Multivariate Time Series: Part 5

Identification in VARs (and ECMs), Bayesian VARs

Gerald P. Dwyer

Clemson University

March 2021

Outline

Caveat

- The underlying papers use x and sometimes y to represent the variables in a VAR
- I use y to represent the variables in the VAR
- This change of letters has no significance other than trying to be consistent in notation with sources and failing to be consistent in my presentation
- My apology

Testing for lag length I

- Any estimate of a VAR requires specifying the lag length
- The results of some analyses, including cointegration and Granger causality, can be affected in an important way by the lag length used
- The "statistical significance" of various tests also can be affected by lag length
  - If one is testing whether a set of coefficients is zero and includes many lagged values with zero coefficients, a test can indicate that the set of coefficients is zero even though the nearby lags are not zero
- How to decide?
- Tests are done on the system of equations
- Common to use the same lag length for all equations and all right-hand side variables
  - This can be justified on the basis that there is no reason to choose different lag lengths for different equations
  - Also, there is no reason to choose different lag lengths for the various right-hand side variables

Testing for lag length II

- One could imagine
  - 1 Deleting the coefficient with the smallest "t-ratio" with a p-value less than 5 percent
    - ⋆ Any coefficient? Probably not. Choose among last lag of the variables across all equations
    - ⋆ To be precise, choose $\min(\text{t-ratio}_{i,j,k_{i,j}})$ where $k_{i,j}$ is the last lag in equation i for variable j and the minimization is to pick the one min for all i and all j
  - 2 Re-estimate the system without that coefficient
  - 3 return to step 1 if one or more coefficients has a p-value less than 5 percent; else done
  - This type of test would have a true significance level much higher than 5 percent
  - Don't do this

Testing for lag length III

- Commonly used procedure which has some of the same problems but not as severe:
- Examine whether the coefficients of the last lag of all variables in all equations are zero
- Let the VAR be

$$
\begin{aligned}
y_{1,t} =& \alpha_1 + \sum_{\ell=1}^{k} \beta_{1,1,\ell} y_{1,t-\ell} + \sum_{\ell=1}^{k} \beta_{1,2,\ell} y_{2,t-\ell} + \ldots + \sum_{\ell=1}^{k} \beta_{1,n,\ell} y_{n,t-\ell} + \varepsilon_{1,t} \\
& \ldots \\
y_{n,t} =& \alpha_n + \sum_{\ell=1}^{k} \beta_{n,1,\ell} y_{1,t-\ell} + \sum_{\ell=1}^{k} \beta_{n,2,\ell} y_{2,t-\ell} + \ldots + \sum_{\ell=1}^{k} \beta_{n,n,\ell} y_{n,t-\ell} + \varepsilon_{n,t}
\end{aligned}
$$

- where $y_{i,t}$ is the t'th observation of variable i, $i = 1, \ldots, n$
- α_i is the constant term in the i'th equation
- $\beta_{i,j,\ell}$ is the coefficient in equation i of $y_{j,t-\ell}$, $j = 1, \ldots, n$
- The number of lags of all variables in all equations is k
- Basic procedure: Choose a value of k for all variables in all equations

Ways to choose lag length I

- Criteria to choose lag length
  - Sequential likelihood-ratio test
  - Akaike Information Criterion
  - Schwarz (Bayesian) Criterion
  - Hannon-Quinn Criterion

Likelihood Ratio to choose lag length I

- The equation for the likelihood ratio tests is

$$LR = (T - k - 1) \log \left[\frac{|\Sigma_{\varepsilon,k-1}|}{|\Sigma_{\varepsilon,k}|} \right]$$

  where $|\Sigma_{\varepsilon,k}|$ is the determinant of the variance-covariance matrix of the errors in the VAR with k lags
- Sims suggested the correction for degrees of freedom in the initial term $(T - k - 1)$
- This is asymptotically distributed Chi-square with degrees of freedom equal to the number of coefficients set to zero in the restricted VAR, k^2
- Start with a number of lags that seems implausibly large, say k^* but not so large as to create a small number of degrees of freedom

Likelihood Ratio to choose lag length II

- Test sequentially whether the number of lags can be reduced from k^* to $k^* - 1$, $k^* - 1$ to $k^* - 2$ etc. until the Chi-square is statistically significant at the 5 percent significance level
- This test has an overall p-value well above 5 percent because of the sequential testing, but it is common to use 5 percent anyway

