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

$$\mathbf{y} = (y_1, y_2, \ldots y_K)'$$

observed at time $t = 1 \ldots T$. Variables y_1, ..., y_K are called *endogenous*.

- VAR model is just a multivariate normal regression of $\mathbf{y}_t$ on its own-lags

$$\mathbf{y}_t \sim \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p}$$

- It is possible to add *exogenous* variables x_1, ..., x_m to the model

$$\mathbf{y}_t \sim \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p}, x_1, \ldots, x_m$$

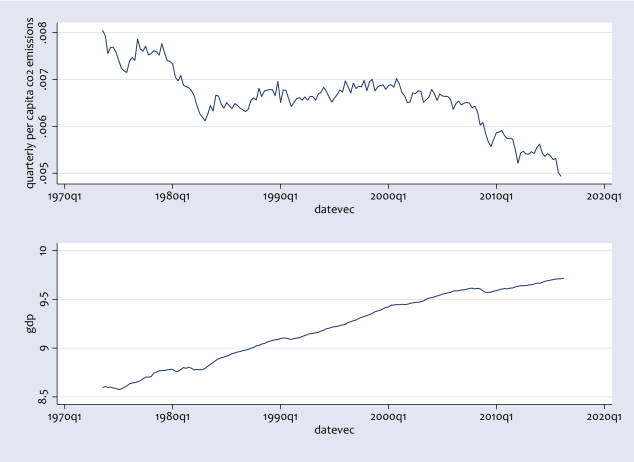
- Advantage: simplicity - no need for special structure of covariates.
- Disadvantage: large number of regression coefficients, $K(Kp + m)$; difficult interpretation.

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
  ---------------------------------------------------------------------
  Variable      Storage   Display    Value
      name         type    format    label      Variable label
  ---------------------------------------------------------------------
  datevec         float   %tq
  co2             double  %9.0g               * quarterly per capita co2
                                                  emissions
  gdp             float   %9.0g                 real per capita gdp
  lp              float   %9.0g                 labour productivity
  tfp             float   %9.0g                 quarterly utilization adjusted
                                                  total factor productivity
  p               float   %9.0g                 relative price of investment
  rc              float   %9.0g                 real personal consumption
                                                  expenditure
  spread          float   %9.0g                 credit spread on corporate bond
                                                  yields
  ```
  ```

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

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{C}_0 + \mathbf{u}_t, \;\; \mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma})$$

  where $\mathbf{A}_l$ and $\boldsymbol{\Sigma}$ are $K \times K$ matrices; $\mathbf{C}_0$ and $\mathbf{u}_t$ are $K \times 1$ vectors; $t = 1 \ldots T$.

- Autoregressive model with m exogenous variables

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{C}\mathbf{x}_t + \mathbf{u}_t, \;\; \mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma})$$

  where $\mathbf{C}$ 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

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{C}\mathbf{x}_t + \mathbf{u}_t, \;\; \mathbf{u}_t \sim N(\mathbf{0}, \mathbf{\Sigma})$$

- The prior for regression coefficient vector

$$\beta = vec(\mathbf{A}_1, \ldots \mathbf{A}_p, \mathbf{C})$$

  is multivariate normal

$$\beta \sim N(\beta_0, \mathbf{\Omega})$$

- The prior for covariance matrix $\mathbf{\Sigma}$ is either $InvWishart(\alpha_0, \mathbf{S}_0)$ or Jeffreys.

bayes: var prior families

- Conjugate Minnesota prior (default)

```
. bayes, minnconjprior ... : var ...
```

- Minnesota prior with fixed covariance $\boldsymbol{\Sigma}$

```
. bayes, minnfixedcovprior ... : var ...
```

- Minnesota prior for β and inverse-Wishart prior for $\boldsymbol{\Sigma}$

```
. bayes, minniwishprior ... : var ...
```

- Minnesota prior for β and Jeffreys prior for $\boldsymbol{\Sigma}$

```
. bayes, minnjeffprior ... : var ...
```

Original Minnesota prior, [Litterman(1980)]

- Regression equations in our example

$$Dgdp = a_{11} L.Dgdp + a_{12} L.Dco2 + c_1 + u_1$$
$$Dco2 = a_{21} L.Dgdp + a_{22} L.Dco2 + c_2 + u_2$$

