Bayesian vector autoregressive models

Nikolay Balov

StataCorp

June, 25th, 2021
Abstract

Vector autoregressive (VAR) models are popular choice for studying the joint dynamics of multiple time series. They require no special structure since the outcome variables are regressed on their own lagged variables. One of the main problems with VAR models is the significant number of regression parameters, which is proportional to the number of lags. As a result, fitted to small data, complex VAR models tend to show poor forecasting performance.

In Stata 17, we introduce a new command, bayes:var, for fitting Bayesian VAR models. Bayesian VAR models apply priors on the regression parameters and variance-covariance of the errors for a fine control over the posterior time-series process. By default, the prior on regression coefficients shrinks them towards a random-walk process that assumes no relationship between time-series variables. This assumption helps avoiding overfitting the data. The Bayesian approach also provides a systematic and unambiguous way of determining the number of lags. We illustrate Bayesian VAR models on some real data and show model interpretations based on their impulse-response functions. We also compute Bayesian forecasts and compare them to classical forecasts.
Outline

- What is VAR and when should be used.
- Data example - CO2 emissions and economic growth.
- Bayesian VAR - from Prior to Posterior.
- Introducing the `bayes: var` command.
- Choosing the number of lags - the Bayesian way.
- Checking VAR model stability with `bayesvarstable` command.
- Impulse-response functions with `bayesirf` commands.
- Forecasting with `bayesfcast` commands.
- How to deal with VAR unstability.
What is VAR?

- VAR models the joint dynamics of multiple time-series.
- Consider a vector of $K$ time-series variables
  \[ y = (y_1, y_2, \ldots, y_K)' \]
  observed at time $t = 1 \ldots T$. Variables $y_1, \ldots, y_K$ are called endogenous.
- VAR model is just a multivariate normal regression of $y_t$ on its own-lags
  \[ y_t \sim y_{t-1}, \ldots, y_{t-p} \]
- It is possible to add exogenous variables $x_1, \ldots, x_m$ to the model
  \[ y_t \sim y_{t-1}, \ldots, y_{t-p}, x_1, \ldots, x_m \]

- Advantage: simplicity - no need for special structure of covariates.
- Disadvantage: large number of regression coefficients, $K(Kp + m)$; difficult interpretation.
CO2 emissions and economy growth

- We use a dataset from *Environmental Econometrics Using Stata, by C. Baum and S. Hurn, Ch. 5. [Grossman and Krueger(1993)]*

- `. use greensolow`
  - `. tsset datevec, quarterly`
  - `. describe`

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Storage type</th>
<th>Display format</th>
<th>Value label</th>
</tr>
</thead>
<tbody>
<tr>
<td>datevec</td>
<td>float</td>
<td>%tq</td>
<td></td>
</tr>
<tr>
<td>co2</td>
<td>double</td>
<td>%9.0g</td>
<td>* quarterly per capita co2 emissions</td>
</tr>
<tr>
<td>gdp</td>
<td>float</td>
<td>%9.0g</td>
<td>real per capita gdp</td>
</tr>
<tr>
<td>lp</td>
<td>float</td>
<td>%9.0g</td>
<td>labour productivity</td>
</tr>
<tr>
<td>tfp</td>
<td>float</td>
<td>%9.0g</td>
<td>quarterly utilization adjusted total factor productivity</td>
</tr>
<tr>
<td>p</td>
<td>float</td>
<td>%9.0g</td>
<td>relative price of investment</td>
</tr>
<tr>
<td>rc</td>
<td>float</td>
<td>%9.0g</td>
<td>real personal consumption expenditure</td>
</tr>
<tr>
<td>spread</td>
<td>float</td>
<td>%9.0g</td>
<td>credit spread on corporate bond yields</td>
</tr>
</tbody>
</table>
Time-series **co0** and **gdp** are nonstationary

- `tsline co2 if tin(1973q3,2018q1), nodraw name(tsline1)`
- `tsline gdp if tin(1973q3,2018q1), nodraw name(tsline2)`
- `graph combine tsline1 tsline2, rows(2)`
Transforming the data [Baum and Hurn(2021)]

```
. gen Dco2 = log(co2) - log(L1.co2)
. gen Dgdp = log(gdp) - log(L1.gdp)

. tsline Dgdp Dco2 if tin(1973q3,2018q1)
```
VAR model specification

- Autoregressive model with $K$ endogenous variables of order $p$

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + C_0 + u_t, \quad u_t \sim N(0, \Sigma)$$

where $A_i$ and $\Sigma$ are $K \times K$ matrices; $C_0$ and $u_t$ are $K \times 1$ vectors; $t = 1 \ldots T$.