Information Criteria to choose lag length I

- The general equation for the information criteria is

$$C(k, T) = \log|\Sigma_{\varepsilon,k}| + c(T)f(k)$$

- where $C(k, T)$ is the value of the criteria with k lags and T observations
- $c(T)$ is the criteria's specific multiplier for a function $f(k)$
- and $f(k)$ is a penalty term for lag length
- As with univariate autoregressions, the functions are estimated for multiple values with fixed T and then the lowest is picked
- The equations for AIC, Schwarz and Hannon-Quinn are

$$AIC(k, T) = \log|\Sigma_{\varepsilon,k}| + \frac{2}{T}kn^2$$

$$SC(k, T) = \log|\Sigma_{\varepsilon,k}| + \frac{\log T}{T}kn^2$$

$$HQ(k, T) = \log|\Sigma_{\varepsilon,k}| + \frac{2\log\log T}{T}kn^2$$

Information Criteria to choose lag length II

- It is **very** important to use the same observations in all estimates for comparing lag lengths

Granger causality and simultaneous systems

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

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

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

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

- 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 price of beer does not “Granger cause” the weather if the price does not help to predict the weather
  - The price of beer does “Granger cause” the weather if the price helps to predict the weather

Definition of Granger Causality

- *Granger Causality* definition
- A variable x is said to *Granger cause* a variable y if and only if
    - Lagged values of x help to predict y and

Granger causality and simultaneous systems I

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

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

- Reduced form

$$\mathbf{y}_t = \mathbf{\Gamma_0^{-1}} \mathbf{\Gamma}_1 \mathbf{y}_{t-1} + \mathbf{\Gamma_0^{-1}} \boldsymbol{\varepsilon}_t$$

- 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

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

- where

$$\alpha\left(\mathsf{L}\right) = \sum_{i=1}^{k} \alpha_i \, \mathsf{L}^{i-1} \qquad \alpha^*\left(\mathsf{L}\right) = \sum_{i=1}^{k} \alpha^{*i} \, \mathsf{L}^{i-1}$$
$$\gamma\left(\mathsf{L}\right) = \sum_{i=1}^{k} \gamma_i \, \mathsf{L}^{i-1} \qquad \gamma^*\left(\mathsf{L}\right) = \sum_{i=1}^{k} \gamma^{*i} \, \mathsf{L}^{i-1}$$

- Income and weather are determined by

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

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 lagged price and quantity help to predict income and the weather
    - This is the same as a test whether lagged price and quantity do not Granger cause income and the weather
- If lagged price and quantity do not help to predict income and the weather, then it cannot be the case that price and quantity Granger cause income and the weather
- If lagged income and weather help to predict the price and quantity, then income and the weather Granger cause price and quantity

Simultaneous systems and block exogeneity III

- If lagged price and quantity do not help to predict income and the weather and lagged income and weather do help to predict price and quantity, then there exists a structural representation of the equations in which income and weather are exogenous to price and quantity and lagged values of price and quantity do not help to predict income and the weather
- Failure to pass this test does not necessarily imply that income and weather are not exogenous
- Suppose that income and the weather are affected by past prices and quantities but not current 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
- Passing this test does not necessarily imply that income and weather are not exogenous

Simultaneous systems and block exogeneity IV

- 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 lagged price and quantity do not help to predict income and the weather and lagged income and weather do help to predict price and quantity, then there exists a structural representation of the equations in which income and weather are exogenous to price and quantity and ...."
  - This structural representation may not be the structural representation we want

Granger causality and block exogeneity I

- A test of whether these lagged values of price and quantity help to predict income and the weather is a test of Granger causality of income and the weather
- General VAR