- Error terms are independent with fixed variances

$$u_1 \sim N(0, \hat{\sigma}_1^2),\ \ u_2 \sim N(0, \hat{\sigma}_2^2)$$

- Prior expectations are

$$E(a_{11}) = E(a_{22}) = 1,\ \ E(a_{12}) = E(a_{21}) = E(c_1) = E(c_2) = 0$$

- Prior variances are

$$Var(a_{11}) = Var(a_{22}) = \lambda_1^2,\ \ Var(a_{12}) = Var(a_{21}) \sim \lambda_1^2 \lambda_2^2$$
$$Var(c_1) = Var(c_2) = \lambda_1^2 \lambda_4^2$$

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 λ_1 ro 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)

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |      Mean   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
Dgdp         |
        Dgdp |
         L1. |  .3321418   .0834622   .002643   .3323443   .1641772   .4912778
             |
        Dco2 |
         L1. |  .0493012   .0317361   .001048   .0489637  -.0125059   .1184353
             |
       _cons |  .0044588   .0008826   .000029   .0044511   .0027347   .0062784
-------------+----------------------------------------------------------------
Dco2         |
        Dgdp |
         L1. |  .8093455   .1995046   .005783   .8051736   .4242036   1.199603
             |
        Dco2 |
         L1. | -.2099469   .0756309   .002635  -.2093206  -.3591086  -.0685509
             |
       _cons | -.0088317   .0021885   .000071   -.008834  -.0133508  -.0046639
------------------------------------------------------------------------------
```

Conjugate Minnesota prior

- Regression equations in our example

$$Dgdp = a_{11} L.Dgdp + a_{12} L.Dco2 + c_1 + u_1$$

$$Dco2 = a_{21} L.Dgdp + a_{22} L.Dco2 + c_2 + u_2$$

- Assumption for error terms, $\mathbf{\Sigma}$ is being estimated

$$(u_1, u_2) \sim MVN((0,0), \mathbf{\Sigma})$$

- Prior expectations and covariance for coefficients are

$$E(a_{11}) = E(a_{22}) = 1,\ E(a_{12}) = E(a_{21}) = E(c_1) = E(c_2) = 0$$

$$Var(a_{11}, a_{12}, a_{21}, a_{22}, c_1, c_2) = \mathbf{\Sigma} \otimes \mathbf{\Phi}_0$$

$$\mathbf{\Phi}_0 = diag(\frac{\lambda_1^2}{\hat{\sigma}_1^2},\ \frac{\lambda_1^2}{\hat{\sigma}_2^2},\ \lambda_1^2 \lambda_4^2)$$

bayes: var with conjugate Minnesota prior

- 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 λ_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)
```

bayes: var with conjugate Minnesota prior - Default

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

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |      Mean   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
Dgdp         |
        Dgdp |
         L1. |  .6210835   .0606735    .00201   .6197146   .4992695   .7426073
             |
        Dco2 |
         L1. |    .01825    .022934   .000725    .018633   -.028942   .0637309
             |
       _cons |  .0024945   .0007528   .000022   .0025054   .0009637   .0039847
-------------+----------------------------------------------------------------
Dco2         |
        Dgdp |
         L1. |  .3563922   .1864985   .005594   .3598283  -.0100042   .6929341
             |
        Dco2 |
         L1. |  .2501885   .0719771   .002437   .2459924   .1163743   .3961878
             |
       _cons | -.0044742   .0022312   .000069  -.0045744  -.0086671   .0001003
-------------+----------------------------------------------------------------
   Sigma_1_1 |  .0000659   7.25e-06   2.1e-07   .0000651   .0000538   .0000822
   Sigma_2_1 |  4.11e-07   .0000154   4.7e-07  -1.44e-07  -.0000299   .0000324
   Sigma_2_2 |  .0006276    .000069   2.2e-06   .0006235   .0005029   .0007709
------------------------------------------------------------------------------
```

MVN-inverse Wishart and MVN-Jeffreys priors

- Assumption for error terms

$$(u_1, u_2) \sim MVN((0,0), \mathbf{\Sigma})$$

- Prior for coefficients is

$$\beta \sim N(\beta_0, \mathbf{\Omega}_0)$$

where β_0 and $\mathbf{\Omega}_0$ are those from the original Minnesota prior.

  - Prior for $\mathbf{\Sigma}$ is $InvWishart(\alpha_0, \mathbf{S}_0)$, where $\alpha_0 = 4$ and $\mathbf{S}_0 = \mathbf{\Sigma}_0$. Use **minniwishprior** option to select this prior.
  - Prior for $\mathbf{\Sigma}$ is $Jeffreys(K)$, Use **minnjeffprior** option to select this prior.

Selecting the number of lags - classical way