- Autoregressive model with $m$ exogenous variables

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + C x_t + u_t, \quad u_t \sim N(0, \Sigma)$$

where $C$ is $K \times m$ matrix.

- Total number of coefficients is $K(Kp + m)$. 
Joint dynamics of CO2 and GDP - frequentist estimation

```
. var Dgdp Dco2, lags(1)
```

```
| Coefficient | Std. err. | z    | P>|z| | [95% conf. interval] |
|-------------|-----------|------|------|-----------------------|
| Dgdp        |           |      |      |                       |
|             | Dgdp      |      |      |                       |
|             | L1.       | .3332524 | .0719999 | 4.63 | 0.000 | .1921352 | .4743696 |
|             | Dco2      |      |      |                       |
|             | L1.       | .0503802 | .0272243 | 1.85 | 0.064 | -.0029784 | .1037388 |
|             | _cons     | .0044558 | .0007448 | 5.98 | 0.000 | .0029961 | .0059155 |
| Dco2        |           |      |      |                       |
|             | Dgdp      |      |      |                       |
|             | L1.       | .8378113 | .1935695 | 4.33 | 0.000 | .4584221 | 1.217201 |
|             | Dco2      |      |      |                       |
|             | L1.       | -.2124933 | .0731916 | -2.90 | 0.004 | -.3559463 | -.0690404 |
|             | _cons     | -.0089927 | .0020023 | -4.49 | 0.000 | -.0129172 | -.0050683 |
```
Bayesian VAR models

- Bayesian models assign priors to all model parameters

\[ y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + C x_t + u_t, \quad u_t \sim N(0, \Sigma) \]

- The prior for regression coefficient vector

\[ \beta = \text{vec}(A_1, \ldots, A_p, C) \]

is multivariate normal

\[ \beta \sim N(\beta_0, \Omega) \]

- The prior for covariance matrix \( \Sigma \) is either \( \text{InvWishart}(\alpha_0, S_0) \) or Jeffreys.
bayes: var prior families

- Conjugate Minnesota prior (default)
  . bayes, minnconjprior ... : var ...

- Minnesota prior with fixed covariance $\Sigma$
  . bayes, minnfixedcovprior ... : var ...

- Minnesota prior for $\beta$ and inverse-Wishart prior for $\Sigma$
  . bayes, minniwishprior ... : var ...

- Minnesota prior for $\beta$ and Jeffreys prior for $\Sigma$
  . bayes, minnjeffprior ... : var ...
Original Minnesota prior, [Litterman(1980)]

- Regression equations in our example
  
  \[ D_{gdp} = a_{11} L.D_{gdp} + a_{12} L.D_{co2} + c_1 + u_1 \]
  
  \[ D_{co2} = a_{21} L.D_{gdp} + a_{22} L.D_{co2} + c_2 + u_2 \]

- Error terms are independent with fixed variances
  
  \[ u_1 \sim N(0, \hat{\sigma}_1^2), \quad u_2 \sim N(0, \hat{\sigma}_2^2) \]

- Prior expectations are
  
  \[ E(a_{11}) = E(a_{22}) = 1, \quad E(a_{12}) = E(a_{21}) = E(c_1) = E(c_2) = 0 \]

- Prior variances are
  
  \[ Var(a_{11}) = Var(a_{22}) = \lambda_1^2, \quad Var(a_{12}) = Var(a_{21}) \sim \lambda_1^2 \lambda_2^2 \]
  
  \[ Var(c_1) = Var(c_2) = \lambda_1^2 \lambda_4^2 \]
bayes: var with fixed covariance

- Use the `minnfixedcovprior` option to specify this prior
- Original Minnesota prior with $\lambda_1 = 0.1$, $\lambda_2 = 0.5$ and $\lambda_4 = 100$
  
  . bayes, minnfixedcovprior: var Dgdp Dco2, lags(1)

- Increasing the self-tightness $\lambda_1$ to 1
  
  . bayes, minnfixedcovprior(selftight(1)): var Dgdp Dco2, lags(1)

- Specifying zero-mean priors for all coefficients
  
  . bayes, minnfixedcovprior(mean(0,0)): var Dgdp Dco2, lags(1)
**bayes: var** with fixed covariance and weakly informative prior

```
. bayes, minnfixedcovprior(mean(0,0) selftight(1)) mcmcsize(1000) rseed(17): ///
   var Dgdp Dco2, lags(1)
```