$$
\begin{aligned}
q_t &= \Pi_{11}(\mathsf{L})\, q_{t-1} + \Pi_{12}(\mathsf{L})\, p_{t-1} + \Pi_{13}(\mathsf{L})\, y_{t-1} + \Pi_{14}(\mathsf{L})\, w_{t-1} + u_{t1} \\
p_t &= \Pi_{21}(\mathsf{L})\, q_{t-1} + \Pi_{22}(\mathsf{L})\, p_{t-1} + \Pi_{23}(\mathsf{L})\, y_{t-1} + \Pi_{24}(\mathsf{L})\, w_{t-1} + u_{t2} \\
y_t &= \Pi_{31}(\mathsf{L})\, q_{t-1} + \Pi_{32}(\mathsf{L})\, p_{t-1} + \Pi_{33}(\mathsf{L})\, y_{t-1} + \Pi_{34}(\mathsf{L})\, w_{t-1} + u_{t3} \\
w_t &= \Pi_{41}(\mathsf{L})\, q_{t-1} + \Pi_{42}(\mathsf{L})\, p_{t-1} + \Pi_{43}(\mathsf{L})\, y_{t-1} + \Pi_{44}(\mathsf{L})\, w_{t-1} + u_{t4}
\end{aligned}
$$

- Test whether price and quantity help to predict income and weather

Granger causality and block exogeneity II

- Restricted set of equations is

$$
\begin{aligned}
q_t &= \Pi_{11}(\mathsf{L})\, q_{t-1} + \Pi_{12}(\mathsf{L})\, p_{t-1} + \Pi_{13}(\mathsf{L})\, y_{t-1} + \Pi_{14}(\mathsf{L})\, w_{t-1} + u_{t1} \\
p_t &= \Pi_{21}(\mathsf{L})\, q_{t-1} + \Pi_{22}(\mathsf{L})\, p_{t-1} + \Pi_{23}(\mathsf{L})\, y_{t-1} + \Pi_{24}(\mathsf{L})\, w_{t-1} + u_{t2} \\
y_t &= \Pi_{33}(\mathsf{L})\, y_{t-1} + \Pi_{34}(\mathsf{L})\, w_{t-1} + u_{t3} \\
w_t &= \Pi_{43}(\mathsf{L})\, y_{t-1} + \Pi_{44}(\mathsf{L})\, w_{t-1} + u_{t4}
\end{aligned}
$$

- Test

$$
\begin{aligned}
\Pi_{31}(\mathsf{L}) = \Pi_{32}(\mathsf{L}) = 0 \\
\Pi_{41}(\mathsf{L}) = \Pi_{42}(\mathsf{L}) = 0
\end{aligned}
$$

- This is a test whether price and quantity help to predict income and the weather
- Also can say it is a test whether price and quantity “Granger cause” income and the weather

Granger causality and block exogeneity III

- If the coefficients are zero, then price and quantity do not help to predict income and the weather
- If the coefficients are zero, then price and quantity do not "Granger cause" income and the weather
- Can do this with a standard F-test for a set of equations or a likelihood-ratio test for the set of equations

Granger causality and cointegrated systems I

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

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

- Granger causality: A variable $y_{i,t}$ "Granger causes" a variable $y_{j,t}$ in an ECM if
  - lagged values of $y_{i,t}$ appear in the equation for $y_{j,t}$, $j \neq i$ either directly as lagged values or in the cointegrating equations $\boldsymbol{\beta}'\mathbf{y}_{t-1}$
- For a two-variable ECM with constant terms suppressed), $y_{2,t}$ "Granger causes" $y_{1,t}$ if

$$\begin{aligned}
\Delta y_{1,t} &= \alpha_1 \left(y_{1,t-1} - \beta y_{2,t-1} \right) + \gamma_{1,1} \Delta y_{1,t-1} + \gamma_{1,2} \Delta y_{2,t-1} + \varepsilon_{1,t} \\
\Delta y_{2,t} &= \gamma_{2,2} \Delta y_{2,t-1} + \varepsilon_{2,t}
\end{aligned}$$

- $y_{2,t}$ is exogenous relative to $y_{1,t}$ if $\text{Cov}\left[\varepsilon_{1,t}, \varepsilon_{2,t}\right] = 0$ and the innovations are serially uncorrelated

Zero restrictions

- An obvious way to proceed to identify a set of structural equations is to assume exogeneity
- Granger causality tests can be a way to see whether such restrictions are plausible
- If lagged values of a variable y_1 do *not help to predict* a second series y_2, it is *more plausible* that the second series y_2 *is exogenous* relative to the first series y_1
- If lagged values of a variable y_1 *do help to predict* a second series y_2, it is *more plausible* that the second series y_2 *is not exogenous* relative to the first series y_1
  - Basically assuming that the appearance of lagged values of y_1 in an equation affects the plausibility of the current value of y_1 appearing in the equation