```
. varsoc, maxlag(4)

Lag-order selection criteria

   Sample: 1974q2 thru 2016q1                          Number of obs = 168
  +---------------------------------------------------------------------------+
  | 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 |   948.6666      0.2500     0.9959
       bvar2 |   943.1799      0.2500     0.0041
       bvar3 |   936.4675      0.2500     0.0000
       bvar4 |   934.1958      0.2500     0.0000
---------------------------------------------

. bayesstats ic bvar1 bvar2 bvar3 bvar4

---------------------------------------------
             |        DIC     log(ML)    log(BF)
-------------+-------------------------------
       bvar1 |  -1983.293   948.6666          .
       bvar2 |  -1971.234   943.1799  -5.486693
       bvar3 |  -1960.184   936.4675  -12.19908
       bvar4 |  -1955.828   934.1958  -14.47085
---------------------------------------------
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

$$\mathbf{y}_t = \mu + \sum_{i=0}^{t-1} \mathbf{\Phi}_i \mathbf{u}_{t-i}$$

- Impulse response matrices $\mathbf{\Phi_i}$ satisfy

$$\mathbf{\Phi}_i = \mathbf{J}\mathbf{A}^i\mathbf{J}', \mathbf{J} = [\mathbf{I}_K : \mathbf{0} : \cdots : \mathbf{0}]$$

  The jk-th coefficient $\phi_{jk,i}$ of $\mathbf{\Phi_i}$ 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 $\mathbf{A}$ 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                    Companion matrix size =    2
                                                  MCMC sample size      = 1000

------------------------------------------------------------------------------
  Eigenvalue |                                                Equal-tailed
     modulus |      Mean   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
           1 |   .636941   .0618234   .001955   .6366748   .5171531   .7522063
           2 |  .2343581   .0760976   .002406   .2295175   .0973634   .3945068
------------------------------------------------------------------------------
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 intrval from equal-tailed to highest posterior density

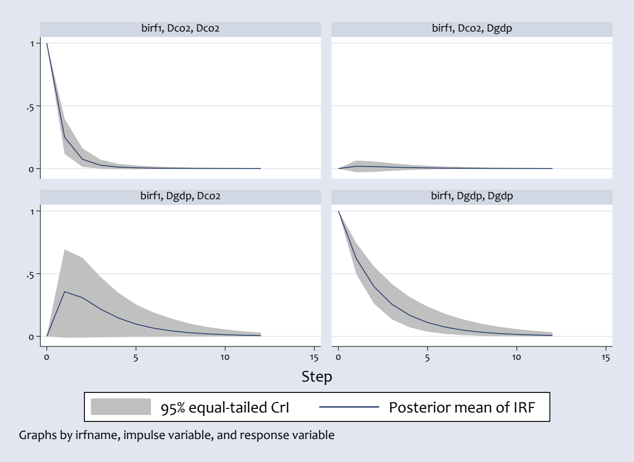
```
. bayesirf create birf1, set(birf) hpd replace
```

- 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)
```

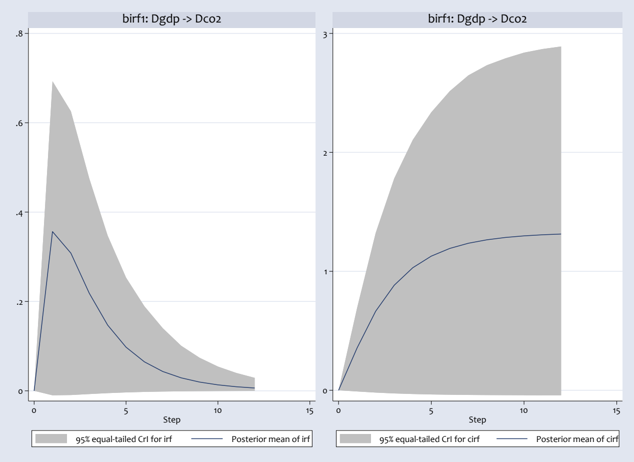
Regular IRFs

