

Financial Econometrics

Multivariate Time Series Analysis: VAR

Gerald P. Dwyer

Trinity College, Dublin

January 2011

Structural equations

- Suppose have simultaneous system for supply and demand

$$\begin{aligned}q_t &= \alpha^d + \beta^d p_t + \gamma^d S_t + \varepsilon_t^d \\q_t &= \alpha^s + \beta^s p_t + \gamma^s T_t + \varepsilon_t^s\end{aligned}\tag{1}$$

- p is price, q is quantity and S and T are other variables that affect demand and supply
 - Equations jointly determine supply and demand
 - Called “structural equations”
- Suppose solve for “reduced form” with time subscript suppressed

$$\begin{aligned}q &= \pi_{10} + \pi_{11} T + \pi_{12} S + u_1 \\p &= \pi_{20} + \pi_{21} T + \pi_{22} S + u_2\end{aligned}\tag{2}$$

Estimation of structural and reduced form equations

- In general, cannot estimate supply and demand structural equations by Ordinary Least Squares (OLS) because p and q are correlated with error terms in

$$q_t = \alpha^d + \beta^d p_t + \gamma^d S_t + \varepsilon_t^d$$
$$q_t = \alpha^s + \beta^s p_t + \gamma^s T_t + \varepsilon_t^s$$

- Can estimate reduced form regressions by OLS

$$q = \pi_{10} + \pi_{11} T + \pi_{12} S + u_1$$
$$p = \pi_{20} + \pi_{21} T + \pi_{22} S + u_2$$

- Cannot necessarily infer effect of price on quantity demanded from reduced form (although it is possible in this case)
- Endogenous variables (p and q) and exogenous variables (S and T)

Exogenous variables and Endogenous variables

- In economics and finance theory
 - A variable is **exogenous** if it is determined outside the theory
 - For example, income in Ireland in supply and demand for books in Ireland
 - Level of real income in the world in an equation estimating risk factors for BMW's stock return
 - A variable is **endogenous** if it is determined by the theory
 - For example, quantity of books sold in Ireland is determined by supply and demand for books in Ireland
 - Risk premia for BMW stock in context of Capital-asset pricing model or Arbitrage pricing theory
- In econometrics, exogeneity is related to that definition but not the same
 - For many practical econometric purposes, the issue is whether the variable can be included in the right-hand side of an equation and one can ignore how it is determined
 - Generally can get consistent estimates of a set of parameters if the covariance of the variable with the error term is zero

- Suppose have simultaneous system for supply and demand

$$q_t = \alpha^d + \beta^d p_t + \varepsilon_t^d$$

$$q_t = \alpha^s + \beta^s p_t + \varepsilon_t^s$$

- No observable variables besides price and quantity
- Reduced form is

$$q = \pi_{10} + u_1$$

$$p = \pi_{20} + u_2$$

- Two constants and error term
- Cannot possibly infer effect of price on quantity or vice versa (at least without some very strong assumptions about the error terms)

Solutions to identification problem

- Economists have ways of solving the problem
- In fact, the effects of changes in price on quantity supplied and demanded can be estimated from the reduced-form equations sometimes

$$q = \pi_{10} + \pi_{11}T + \pi_{12}S + u_1$$

$$p = \pi_{20} + \pi_{21}T + \pi_{22}S + u_2$$

- For example, in these equations, can infer coefficients in structural equations from the reduced-form equations

All of this can be illustrated with a couple of graphs

- **It's not mysterious**

- Simultaneous equations

$$q_t = \alpha^d + \beta^d p_t + \gamma^d S_t + \varepsilon_t^d$$

$$q_t = \alpha^s + \beta^s p_t + \gamma^s T_t + \varepsilon_t^s$$

- can be estimated by
 - Indirect least squares
 - Use parameters in reduced form to infer coefficients in structural equations above
 - Instrumental variables
 - Find variable correlated with price but not with error term in demand equation (ε_t^d) and not included in demand equation
 - Run reduced-form regression of price on exogenous variables
 - Use predicted value of price in regression for demand equation to get estimated coefficient
 - Similarly supply
 - Two-stage least squares