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

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

- and assume $\gamma_0 = 0$ and

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

- 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 buy the quantity supplied
  - A vertical supply curve this period, i.e. $\gamma_0 = 0$
- Common to write it as

$$p_t = \alpha_0^{-1}\left[q_t - \sum_{i=1}^{k}\alpha_i p_{t-i} - \sum_{i=1}^{k}\alpha_i^* q_{t-i} - \beta y_t - \varepsilon_t^d\right]$$

$$q_t = \sum_{i=1}^{k}\gamma_i p_{t-i} + \sum_{i=1}^{k}\gamma_i^* q_{t-i} + \delta w_t + \varepsilon_t^s$$

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 **L** is a lower triangular matrix
  - A lower triangular matrix **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

Sign restrictions I

- Basic intuition of sign restriction literature can be illustrated with supply and demand
- Suppose there is a general representation of demand and supply with no exogenous variables

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

- 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
- There are no exogenous variables so identification by that route is not feasible
- Predetermined variables will not help because they appear in both equations

Sign restrictions II

- Could assume a recursive VAR but clearly this will not recover the demand and supply equations above
- We do know the signs of the effects of shocks on the variables
- In this model, we expect $\varepsilon_t^d > 0$ to have positive effects on p_t and q_t
- In this model, we expect $\varepsilon_t^s > 0$ 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*
- Suppose we only look at structures (sets of coefficient values) for this model that have these signs
- That will identify the set of structures consistent with our priors about the signs of the effects of the shocks

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

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

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

$$\mathbf{B}_0 \mathbf{y}_t = \mathbf{B}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$
$$\mathrm{E}\, \boldsymbol{\varepsilon}_t = \mathbf{0}, \quad \mathrm{E}\, \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' = \boldsymbol{\Sigma}, \quad \mathrm{E}\, \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s' = \mathbf{0} \;\forall\; t \neq s$$
$$\mathrm{E}\, \varepsilon_{it} \varepsilon_{jt} = \sigma_i^2 \text{ for } i = j \text{ and } \mathrm{E}\, \varepsilon_{it} \varepsilon_{jt} = 0 \text{ for } i \neq j$$

- 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

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

- $\mathbf{D}_0 = \mathbf{I}$ and therefore

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

  - For a first-order VAR, $\mathbf{D}_j = \mathbf{A}_1^j$

- The moving average representation of the SVAR is given by

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

Basic setup of VAR for further identification discussion III

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

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

and therefore

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

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

$$\mathbf{y}_0 = \mathbf{e}_0$$
$$\mathbf{y}_1 = \mathbf{D}_1 \mathbf{e}_0$$
$$\mathbf{y}_2 = \mathbf{D}_2 \mathbf{e}_0$$
$$\ldots$$

Basic setup of VAR for further identification discussion IV

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

$$\mathbf{y}_0 = \mathbf{C}_0 \varepsilon_0$$
$$\mathbf{y}_1 = \mathbf{C}_1 \varepsilon_0$$
$$\mathbf{y}_2 = \mathbf{C}_2 \varepsilon_0$$
$$\ldots$$

- 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 $\mathbf{C}_0$

Basic setup of VAR for further identification discussion V

- 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

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

- 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 $\varepsilon_t^s > 0$ to have a positive effect on p_t and a negative effect on q_t

Sign restriction approach to identification I

- A supply and demand example with no exogenous variables

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

- 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 $\varepsilon_t^s > 0$ to have a positive effect on p_t and a negative effect on q_t
- The differential signs of the effects of demand and supply shocks might help to identify the structures consistent with our priors
- Sign restriction: Use the shocks from the VAR e_{it} to compute $\hat{\varepsilon}_t^d$ and $\hat{\varepsilon}_t^s$

Sign restriction approach to identification II

- Estimate the VAR and obtain the coefficients and shocks

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

- First iteration: 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
- 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

Sign restriction approach to identification III

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

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

Sign restriction approach to identification IV

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

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

Sign restriction approach to identification V

- The connection with the VAR shocks is

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

- 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 structures 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 structures consistent with priors about IRF
- Not same as confidence intervals for estimated parameters for a single model
- This is the set of IRFs for structures consistent with prior *given* estimated parameters of VAR