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. dev.</th>
<th>MCSE</th>
<th>Median</th>
<th>[95% cred. interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dgdp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dgdp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.332142</td>
<td>.0834622</td>
<td>.002643</td>
<td>.3323443</td>
<td>.1641772</td>
</tr>
<tr>
<td>Dco2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.049301</td>
<td>.0317361</td>
<td>.001048</td>
<td>.0489637</td>
<td>-.0125059</td>
</tr>
<tr>
<td>_cons</td>
<td>.004459</td>
<td>.0008826</td>
<td>.000029</td>
<td>.0044511</td>
<td>.0027347</td>
</tr>
<tr>
<td>Dco2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dgdp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.809345</td>
<td>.1995046</td>
<td>.005783</td>
<td>.8051736</td>
<td>.4242036</td>
</tr>
<tr>
<td>Dco2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>-.209947</td>
<td>.0756309</td>
<td>.002635</td>
<td>-.2093206</td>
<td>-.3591086</td>
</tr>
<tr>
<td>_cons</td>
<td>-.008832</td>
<td>.0021885</td>
<td>.000071</td>
<td>-.008834</td>
<td>-.0133508</td>
</tr>
</tbody>
</table>

Nikolay Balov (Stata)
Conjugate Minnesota prior

- Regression equations in our example

\[ D_{gdp} = a_{11} L.D_{gdp} + a_{12} L.D_{co2} + c_1 + u_1 \]
\[ D_{co2} = a_{21} L.D_{gdp} + a_{22} L.D_{co2} + c_2 + u_2 \]

- Assumption for error terms, \( \Sigma \) is being estimated

\[ (u_1, u_2) \sim MVN((0, 0), \Sigma) \]

- Prior expectations and covariance for coefficients are

\[ E(a_{11}) = E(a_{22}) = 1, \quad E(a_{12}) = E(a_{21}) = E(c_1) = E(c_2) = 0 \]
\[ \text{Var}(a_{11}, a_{12}, a_{21}, a_{22}, c_1, c_2) = \Sigma \otimes \Phi_0 \]
\[ \Phi_0 = \text{diag}(\frac{\lambda_1^2}{\hat{\sigma}_1^2}, \frac{\lambda_1^2}{\hat{\sigma}_2^2}, \lambda_1^2 \lambda_4^2) \]
Use the **minnconjprior** option to specify this prior

Original Minnesota prior with $\lambda_1 = 0.1$ and $\lambda_4 = 100$

```
bayes, minnconjprior: var Dgdp Dco2, lags(1)
```

Increasing the self-tightness $\lambda_1$ to 1

```
bayes, minnconjprior(selftight(1)): var Dgdp Dco2, lags(1)
```

Specifying zero-mean priors for all coefficients

```
bayes, minnconjprior(mean(0,0)): var Dgdp Dco2, lags(1)
```
```stata
bayes: var with conjugate Minnesota prior - Default

. bayes, mcmcsize(1000) rseed(17): var Dgdp Dco2, lags(1)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. dev.</th>
<th>MCSE</th>
<th>Median</th>
<th>[95% cred. interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dgdp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dgdp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.6210835</td>
<td>.0606735</td>
<td>.00201</td>
<td>.6197146</td>
<td>.4992695 .7426073</td>
</tr>
<tr>
<td>Dco2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dco2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.01825</td>
<td>.022934</td>
<td>.000725</td>
<td>.018633</td>
<td>-.028942 .0637309</td>
</tr>
<tr>
<td>_cons</td>
<td>.0024945</td>
<td>.0007528</td>
<td>.000022</td>
<td>.0025054</td>
<td>.0009637 .0039847</td>
</tr>
<tr>
<td>Dco2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dgdp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.3563922</td>
<td>.1864985</td>
<td>.005594</td>
<td>.3598283</td>
<td>-.0100042 .6929341</td>
</tr>
<tr>
<td>Dco2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dco2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.2501885</td>
<td>.0719771</td>
<td>.002437</td>
<td>.2459924</td>
<td>.1163743 .3961878</td>
</tr>
<tr>
<td>_cons</td>
<td>-.0044742</td>
<td>.0022312</td>
<td>.000069</td>
<td>-.0045744</td>
<td>-.0086671 .0001003</td>
</tr>
<tr>
<td>Sigma_1_1</td>
<td>.0000659</td>
<td>7.25e-06</td>
<td>2.1e-07</td>
<td>.0000651</td>
<td>.0000538 .0000822</td>
</tr>
<tr>
<td>Sigma_2_1</td>
<td>4.11e-07</td>
<td>.0000154</td>
<td>4.7e-07</td>
<td>-1.44e-07</td>
<td>-.000299 .0000324</td>
</tr>
<tr>
<td>Sigma_2_2</td>
<td>.0006276</td>
<td>.000069</td>
<td>2.2e-06</td>
<td>.0006235</td>
<td>.0005029 .0007709</td>
</tr>
</tbody>
</table>
```
MVN-inverse Wishart and MVN-Jeffreys priors

- Assumption for error terms
  \[(u_1, u_2) \sim MVN((0, 0), \Sigma)\]