```
. bayesirf graph irf
```



Combining regular and cumulative IRFs

```
. bayesirf cgraph (birf1 Dgdp Dco2 irf) (birf1 Dgdp Dco2 cirf)
```



Regular and orthogonal IRFs

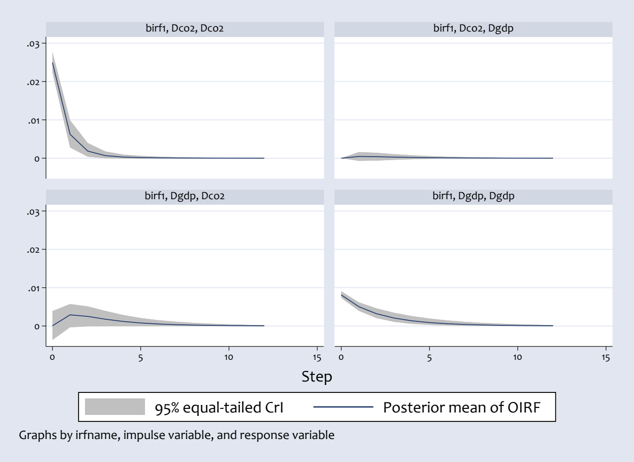
- Regular IRFs are the coefficients of $\mathbf{\Phi_i}$ in the moving-average representation

$$\mathbf{y}_t = \mu + \sum_{i=0}^{t-1} \mathbf{\Phi}_i \mathbf{u}_{t-i}$$

  If the components of $\mathbf{u}$ are correlated ($\mathbf{\Sigma} \neq \mathbf{I}_K$), then shocks are not independent.
- Orthogonal IRFs are coefficients of the transformed matrices $\mathbf{\Theta}_i = \mathbf{\Phi}_i \mathbf{P_c}$, where $\mathbf{\Sigma} = \mathbf{P}_c \mathbf{P}_c'$.
- $\mathbf{P}_c$ depends on the order of endogenous variables.
- Orthogonal shocks are independent which allows for more realistic impulse response analysis.

Orthogonal IRFs

```
. bayesirf graph oirf
```



Forecast error decompositions

```
. bayesirf table fevd, nocri

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

--------------------------------
         |      (1)         (2)
    Step |       dm          dm
---------+----------------------
       0 | -.002208    -.004469
       1 | -.001413    -.001761
       2 | -.000892    -.000897
       3 | -.000566    -.000519
       4 | -.000362    -.000318
       5 | -.000234    -.000201
       6 | -.000152    -.000129
       7 |   -.0001    -.000084
       8 | -.000067    -.000056
--------------------------------
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

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

- For each draw $\theta^s = (\beta^s, \mathbf{\Sigma}^s)$ from **bayes:var**'s MCMC sample

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

   ...

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

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

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 **Dgdp** and **Dco2**

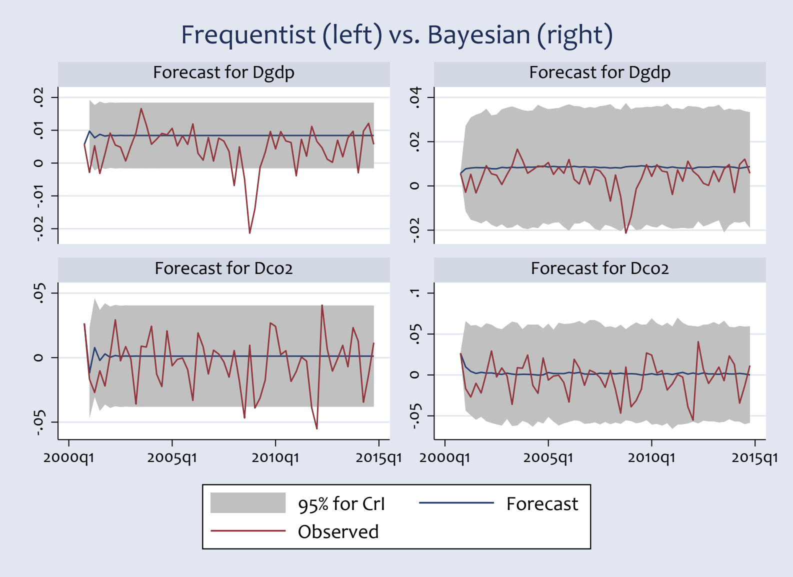
Fit VAR(1) model to a smaller subsample of 62 observations.

```
. 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_*
    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
```

