Multivariate Time Series: Part 5

Granger Causality, Identification in VARs (and ECMs), Bayesian VARs

Gerald P. Dwyer

Clemson University

April 2016

Outline

Granger causality and simultaneous systems

- Simultaneous system with n variables and lagged values (common in time series) of all variables

$$\mathbf{\Gamma}_0 \mathbf{x}_t = \mathbf{\Gamma}_1 \mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t$$

  - Think of variables as being price, quantity, income and weather
- A vector autoregression from this is

$$\mathbf{x}_t = \mathbf{\Gamma}_0^{-1} \mathbf{\Gamma}_1 \mathbf{x}_{t-1} + \mathbf{\Gamma}_0^{-1} \boldsymbol{\varepsilon}_t$$

- Plausible that weather and income are not affected by the price and quantity of some goods – say beer
- Granger causality tests that supposition with the data
  - The weather does not "Granger cause" the price if weather does not help to predict the price
  - The weather does "Granger cause" the price if the weather helps to predict the price

Granger causality and simultaneous systems I

- Simultaneous system with n variables and lagged values, common in time series (with constant suppressed)

$$\mathbf{\Gamma}_0 \mathbf{x}_t = \mathbf{\Gamma}_1 \mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t$$

- Reduced form

$$\mathbf{x}_t = \mathbf{\Gamma}_{\mathbf{0}}^{-\mathbf{1}} \mathbf{\Gamma}_1 \mathbf{x}_{t-1} + \mathbf{\Gamma}_{\mathbf{0}}^{-\mathbf{1}} \boldsymbol{\varepsilon}_t$$

- Have to allow for inter-relationships across equations
- When more than two variables, can look at block diagonality of simultaneous system
  - If, in the structural equations, weather appears in the structural price and quantity equations but the price and quantity do not appear in the weather equation
  - Then, in the reduced form, lagged values of the price and quantity will not help to predict the weather
- The relationship between Granger causality and econometric exogeneity is complicated

Simultaneous systems and block exogeneity I

- Consider the four-variable set of equations from earlier
- Demand and supply

$$q_t = [\alpha_0 + \alpha(\mathsf{L})]\, p_t + \alpha^*(\mathsf{L})\, q_{t-1} + \beta y_t + \varepsilon_t^d$$
$$q_t = [\gamma_0 + \gamma(\mathsf{L})]\, p_t + \gamma^*(\mathsf{L})\, q_{t-1} + \delta w_t + \varepsilon_t^s$$

- where

$$\alpha_0 + \alpha(\mathsf{L}) = \alpha_0 + \sum_{i=1}^{k} \alpha_i \mathsf{L}^i \qquad \alpha^*(\mathsf{L}) = \sum_{i=0}^{k} \alpha^{*i} \mathsf{L}^i$$
$$\gamma_0 + \gamma(\mathsf{L}) = \gamma_0 + \sum_{i=1}^{k} \gamma_i \mathsf{L}^i \qquad \gamma^*(\mathsf{L}) = \sum_{i=0}^{k} \gamma^{*i} \mathsf{L}^i$$

- Income and weather are determined by

$$y_t = \Pi_{33}(\mathsf{L})\, y_{t-1} + u_{t3}$$
$$w_t = \Pi_{44}(\mathsf{L})\, w_{t-1} + u_{t4}$$

Simultaneous systems and block exogeneity II

- The reduced form is

$$
\begin{aligned}
q_t &= \Pi_{11}(\mathsf{L})\, q_{t-1} + \Pi_{12}(\mathsf{L})\, p_{t-1} + \Pi_{13}(\mathsf{L})\, y_{t-1} + \Pi_{14}(\mathsf{L})\, w_{t-1} + u_{t1} \\
p_t &= \Pi_{21}(\mathsf{L})\, q_{t-1} + \Pi_{22}(\mathsf{L})\, p_{t-1} + \Pi_{23}(\mathsf{L})\, y_{t-1} + \Pi_{24}(\mathsf{L})\, w_{t-1} + u_{t2} \\
y_t &= \Pi_{33}(\mathsf{L})\, y_{t-1} + u_{t3} \\
w_t &= \Pi_{44}(\mathsf{L})\, w_{t-1} + u_{t4}
\end{aligned}
$$

