Andrew, John B. Carlin, Hal S. Stern and Donald B. Rubin. 2004. *Bayesian Data Analysis*. Second ed. Boca Raton: Chapman & Hall / CRC.

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