- Prior for coefficients is
  \[\beta \sim N(\beta_0, \Omega_0)\]
  where \(\beta_0\) and \(\Omega_0\) are those from the original Minnesota prior.
  - Prior for \(\Sigma\) is \(InvWishart(\alpha_0, S_0)\), where \(\alpha_0 = 4\) and \(S_0 = \Sigma_0\). Use \texttt{minniwishprior} option to select this prior.
  - Prior for \(\Sigma\) is \(Jeffreys(K)\), Use \texttt{minnjeffprior} option to select this prior.
Selecting the number of lags - classical way

. varsoc, maxlag(4)

Lag-order selection criteria

| Lag | LL     | LR      | df | p       | FPE     | AIC     | HQIC    | SBIC    |
|-----+--------+---------+----+---------+---------+---------+---------+---------|
| 0   | 991.048| 2.6e-08 | -11.7744 | -11.7593 | -11.7372 |
| 1   | 1015.73 | 49.365  | 4  | 0.000   | 2.1e-08 | -12.0206 | -11.9753 | -11.909*|
| 2   | 1021.77 | 12.072* | 4  | 0.017   | 2.0e-08*| -12.0448*| -11.9694 | -11.8589|
| 3   | 1024.18 | 4.822   | 4  | 0.306   | 2.1e-08 | -12.0259 | -11.9203 | -11.7656|
| 4   | 1025.74 | 3.1279  | 4  | 0.537   | 2.1e-08 | -11.9969 | -11.8611 | -11.6622|

* optimal lag
Endogenous: Dgdp Dco2
Exogenous: _cons

**BIC** suggests 1-lag model, but **AIC** prefers 2-lag model.
Selecting the number of lags - Bayesian way

• In Bayes, we use posterior model probabilities and Bayes factors to select the optimal number of lags.
• First, fit candidate models with lags from 1 to 4 and save the results.
• Then, use **bayesstats ic** or **bayestest model** to compare the models.

```
.bayes, mcmcsize(1000) rseed(17) saving(bsim, replace): var Dgdp Dco2, lags(1/1)
.estimates store bvar1
.bayes, mcmcsize(1000) rseed(17) saving(bsim, replace): var Dgdp Dco2, lags(1/2)
.estimates store bvar2
.bayes, mcmcsize(1000) rseed(17) saving(bsim, replace): var Dgdp Dco2, lags(1/3)
.estimates store bvar3
.bayes, mcmcsize(1000) rseed(17) saving(bsim, replace): var Dgdp Dco2, lags(1/4)
.estimates store bvar4
```
The VAR(1) model is best

. bayestest model bvar1 bvar2 bvar3 bvar4

|         | log(ML) | P(M)  | P(M|y) |
|---------|---------|-------|-------|
| bvar1   | 948.666 | 0.250 | 0.9959|
| bvar2   | 943.180 | 0.250 | 0.0041|
| bvar3   | 936.468 | 0.250 | 0.0000|
| bvar4   | 934.196 | 0.250 | 0.0000|

. bayesstats ic bvar1 bvar2 bvar3 bvar4

<table>
<thead>
<tr>
<th></th>
<th>DIC</th>
<th>log(ML)</th>
<th>log(BF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bvar1</td>
<td>-1983.3</td>
<td>948.666</td>
<td></td>
</tr>
<tr>
<td>bvar2</td>
<td>-1971.2</td>
<td>943.179</td>
<td>-5.486693</td>
</tr>
<tr>
<td>bvar3</td>
<td>-1960.2</td>
<td>936.467</td>
<td>-12.19908</td>
</tr>
<tr>
<td>bvar4</td>
<td>-1955.8</td>
<td>934.196</td>
<td>-14.47085</td>
</tr>
</tbody>
</table>

Note: Marginal likelihood (ML) is computed using Laplace-Metropolis approximation.
Impulse-response functions and stability condition

- A VAR($p$) is **stable** if it has moving-average representation

\[ y_t = \mu + \sum_{i=0}^{t-1} \Phi_i u_{t-i} \]

- Impulse response matrices $\Phi_i$ satisfy

\[ \Phi_i = JA_iJ', \quad J = [I_K : 0 : \cdots : 0] \]

The $jk$-th coefficient $\phi_{jk,i}$ of $\Phi_i$ represents the response of the $j$-th outcome variable to a unit impulse in the $k$-th endogenous variable performed $i$ lags ago.