- The test for block exogeneity is a test whether price and quantity help to predict income and the weather
  - This is the same as a test whether price and quantity do not Granger cause income and the weather
- If price and quantity do not help to predict income and the weather, then there exists a structural representation of the equations in which income and weather are exogenous to price and quantity
- Failure to pass this test does not necessarily imply that income and weather are exogenous

Simultaneous systems and block exogeneity III

- Suppose that income and the weather are affected by past prices and quantities
  - Then the reduced form will include these lagged variables even though there is no contemporaneous relationship
  - If the errors across equations are uncorrelated, income and the weather are exogenous to price and quantity despite the effects of lagged prices and income
- Suppose that the errors in the price and quantity equations are correlated with the errors in the income and the weather equations
  - Then income and the weather are not exogenous in this system relative to the price and quantity even though the reduced form is block diagonal
  - This caveat is part of the reason the statement above says
  - "If price and quantity do not help to predict income and the weather, then there exists a structural representation of the equations in which income and weather are exogenous to price and quantity"
  - This may not be the structural representation we want

Granger causality and cointegrated systems

- If variables are cointegrated, there is a Vector Error Correction Mechanism relating them (with constant suppressed)

$$\Delta \mathbf{x}_t = \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{x}_{t-1} + \sum_{i=1}^{\infty} \boldsymbol{\Gamma}_i \Delta \mathbf{x}_{t-i} + \mathbf{u}_t$$

- Granger causality: Coefficients that must be zero are coefficients of lagged values of, say, z_t and any cointegrating relationships in which the variable z_t appears in $\boldsymbol{\beta}' \mathbf{x}_{t-1}$

Zero restrictions

- An obvious way to proceed to identify a set of structural equations is assume exogeneity
- Granger causality tests can be a way to see whether such restrictions are plausible

Recursive Systems I

- Assuming a recursive system is a way to identify a set of equations that is estimable by OLS
- Also called a "Wold causal chain"
- Assume that demand and supply are given by

$$
\begin{aligned}
q_t &= \alpha_0 p_t + \alpha \left(\mathrm{L}\right) p_{t-1} + \alpha^* \left(\mathrm{L}\right) q_{t-1} + \beta y_t + \varepsilon_t^d \\
p_t &= \gamma \left(\mathrm{L}\right) p_{t-1} + \gamma_0^* q_t + \gamma^* \left(\mathrm{L}\right) q_{t-1} + \delta w_t + \varepsilon_t^s
\end{aligned}
$$

- and assume $\alpha_0 = 0$ and

$$
\begin{aligned}
&\mathrm{Cov}\left[y_t, \varepsilon_t^d\right] = \mathrm{Cov}\left[y_t, \varepsilon_t^s\right] = \mathrm{Cov}\left[w_t, \varepsilon_t^d\right] = \mathrm{Cov}\left[w_t, \varepsilon_t^s\right] \\
&= \mathrm{Cov}\left[\varepsilon_t^d, \varepsilon_t^s\right] = 0
\end{aligned}
$$

- Then demand and supply are identified and they can be estimated by OLS

Recursive Systems II

- Often discussed in terms of "ordering" in which current price does not affect current quantity and current quantity is affected by current price
- Here, quantity supplied is independent of the price this period and the price has to be just right to induce buyers to produce the quantity supplied
  - A vertical supply curve this period
- Common to write it as

$$q_t = \alpha_0 p_t + \sum_{i=0}^{k} \alpha_i p_{t-i} \sum_{i=0}^{k} \alpha_i^* q_{t-i} + \beta y_t + \varepsilon_t^d$$

$$p_t = \sum_{i=0}^{k} \gamma_i p_{t-i} + \sum_{i=0}^{k} \gamma_i^* q_{t-i} + \delta w_t + \varepsilon_t^s$$

with the variable determined "first" at the bottom

Choleski Decomposition and Recursive Models

- A Choleski decomposition often is referenced as how this is done
- A Choleski decompositions used to be used to invert matrices
  - Singular value decompositions were discovered later and are a better way to invert a matrix
- For a real positive definite matrix

$$\mathbf{A} = \mathbf{L}\mathbf{L}'$$

- where $\mathbf{L}$ is a lower triangular matrix
  - A lower triangular matrix $\mathbf{L}$ has all entries above the main diagonal equal to zero
- Imposing a recursive system takes a general matrix and imposes a triangular structure on the structural coefficients for contemporaneous values
- It imposes the same structure on the relationship between the reduced-form error terms
- Can use that structure to compute coefficients in structural equation without re-estimating equations
  - Not worth worrying about – not that hard an estimation problem