Summary of Sign-restriction Strategy

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

Global VARs I

- Want to have models of many countries with interdependencies reflected
- Two problems
  - There could be many, many coefficients
    - Six equations and variables for each country with 4 lags and 20 countries implies $6 \times 4 \times 20 = 480$ coefficients in each regression
    - Presumably most of them would be small and unimportant
    - How to decide what is unimportant?
  - What is exogenous to what?
    - Use standard device of large and small countries
    - United States, Eurozone, China maybe other countries large
    - Panama or Albania small
- Construct a GVAR model for country i
- The vector of k_i variables for country i is $\mathbf{y}_t$ where $\mathbf{y}_t$ is k_i by 1
- Leave off subscript i for simplicity but recall the model is for a particular country i

Global VARs II

- A VAR without variables from the rest of the world, no constant term and one lag would be

$$\mathbf{y}_t = \boldsymbol{\beta}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

- We now want to allow for foreign variables, say k_j of them
- Suppose that these variables are included in the vector $\mathbf{y}_t^*$ where the asterisk denotes foreign variables
- Each variable is an aggregate of the variables for other countries
  - Another common device besides asterisks is to use a superscript f for foreign variables
- Then the VAR can be written

$$\mathbf{y}_t = \boldsymbol{\beta}\mathbf{y}_{t-1} + \gamma_0\mathbf{y}_t^* + \gamma_1\mathbf{y}_{t-1}^* + \boldsymbol{\varepsilon}_t$$

- Obviously the restriction to one lag is solely for simplicity
- The inclusion of the variables for *other* countries as an aggregated entity is not solely for simplicity

Global VARs III

- A variable $y^*_{i,t-1}$ is an aggregated value for the other countries

$$y^*_{i,t-1} = \sum_{j=1, j \neq i}^{N} w_{ij} y^j_{i,t-1}$$

where w_{ij} is a weight applied to each country j for country i
  - The weight typically is based on trade or capital flows between i and j
  - This implies that the variables are different for each country
- The assumption made is that $\mathbf{y}^*_t$ is exogenous to $\mathbf{y}_t$
- In which case, OLS is fine for estimation
- A problem:
  - If one wants to estimate a GVAR model for each of the N countries, the model is internally inconsistent
  - It is impossible for all countries to be exogenous to each and every other country

Global VARs IV

- Variable $y_{i,t}^{j+1}$ is exogenous to $y_{i,t}^{j}$ in $y_{i,t}^{j}$ equation only if $y_{i,t}^{j}$ does not affect $y_{i,t}^{j+1}$ in $y_{i,t}^{j+1}$ equation (Ouliaris Pagan Restrepo 2016, pp. 129-31)
- A GVAR with all countries symmetric is a contradiction

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
- A supply and demand example with no exogenous variables

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

- The reduced form is

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

- There are six parameters in the structural equations and four in the reduced form
- There are two free parameters in the structural variance-covariance matrix and three in the reduced form

Blanchard-Quah II

- 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
- We still have seven total parameters in the reduced form
- Estimation strategy

Blanchard-Quah III

- 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 (BVAR) with $\mathbf{A}$ as the matrix of parameters

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

- Apply Bayes rule

$$\mathrm{p}\left(\mathbf{A}_1 | \mathbf{y}\right) = \frac{\mathrm{p}\left(\mathbf{y} | \mathbf{A}_1\right) \mathrm{p}\left(\mathbf{A}_1\right)}{\mathrm{p}\left(\mathbf{y}\right)}$$

- 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 summary is provided by Helmut Lütkephohl, New Introduction to Multiple Time Series Analysis, pp. 222-29, 309-15
- Prior and implementation: Litterman (1986), also Doan, Litterman and Sims (1984)

Bayesian VAR II

- "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 lag length is a multivariate generalization of determining lag length in an autoregression
- Testing for Granger causality can be a part of specifying a simultaneous-equations model
  - If one series does *not* Granger cause another series, this can be consistent with exogeneity of the second series
- The most common way to identify structural equations historically is by exclusion restrictions
  - i.e. zero restrictions on coefficients of variables
- 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 not be a useful way to look at the underlying structural equations
- Other ways of identifying equations are
- Sign restrictions
  - Estimate a VAR

Summary II

    - Start from a recursive VAR based on the estimated VAR
    - Look at structures that produce impulse response functions consistent with priors
- GVAR
    - A Global VAR is a way of estimating a VAR for countries that are
        - part of the world economy and
        - small relative to the rest of the world
- 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

Summary III

- 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