- **Stability condition**: moduli of all eigenvalues of $A$ are less than 1.
Stability check with `bayesvarstable` command

```
. bayes, mcmcsize(1000) rseed(17): var Dgdp Dco2, lags(1)
. bayes, saving(bsim, replace)
note: file bsim.dta saved.

. bayesvarstable

Eigenvalue stability condition

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>Equal-tailed modulus</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>MCSE</th>
<th>Median</th>
<th>[95% cred. interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.636941</td>
<td>.0618234</td>
<td>.001955</td>
<td>.6366748</td>
<td>.5171531</td>
<td>.7522063</td>
</tr>
<tr>
<td>2</td>
<td>.2343581</td>
<td>.0760976</td>
<td>.002406</td>
<td>.2295175</td>
<td>.0973634</td>
<td>.3945068</td>
</tr>
</tbody>
</table>

Pr(eigenvalues lie inside the unit circle) = 1.0000
```
Impulse response functions (IRF)

- IRFs estimate the marginal effect of a shock to one of the endogenous variables.
- IRFs of a VAR model include:
  - Regular impulse-response functions: `irf`
  - Cumulative IRFs: `cirf`
  - Orthogonal IRFs: `oirf`
  - Cumulative orthogonal IRFs: `coirf`
  - Dynamic-multiplier functions: `dm`
  - Cumulative dynamic-multiplier functions: `cdm`
  - Cholesky forecast-error decompositions: `fevd`
- There is one set of IRF results for each impulse-response pair.
- Bayesian estimates of a IRF include posterior mean and standard deviations, posterior median, and lower and upper credible interval bounds.
**bayesirf** subcommands

- **create** IRFs, dynamic-multipliers, and FEVDs
- **set** active IRF file
- **graph** of IRF results
- **cgraph** of combined IRF results
- **ograph** of overlaid IRF results
- **table** of IRF results
- **ctable** of combined IRF results
- **describe** content of active file
- **add** IRF results to active file
- **drop** IRF results from active file
- **rename** IRF results within a file
bayesirf create

- Creating IRF results with name `birf1` in IRF file `birf.irf`
  
  . bayesirf create birf1, set(birf)

- Changing level of credible intervals
  
  . bayesirf create birf1, set(birf) clevel(90) replace

- Changing the type of credible interval from equal-tailed to highest posterior density
  
  . bayesirf create birf1, set(birf) hpd replace

- Requesting 12 steps instead of the default 8 and saving MCMC results in `birfmcmc` dataset
  
  . bayesirf create birf1, set(birf) step(12) mcmcsaving(birfmcmc, replace)
Regular IRFs

`. bayesirf graph irf`

Graphs by irfname, impulse variable, and response variable
Combining regular and cumulative IRFs

`. bayesirf cgraph (birf1 Dgdp Dco2 irf) (birf1 Dgdp Dco2 cirf)`
Regular and orthogonal IRFs

- Regular IRFs are the coefficients of $\Phi_i$ in the moving-average representation

\[ y_t = \mu + \sum_{i=0}^{t-1} \Phi_i u_{t-i} \]

If the components of $u$ are correlated ($\Sigma \neq I_K$), then shocks are not independent.

- Orthogonal IRFs are coefficients of the transformed matrices $\Theta_i = \Phi_i P_c$, where $\Sigma = P_c P'_c$.

- $P_c$ depends on the order of endogenous variables.

- Orthogonal shocks are independent which allows for more realistic impulse response analysis.
Orthogonal IRFs

```
bayesirf graph oirf
```

Graphs by irfname, impulse variable, and response variable

95% equal-tailed CrI
Posterior mean of OIRF
**Forecast error decompositions**

```
.bayesirf table fevd, nocri
```

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fevd</td>
<td>fevd</td>
<td>fevd</td>
<td>fevd</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>.005596</td>
<td>0</td>
<td>.994404</td>
</tr>
<tr>
<td>2</td>
<td>.994134</td>
<td>.021325</td>
<td>.005866</td>
<td>.978675</td>
</tr>
<tr>
<td>3</td>
<td>.990852</td>
<td>.032362</td>
<td>.009148</td>
<td>.967638</td>
</tr>
<tr>
<td>4</td>
<td>.989353</td>
<td>.037859</td>
<td>.010647</td>
<td>.962141</td>
</tr>
<tr>
<td>5</td>
<td>.98869</td>
<td>.040397</td>
<td>.01131</td>
<td>.959603</td>
</tr>
<tr>
<td>6</td>
<td>.988394</td>
<td>.041562</td>
<td>.011606</td>
<td>.958438</td>
</tr>
<tr>
<td>7</td>
<td>.98826</td>
<td>.042104</td>
<td>.01174</td>
<td>.957896</td>
</tr>
<tr>
<td>8</td>
<td>.988198</td>
<td>.042363</td>
<td>.011802</td>
<td>.957637</td>
</tr>
<tr>
<td>9</td>
<td>.988168</td>
<td>.042489</td>
<td>.011832</td>
<td>.957511</td>
</tr>
<tr>
<td>10</td>
<td>.988154</td>
<td>.042552</td>
<td>.011846</td>
<td>.957448</td>
</tr>
<tr>
<td>11</td>
<td>.988147</td>
<td>.042585</td>
<td>.011853</td>
<td>.957415</td>
</tr>
<tr>
<td>12</td>
<td>.988143</td>
<td>.042602</td>
<td>.011857</td>
<td>.957398</td>
</tr>
</tbody>
</table>

Posterior means reported.