Basic setup of VAR for further identification discussion I

- I will follow Fry and Pagan's (Jnl Econ Lit, 2011) discussion with some reversion to the notation in Enders and prior notes
- Start from a first-order VAR for n variables

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{e}_t$$
$$\mathrm{E}\, \mathbf{e}_t = \mathbf{0}, \quad \mathrm{E}\, \mathbf{e}_t \mathbf{e}_t' = \mathbf{\Omega}, \quad \mathrm{E}\, \mathbf{e}_t \mathbf{e}_s' = \mathbf{0} \;\forall\; t \neq s$$

- A structural VAR (SVAR) for these data is given by

$$\mathbf{B}_0 \mathbf{y}_t = \mathbf{B}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$
$$\mathrm{E}\, \boldsymbol{\varepsilon}_t = \mathbf{0}, \quad \mathrm{E}\, \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' = \mathbf{\Sigma}, \quad \mathrm{E}\, \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s' = \mathbf{0} \;\forall\; t \neq s$$
$$\mathrm{E}\, \varepsilon_{it} \varepsilon_{jt} = \sigma_i \text{ for } i = j \text{ and } \mathrm{E}\, \varepsilon_{it} \varepsilon_{jt} = 0 \text{ for } i \neq j$$

- If $\mathbf{B}_0$ is invertible, the SVAR implies

$$\mathbf{y}_t = \mathbf{B}_0^{-1} \mathbf{B}_1 \mathbf{y}_{t-1} + \mathbf{B}_0^{-1} \boldsymbol{\varepsilon}_t$$

Basic setup of VAR for further identification discussion II

- and therefore

$$\mathbf{B}_0^{-1}\mathbf{B}_1 = \mathbf{A}_1 \text{ and } \mathbf{B}_0^{-1}\varepsilon_t = \mathbf{e}_t$$

- The moving average representation of the VAR is given by

$$\mathbf{y}_t = \mathbf{D}_0\mathbf{e}_t + \mathbf{D}_1\mathbf{e}_{t-1} + \mathbf{D}_2\mathbf{e}_{t-2} + \ldots$$

- The moving average representation of the SVAR is given by

$$\mathbf{y}_t = \mathbf{C}_0\varepsilon_t + \mathbf{C}_1\varepsilon_{t-1} + \mathbf{C}_2\varepsilon_{t-2} + \ldots$$

- The error terms in the VAR and SVAR are related by

$$\mathbf{e}_t = \mathbf{B}_0^{-1}\varepsilon_t$$

  and therefore

$$\mathbf{C}_0 = \mathbf{B}_0^{-1}$$
$$\mathbf{C}_j = \mathbf{D}_j\mathbf{B}_0^{-1} = \mathbf{D}_j\mathbf{C}_0$$

Basic setup of VAR for further identification discussion III

- The impulse response functions
- The k'th period-ahead impulse responses for the VAR with an impulse at 0 are given by

$$
\begin{aligned}
\mathbf{y}_0 &= \mathbf{D}_0 \mathbf{e}_0 \\
\mathbf{y}_1 &= \mathbf{D}_1 \mathbf{e}_0 \\
\mathbf{y}_2 &= \mathbf{D}_2 \mathbf{e}_0 \\
&\dots
\end{aligned}
$$

- The k'th period-ahead impulse responses for the SVAR with an impulse at 0 are given by

$$
\begin{aligned}
\mathbf{y}_0 &= \mathbf{C}_0 \varepsilon_0 \\
\mathbf{y}_1 &= \mathbf{C}_1 \varepsilon_0 \\
\mathbf{y}_2 &= \mathbf{C}_2 \varepsilon_0 \\
&\dots
\end{aligned}
$$