Bayesian forecasting of **Dgdp** and **Dco2**



Sources of nonstationarity for vector time series

1. Deterministic trends
   - VAR with time trend

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{C}_1 t + \mathbf{C}_0 + \mathbf{u}_t$$

   - VAR with (exogenous) explanatory factors **x**

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{C}\mathbf{x}_t + \mathbf{u}_t$$

2. Stochastic trends
   - difference stationary VAR process

$$\mathbf{D}.\mathbf{y}_t = \mathbf{A}_1\mathbf{D}.\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{D}.\mathbf{y}_{t-p} + \mathbf{C}_0 + \mathbf{u}_t$$

   - cointegrated VAR process - if there is B such that $B\mathbf{y}_t$ is stationary. Not supported by **bayes: var**.

Example of unstable model due to nonstationarity

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

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

Verifying unstability - probability of inclusion is low

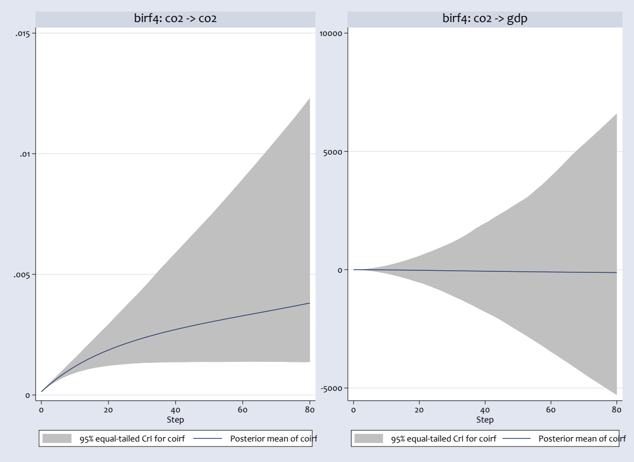
```
. bayes, saving(bsim3, replace)
. bayesvarstable
```

```
Eigenvalue stability condition                   Companion matrix size =    2
                                                 MCMC sample size      = 1000

------------------------------------------------------------------------------
  Eigenvalue |                                                Equal-tailed
     modulus |      Mean   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
           1 |  1.002658   .0031518     .0001   1.002474   .9984661   1.007404
           2 |  .9511107   .0249542   .000789    .951537   .9035589    1.00106
------------------------------------------------------------------------------
Pr(eigenvalues lie inside the unit circle) = 0.0930
```