(1) irfname = birf1, impulse = Dgdp, and response = Dgdp.
(2) irfname = birf1, impulse = Dgdp, and response = Dco2.
(3) irfname = birf1, impulse = Dco2, and response = Dgdp.
(4) irfname = birf1, impulse = Dco2, and response = Dco2.
Dynamic multipliers - IRFs for exogenous variables

```
. bayes, mcmcsize(1000) rseed(17): var Dgdp Dco2, lags(1) exog(p)
. bayes, saving(bsim2, replace)
. bayesirf create birf2, set(birf)
. bayesirf table dm, irf(birf2) impulse(p) nocri
```

Results from `birf2`

```
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step</td>
<td>dm</td>
<td>dm</td>
</tr>
<tr>
<td>0</td>
<td>-.002208</td>
<td>-.004469</td>
</tr>
<tr>
<td>1</td>
<td>-.001413</td>
<td>-.001761</td>
</tr>
<tr>
<td>2</td>
<td>-.000892</td>
<td>-.000897</td>
</tr>
<tr>
<td>3</td>
<td>-.000566</td>
<td>-.000519</td>
</tr>
<tr>
<td>4</td>
<td>-.000362</td>
<td>-.000318</td>
</tr>
<tr>
<td>5</td>
<td>-.000234</td>
<td>-.000201</td>
</tr>
<tr>
<td>6</td>
<td>-.000152</td>
<td>-.000129</td>
</tr>
<tr>
<td>7</td>
<td>-.0001</td>
<td>-.000084</td>
</tr>
<tr>
<td>8</td>
<td>-.000067</td>
<td>-.000056</td>
</tr>
</tbody>
</table>
```

Posterior means reported.

(1) `irfname = birf2`, `impulse = p`, and `response = Dgdp`.
(2) `irfname = birf2`, `impulse = p`, and `response = Dco2`. 
Bayesian forecasting

- For a VAR(1) model, dynamic forecasts starting at time $T$ with horizon $h$ are drawn from the posterior predictive distribution

$$p(y_{T+1:T+h}|y_{1:T}) = \int f(y_{T+1:T+h}|y_{1:T}, \theta)p(\theta)d\theta$$

- For each draw $\theta^s = (\beta^s, \Sigma^s)$ from bayes:var’s MCMC sample

  1. $\tilde{y}^s_{T+1} = A^s_1 y_T + \cdots + A^s_p y_{T-p+1} + u^1, u^1 \sim N(0, \Sigma^s)$
  2. $\tilde{y}^s_{T+2} = A^s_1 \tilde{y}^s_{T+1} + \cdots + A^s_p y_{T-p} + u^2, u^2 \sim N(0, \Sigma^s)$

  ...  

  $h$. $\tilde{y}^s_{T+h} = A^s_1 \tilde{y}^s_{T+h-1} + \cdots + A^s_p \tilde{y}^s_{T-p+h} + u^h, u^h \sim N(0, \Sigma^s)$

Save dynamic forecast $(\tilde{y}^s_{T+1}, \tilde{y}^s_{T+2}, \ldots, \tilde{y}^s_{T+h})$
**bayesfcast subcommands**

- **compute** obtains dynamic forecasts
- **graph** dynamic forecasts, both frequentist and Bayesian

  . bayesfcast compute B_

  . bayesfcast compute B_, dynamic(tq(2001q1)) step(10)

  . bayesfcast compute B_, median

  . bayesfcast compute B_, hpd clevel(90)
Bayesian forecasting of $D_{gdp}$ and $D_{co2}$

Fit VAR(1) model to a smaller subsample of 62 observations.

\begin{verbatim}
. bayes, mcmcsize(1000) rseed(17): var Dgdp Dco2 if tin(1985q3,2000q4), lags(1)
. bayes, saving(bsim3, replace)

. bayesfcast compute B_, dynamic(tq(2001q1)) step(56)