Basic setup of VAR for further identification discussion IV

- Because

$$\mathbf{C}_0 = \mathbf{B}_0^{-1}$$
$$\mathbf{C}_j = \mathbf{D}_j \mathbf{B}_0^{-1} = \mathbf{D}_j \mathbf{C}_0$$

- it is clear that obtaining impulse responses from the SVAR is a matter of
  - knowing the matrix $\mathbf{C}_0$
  - estimating the $\mathbf{D}_j$, which can be computed from $\mathbf{A}_i$ which is estimated by OLS
    - For the VAR(1), $\mathbf{D}_j = \mathbf{A}_1^j$
    - For a VAR(k), the computations simply are more involved but $\mathbf{D}_j$ can be determined from the coefficients in the VAR
- A supply and demand example with no exogenous variables

$$q_t = -\alpha p_t + \gamma_{qq} q_{t-1} + \gamma_{qp} p_{t-1} + \varepsilon_t^d$$
$$p_t = \beta q_t + \gamma_{pq} q_{t-1} + \gamma_{pp} p_{t-1} + \varepsilon_t^s$$

Basic setup of VAR for further identification discussion V

- where $\alpha > 0$, the "errors" (or "innovations" or "shocks") have expected values of zero, constant variances of σ_d and σ_s, are serially uncorrelated and are uncorrelated with each other
- In this model, we expect $\varepsilon_t^d > 0$ to have positive effects on p_t and q_t
- In this model, we expect $\varepsilon_t^e > 0$ to have a positive effect on p_t and a negative effect on q_t
  - It is odd, but ε_t^s is an inverse measure of supply shifting right or left
  - $\varepsilon_t^s > 0$ is associated with a shift of the supply curve up in the (p,q) plane, which would commonly be called a decrease
  - So think of it as a shift of the supply curve *up*
- There are no exogenous variables so identification by that route is not feasible
- Could assume a recursive VAR but clearly this will not recover the demand and supply equations above

Sign restriction approach to identification I

- The differential signs of the effects of demand and supply shocks might help to identify these shocks
- Sign restriction: Use the shocks from the VAR e_{it} to estimate $\hat{\varepsilon}_t^d$ and $\hat{\varepsilon}_t^s$
- Estimate the VAR and obtain the coefficients and shocks

$$q_t = a_{qq} q_{t-1} + a_{qp} p_{t-1} + e_{1t}$$
$$p_t = a_{pq} q_{t-1} + a_{pp} p_{t-1} + e_{2t}$$

- Estimate a recursive VAR with $\hat{\mathbf{B}}_0$ lower triangular
  - so current q affects current p
  - but current p does not affect current q, i.e. $\beta = 0$
- This recursive system produces errors for each equation $\hat{v}_{it}$ such that

$$\hat{e}_t = \hat{B}_0^{-1} \hat{v}_t$$

- Will impose that all $\hat{\varepsilon}_t^d$ and $\hat{\varepsilon}_t^s$ are uncorrelated with each other

Sign restriction approach to identification II

- There are a large number of possible combinations of the shocks that will produce uncorrelated $\hat{\varepsilon}$'s
  - In fact, the recursive model is one of them
- It is convenient to work with shocks that have unit variance
- Let $\hat{S}$ be a matrix with the estimated standard deviations of the $\hat{v}_t$ on the diagonals and zero elsewhere
- Then
$$\hat{e}_t = \hat{B}_0^{-1}\hat{S}\hat{S}^{-1}\hat{v}_t$$
- Define
$$\hat{T} = \hat{B}_0^{-1}\hat{S} \qquad \hat{\eta}_t = \hat{S}^{-1}\hat{v}_t$$
where $\hat{\eta}_t$ has unit variance and $\hat{\eta}_{1t}$ and $\hat{\eta}_{2t}$ are uncorrelated by construction
- and T is a transformation matrix from uncorrelated shocks to the correlated shocks in the VAR