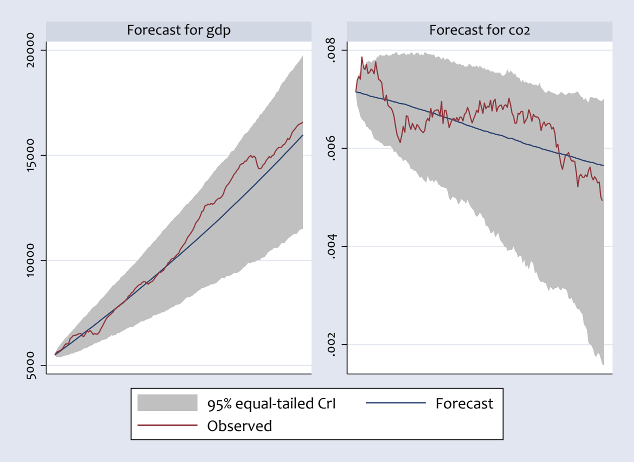
Cummulative OIRFs of unstable model do not reach saturation

```
. 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

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



Deterministic trend modeled by exogenous variables

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

------------------------------------------------------------------------------
             |                                                Equal-tailed
             |      Mean   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
gdp          |
         gdp |
         L1. |  .8535136   .0336989   .001084   .8537628   .7885342   .9184529
             |
         co2 |
         L1. | -15989.47   13506.26   427.105  -16293.21  -41232.48   11829.05
             |
           p | -454.0642    90.6652   2.76536   -455.861  -633.2077  -278.1159
          lp |  1.605593   1.985492   .062787   1.634635  -2.262691   5.429637
          rc |   110.742   25.53135   .829804    111.403    61.1454   161.2082
       _cons | -182.1933   145.6657   4.60636  -178.4868  -501.1311    98.1174
-------------+----------------------------------------------------------------
co2          |
         gdp |
         L1. | -2.70e-08   6.52e-08   1.9e-09  -2.62e-08  -1.48e-07   1.05e-07
             |
         co2 |
         L1. |  .9217557    .025128   .000824   .9228897   .8715538   .9682776
             |
           p | -.0005047   .0001759   5.6e-06  -.0005043  -.0008504  -.0001696
          lp |  4.04e-06   4.03e-06   1.2e-07   3.94e-06  -3.40e-06    .000012
          rc |   .000014   .0000491   1.6e-06   .0000144  -.0000839   .0001096
        ...
```

The model is stable

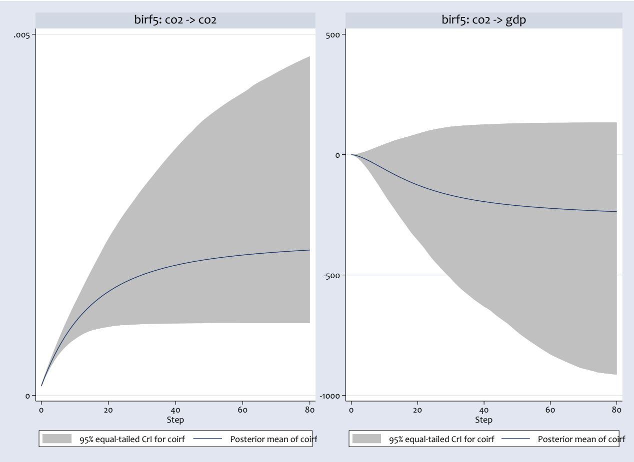
```
. bayes, saving(bsim5, replace)
. bayesvarstable

Eigenvalue stability condition                  Companion matrix size =    2
                                                MCMC sample size      = 1000

------------------------------------------------------------------------------
  Eigenvalue |                                                Equal-tailed
     modulus |       Mean   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
           1 |  .9268752   .0287381   .000909   .9290643   .8684965   .9794203
           2 |  .8485238   .0350846   .001109   .8473363   .7803878   .9142691
------------------------------------------------------------------------------
Pr(eigenvalues lie inside the unit circle) = 0.9990
```

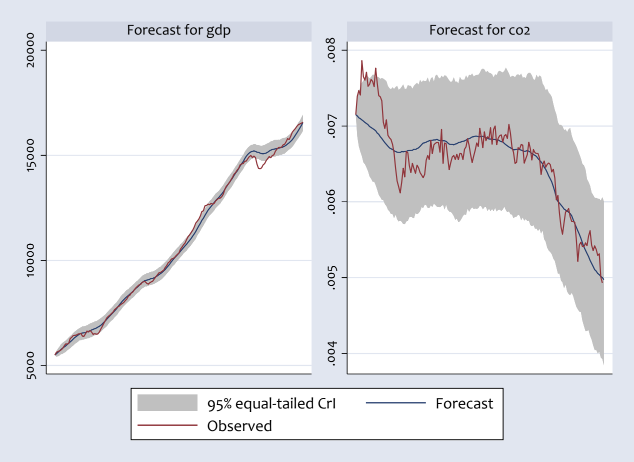
Cummulative OIRFs of stable model

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



Forecasts of stable model

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



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

Baum, C. F., and S. Hurn. 2021.
Environmental Econometrics Using Stata.
College Station, TX: Stata Press.

Grossman, G., and A. Krueger. 1993.
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.