. summ B_*
\end{verbatim}

\begin{verbatim}
Variable | Obs  Mean    Std. dev.  Min          Max
-------------+---------------------------------------------------------
   B_Dgdp    |  57  .0083962  .0004829  .0056679          .0090997
   B_Dgdp_sd |  57  .0129135  .0018603     0          .0139057
   B_Dgdp_lb |  57  -.0174288  .0035105  -.0209723          .0056679
   B_Dgdp_ub |  57   .0344242  .0042156  .0056679          .0373846
   B_Dco2    |  57  .0021036  .0035946  -.0001626        .0265668
   B_Dco2_sd |  57  .0298305  .0041074          0        .0323321
   B_Dco2_lb |  57  -.0562927  .0117139  -.0658958        .0265668
   B_Dco2_ub |  57   .0608982  .0056469  .0265668        .0698145
\end{verbatim}

Nikolay Balov (Stata) bayes:var June, 25th, 2021 35 / 46
Bayesian forecasting of Dgdp and Dco2

Frequentist (left) vs. Bayesian (right)

Forecast for Dgdp
Forecast for Dgdp
Forecast for Dco2
Forecast for Dco2

95% for CrI
Forecast
Observed
Sources of nonstationarity for vector time series

1. Deterministic trends
   - VAR with time trend
     \[ y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + C_1 t + C_0 + u_t \]
   - VAR with (exogenous) explanatory factors \( x \)
     \[ y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + C x_t + u_t \]

2. Stochastic trends
   - difference stationary VAR process
     \[ D \cdot y_t = A_1 D \cdot y_{t-1} + \cdots + A_p D \cdot y_{t-p} + C_0 + u_t \]
   - cointegrated VAR process - if there is \( B \) such that \( B y_t \) is stationary. Not supported by \texttt{bayes: var}. 
Example of unstable model due to nonstationarity

```
.bayes, rseed(17): var gdp co2, lags(1)
```

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. dev.</th>
<th>MCSE</th>
<th>Median</th>
<th>[95% cred. interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>gdp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gdp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>1.002089</td>
<td>.0023919</td>
<td>.000074</td>
<td>1.002153</td>
<td>.997157  1.006572</td>
</tr>
<tr>
<td><strong>co2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>co2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>-1477.896</td>
<td>13027.68</td>
<td>419.657</td>
<td>-1519.133</td>
<td>-25738.92  25721.34</td>
</tr>
<tr>
<td>_cons</td>
<td>52.841</td>
<td>106.1804</td>
<td>3.3495</td>
<td>53.52128</td>
<td>-162.343  248.5084</td>
</tr>
<tr>
<td><strong>co2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gdp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>-7.27e-09</td>
<td>4.25e-09</td>
<td>1.3e-10</td>
<td>-7.23e-09</td>
<td>-1.53e-08  1.37e-09</td>
</tr>
<tr>
<td>co2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.9516799</td>
<td>.0254478</td>
<td>.000801</td>
<td>.9516693</td>
<td>.9022775  1.003025</td>
</tr>
<tr>
<td>_cons</td>
<td>.0003766</td>
<td>.0002036</td>
<td>5.6e-06</td>
<td>.000374</td>
<td>-.0000575  .0007712</td>
</tr>
<tr>
<td><strong>Sigma_1_1</strong></td>
<td>5503.311</td>
<td>605.8631</td>
<td>18.5625</td>
<td>5479.42</td>
<td>4421.519  6809.851</td>
</tr>
<tr>
<td><strong>Sigma_2_1</strong></td>
<td>.0014116</td>
<td>.0007405</td>
<td>.000023</td>
<td>.0014273</td>
<td>-.0000697  .002903</td>
</tr>
<tr>
<td><strong>Sigma_2_2</strong></td>
<td>1.86e-08</td>
<td>2.08e-09</td>
<td>7.0e-11</td>
<td>1.84e-08</td>
<td>1.51e-08  2.33e-08</td>
</tr>
</tbody>
</table>

Nikolay Balov (Stata) bayes:var June, 25th, 2021 38 / 46
Verifying unstability - probability of inclusion is low

```
. bayes, saving(bsim3, replace)
. bayesvarstable
```

Eigenvalue stability condition

<table>
<thead>
<tr>
<th>Eigenvalue</th>
<th>Equal-tailed</th>
</tr>
</thead>
<tbody>
<tr>
<td>modulus</td>
<td>Mean</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>1</td>
<td>1.002658</td>
</tr>
<tr>
<td>2</td>
<td>.9511107</td>
</tr>
</tbody>
</table>

Pr(eigenvalues lie inside the unit circle) = 0.0930
Cummulative OIRFs of unstable model do not reach saturation

. bayesirf create birf4, set(birf) replace step(80)
. bayesirf cgraph (birf4 co2 co2 coirf) (birf4 co2 gdp coirf)
Forecasts of unstable model have ever expanding CrIs

```stata
.bayesfcast compute B_, dynamic(tq(1976q1)) step(162)
bayesfcast graph B_gdp B_co2, observed xlabels(none)
```