Sign restriction approach to identification III

- We can think of these shocks as coming from a base "structural system" which is recursive and is given by

$$T^{-1}y_t = T^{-1}B_1 y_{t-1} + \eta_t$$

- The idea is to find a more plausible set of shocks if these are not plausible
- New shocks can be written as linear functions of the base shocks as

$$\hat{\eta}_t^* = Q\hat{\eta}_t$$

- where Q is a square $n \times n$ matrix

Sign restriction approach to identification IV

- We want the new shocks to be uncorrelated so Q must be restricted

$$\begin{aligned} \mathrm{Var}\left[\hat{\eta}_t^*\right] &= \mathrm{Var}\left[Q\hat{\eta}_t\right] \\ &= \mathrm{E}\left[Q\hat{\eta}_t\hat{\eta}_t' Q'\right] \\ &= Q\,\mathrm{E}\left[\hat{\eta}_t\hat{\eta}_t'\right] Q' \\ &= Q I_n Q' \\ &= I_n \quad \text{if } QQ' = I_n \end{aligned}$$

- The connection with the VAR shocks is

$$\begin{aligned} \hat{e}_t &= \hat{T}\hat{\eta}_t \\ &= \hat{T} Q^{-1} Q \hat{\eta}_t \end{aligned}$$

Sign restriction approach to identification V

- If $Q^{-1} = Q'$ then

$$\begin{aligned} \hat{e}_t &= \hat{T}\hat{\eta}_t \\ &= \hat{T}Q'Q\hat{\eta}_t \\ &= \hat{T}^*\hat{\eta}_t^* \end{aligned}$$

  where $\hat{T}^* = \hat{T}Q'$ and $\hat{\eta}_t^* = Q\hat{\eta}_t$
- So we want to restrict ourselves to transformation matrices such that the transpose of the transformation is its inverse
- Effectively, what we are doing is rotating the vector of data in a two-dimensional plane and keeping, in two dimensions, the axes at a 90-degree angle

Givens transformations I

- In two dimensions, a Givens rotation is

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

- In higher dimensions, there is more than one Givens rotation array
- Choose θ and get a set of shocks
- Check if these shocks are consistent with being demand and supply shocks

Householder transformation I

- Generate an $n \times n$ matrix W, say by using normally distributed innovations
- Take the matrix W generated and apply a QR decomposition

$$W = Q_R R$$

  where Q_R is an orthogonal matrix and R is a triangular matrix
- Search over W's and compute Q_R matrices

Identification by Sign Restriction

- This strategy finds identifications that are consistent with prior about impulse response functions (IRFs)
- Obviously can have an embarrassment of possibilities
- Solution is to present the set of impulse response functions for models consistent with priors about IRF
- Not same as confidence intervals for estimated parameters for a single model
- This is the set of IRFs for models consistent with prior *given* estimated parameters of VAR

Summary of Strategy

- Summarize priors of IRFs given general theoretical framework
- Estimate VAR
- Compute a recursive model with orthogonal innovations
- Sample identified models and select those with IRFs consistent with prior about sign of effect of innovations

Sims-Bernanke

- Basically the idea from Sims (1986) and Bernanke (1986) is to identify parameters in terms of contemporaneous relationships using residuals from VAR
- Identify by restrictions on the residuals across variables

Blanchard-Quah I

- Blanchard-Quah distinguish shocks with long-run effects and those without long-run effects
- The supply and demand example
- A supply and demand example with no exogenous variables

$$
\begin{aligned}
q_t &= -\alpha p_t + \gamma_{qq} q_{t-1} + \gamma_{qp} p_{t-1} + \varepsilon_t^d \\
p_t &= \beta q_t + \gamma_{pq} q_{t-1} + \gamma_{pp} p_{t-1} + \varepsilon_t^s
\end{aligned}
$$

- The reduced form is

$$
\begin{aligned}
q_t &= a_{qq} q_{t-1} + a_{qp} p_{t-1} + e_{1t} \\
p_t &= a_{pq} q_{t-1} + a_{pp} p_{t-1} + e_{2t}
\end{aligned}
$$

- There are six parameters in the structural equations and four in the reduced form

Blanchard-Quah II

- There are two free parameters in the structural variance-covariance matrix and three in the reduced form
- There are eight free parameters in the structural equations and variance-covariance matrix
- There are seven free parameters in the reduced form equations and the variance-covariance matrix
  - Clearly not identified
- Suppose that demand shocks have no permanent effect on the price
- Implies that the supply curve is horizontal
- This implies that $\beta = -\gamma_{pq}$
- The structural equations now are

$$q_t = -\alpha p_t + \gamma_{qq} q_{t-1} + \gamma_{qp} p_{t-1} + \varepsilon_t^d$$
$$p_t = \beta \Delta q_t + \gamma_{pp} p_{t-1} + \varepsilon_t^s$$

- We now have seven total free parameters in the structural equations

Blanchard-Quah III

- We still have seven total parameters in the reduced form
- Estimation strategy
  - Maximum likelihood
  - Instrumental variables or GMM
    - The lagged quantity q_{t-1} is a valid instrument for Δq_t
    - Estimate the supply equation and $\hat{\varepsilon}_t^s$ is a valid instrument for the price
- Blanchard-Quah: Permanent effect of one or more shocks and only transitory effect of others
- Works most obviously with two shocks

Over-identifying restrictions

- Possible to have over-identifying restrictions
- For example, two jointly dependent variables and one excluded exogenous variable provides identification
- Suppose there are four excluded exogenous variables
- Three excluded exogenous variables provide over-identifying restrictions
- Can test these over-identifying restrictions using a likelihood-ratio test

Over-identifying restrictions

- Possible to have over-identifying restrictions
- For example, two jointly dependent variables and one excluded exogenous variable provides identification
- Suppose there are four excluded exogenous variables
- Three excluded exogenous variables provide over-identifying restrictions
- Can test these over-identifying restrictions using a likelihood-ratio test
- *Never* can test identifying restrictions

Bayesian VAR I

- Bayesian VAR with A as the matrix of parameters

$$\begin{gathered} y_t = A_1 y_{t-1} + e_t \\ \mathrm{E}\, e_t = 0, \quad \mathrm{E}\, e_t e_t' = \Omega, \quad \mathrm{E}\, e_t e_s' = 0 \;\forall\; t \neq s \end{gathered}$$

- Let $a = \operatorname{vec} A$ be the vector of parameters
- Apply Bayes rule

$$\mathrm{p}\left(A_1 | y\right) = \frac{\mathrm{p}\left(y | A_1\right) \mathrm{p}\left(A_1\right)}{\mathrm{p}\left(y\right)}$$

- Common to use normal distribution for innovations
- Purpose is to narrow posterior standard deviations when estimating many coefficients, many of which may be close to zero
- A very nice analysis is provided by Helmut Lütkephohl, New Introduction to Multiple Time Series Analysis, pp. 222-29, 309-15

Bayesian VAR II

- Prior and implementation: Litterman (1986), also Doan, Litterman and Sims (1984)
- "Minnesota" prior is a prior mean of one on the first lag of the variable on left-hand side and zero on all other coefficients
  - In other words, start with a prior that the variables are a collection of random walks
- Set up prior variances based on a presumption that longer lags are more likely to be small
- Lütkephohl also summarizes inference starting from a presumption that the variables are stationary

Summary I

- Testing for Granger causality can be a part of specifying a simultaneous-equations model
- Failure of a series to Granger cause other series can be consistent with their exogeneity
- The most common way to identify structural equations historically is by exclusion restrictions
- A recursive system can be an identified VAR with uncorrelated errors
  - This presupposes a particular structure
  - If that structure is not correct, the estimates may have little or nothing to do with the underlying structural equations
- Other ways of identifying equations are
- Sign restrictions
  - Estimate a VAR
  - Start from a recursive VAR based on the estimated VAR
  - Look at models that produce impulse response functions consistent with priors

Summary II

- Sims-Bernanke
  - Impose restrictions on variance-covariance matrix of VAR's errors
- Blanchard-Quah
  - Distinguish shocks that have transitory effects on variables and shocks that have permanent effects
- Over-identifying restrictions
  - If there are more restrictions than are necessary to identify a model, these restrictions can be tested
  - Identifying restrictions identify a model and cannot be tested
- Bayesian VAR
  - A Bayesian VAR uses Bayesian techniques to estimate the VAR
  - The prior can reflect the uncertainty about the coefficients in the model
  - A "Minnesota" prior is a prior that the series are uncorrelated random walks