![Graph showing forecasts for GDP and CO2](image)

- **Forecast for GDP**
  - 95% equal-tailed CrI
  - Forecast
  - Observed

- **Forecast for CO2**
  - 95% equal-tailed CrI
  - Forecast
  - Observed
Deterministic trend modeled by exogenous variables

`. bayes, rseed(17): var gdp co2, lags(1) exog(p lp rc)`

<table>
<thead>
<tr>
<th>Equal-tailed</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>MCSE</th>
<th>Median</th>
<th>[95% cred. interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>gdp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gdp</td>
<td>.8535136</td>
<td>.0336989</td>
<td>.001084</td>
<td>.8537628</td>
<td>.7885342 – .9184529</td>
</tr>
<tr>
<td>L1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>co2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>-15989.47</td>
<td>13506.26</td>
<td>427.105</td>
<td>-16293.21</td>
<td>-41232.48 – 11829.05</td>
</tr>
<tr>
<td><strong>p</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>-454.0642</td>
<td>90.6652</td>
<td>2.76536</td>
<td>-455.861</td>
<td>-633.2077 – -278.1159</td>
</tr>
<tr>
<td><strong>lp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>1.605593</td>
<td>1.985492</td>
<td>.062787</td>
<td>1.634635</td>
<td>-2.262691 – 5.429637</td>
</tr>
<tr>
<td><strong>rc</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>110.742</td>
<td>25.53135</td>
<td>.829804</td>
<td>111.403</td>
<td>61.1454 – 161.2082</td>
</tr>
<tr>
<td><strong>_cons</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>co2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>-2.70e-08</td>
<td>6.52e-08</td>
<td>1.9e-09</td>
<td>-2.62e-08</td>
<td>-1.48e-07 – 1.05e-07</td>
</tr>
<tr>
<td><strong>p</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>-0.0005047</td>
<td>.0001759</td>
<td>5.6e-06</td>
<td>-.0005043</td>
<td>-0.0008504 – -0.0001696</td>
</tr>
<tr>
<td><strong>lp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>4.04e-06</td>
<td>4.03e-06</td>
<td>1.2e-07</td>
<td>3.94e-06</td>
<td>-3.40e-06 – 0.000012</td>
</tr>
<tr>
<td><strong>rc</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>0.000014</td>
<td>0.0000491</td>
<td>1.6e-06</td>
<td>0.0000144</td>
<td>-0.0000839 – 0.0001096</td>
</tr>
</tbody>
</table>

Nikolay Balov (Stata)
The model is stable

`. bayes, saving(bsim5, replace)`
`. bayesvarstable`

### Eigenvalue stability condition

<table>
<thead>
<tr>
<th>Eigenvalue modulus</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>MCSE</th>
<th>Median</th>
<th>[95% cred. interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.9268752</td>
<td>.0287381</td>
<td>.000909</td>
<td>.9290643</td>
<td>.8684965 .9794203</td>
</tr>
<tr>
<td>2</td>
<td>.8485238</td>
<td>.0350846</td>
<td>.001109</td>
<td>.8473363</td>
<td>.7803878 .9142691</td>
</tr>
</tbody>
</table>

Pr(eigenvalues lie inside the unit circle) = 0.9990
Cummulative OIRFs of stable model

```stata
. bayesirf create birf5, set(birf) replace step(80)
. bayesirf cgraph (birf5 co2 co2 coirf) (birf5 co2 gdp coirf)
```

![Graph showing Cummulative OIRFs for stable model](image)

- For `birf5: co2 -> co2`, the 95% equal-tailed CrI for `coirf` is shown with the posterior mean.
- For `birf5: co2 -> gdp`, the 95% equal-tailed CrI for `coirf` is also shown with the posterior mean.
Forecasts of stable model

- `. bayesfcast compute B_, dynamic(tq(1976q1)) step(162)`
- `. bayesfcast graph B_gdp B_co2, observed xlabels(none)`

---

![Forecast graphs for GDP and CO2](image)

- **Forecast for GDP**
- **Forecast for CO2**

- **Observed**
- **Forecast**

- **95% equal-tailed CrI**

---

Nikolay Balov (Stata)
Conclusion - why you should use Bayesian VARs

- Flexibility in providing a variety of prior information
- Prior support for weak VAR inference when based on small datasets
- Sensitivity analysis and stability control through prior variation
- Methodological and consistent way of selecting number of lags
- More reliable forecasting using predictive posterior distribution
*Environmental Econometrics Using Stata.*
College Station, TX: Stata Press.

Environmental impacts of a North American Free Trade Agreement.

Litterman, R. B. 1980.
A Bayesian procedure for forecasting with vector autoregressions.
Mit working paper, Massachusetts Institute of Technology.