- Simple autoregression

$$y_t = \phi_0 + \phi_1 y_t + \varepsilon_t$$

First-order vector autoregression

- For a vector? First-order vector autoregression is

$$\mathbf{y}_t = \boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{u}_t$$

where there are k returns in the vector \mathbf{r}_t

$$\begin{aligned}\mathbf{y}'_t &= [y_{1t}, \dots, y_{kt}] \\ \boldsymbol{\phi}'_0 &= [\phi_{10}, \dots, \phi_{k0}] \\ \boldsymbol{\Phi}_1 &= \begin{bmatrix} \phi_{11} & \dots & \phi_{1k} \\ & \dots & \\ \phi_{k1} & & \phi_{kk} \end{bmatrix}, \\ \mathbf{u}'_t &= [u_{1t}, \dots, u_{kt}]\end{aligned}$$

where for ϕ_{ij} , i is equation, j is variable in equation

- Common to put vectors and matrices in bold and I'll do that

Two-variable example

- First-order vector autoregression is

$$\mathbf{y}_t = \boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{u}_t$$

- For two variables, this is

$$y_{1t} = \phi_{10} + \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + u_{1,t}$$

$$y_{2t} = \phi_{20} + \phi_{21}y_{1,t-1} + \phi_{22}y_{2,t-1} + u_{2,t}$$

- Common practice is to use Ordinary Least Squares (OLS) to estimate this set of equations
 - Why?
 - Assume that error terms are not correlated with past values of variables
 - Can be correlated with future values
 - Same variables in every equation so OLS is equivalent to Seemingly Unrelated Regression
 - VAR is like a reduced form
 - No current values of variables on left-hand side of any equation are on the right-hand side of any other equation

- It also is possible to have a second-order autoregression

$$\mathbf{y}_t = \boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{y}_{t-2} + \mathbf{u}_t$$

- or more generally, a k th order autoregression

$$\mathbf{y}_t = \boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{y}_{t-2} + \dots + \boldsymbol{\Phi}_k \mathbf{y}_{t-k} + \mathbf{u}_t$$

Higher-order autoregressions and lag length

- Can write a p -th order autoregression as

$$\mathbf{y}_t = \boldsymbol{\phi}_o + \sum_{i=1}^p \boldsymbol{\Phi}_i \mathbf{y}_{t-i} + \mathbf{a}_t$$

- Lag operator can be handy, $L \mathbf{y}_t = \mathbf{y}_{t-1}$ and $L^i \mathbf{y}_t = \mathbf{y}_{t-i}$
- This implies $\boldsymbol{\Phi}_i \mathbf{y}_{t-i} = \boldsymbol{\Phi}_i L^i \mathbf{y}_t$ and

$$\sum_{i=1}^p \boldsymbol{\Phi}_i \mathbf{y}_{t-i} = \sum_{i=1}^p \boldsymbol{\Phi}_i L^{i-1} \mathbf{y}_t = \boldsymbol{\Phi}(L) \mathbf{y}_{t-1}$$

so

$$\mathbf{y}_t = \boldsymbol{\phi}_o + \boldsymbol{\Phi}(L) \mathbf{y}_{t-1} + \mathbf{a}_t$$

- Higher-order autoregression doesn't add anything of substance, at least formally
 - They do add a lot of parameters

Estimation of higher-order autoregressions

- A fairly common practice in economics is to use the same lag length for all equations for all variables
 - Contributed to Arnold Zellner calling them “Very Awful Regressions”
 - How decide reliably that some coefficients should be zero?
 - Suppose that lags one and four have t-ratios of 3 and lags two and three have t-ratios around one
 - Are the coefficients really zero or just imprecisely estimated?
 - Many extraneous coefficients leads to wide confidence intervals for forecasts
- Determine lag length same way as for univariate autoregressions
 - χ^2 test until five percent “significance”
 - Akaike information criterion
 - (Schwarz) Bayesian information criterion

Bayesian Vector Autoregression

- A Bayesian Vector Autoregression commonly starts from a “Minnesota prior”: The prior supposes that each series is a random walk
 - Data then update this prior
 - Improves precision of estimated coefficients and, to some degree, starts with a simple characterization

Reduced form and structural equations

- Can interpret a VAR as a reduced form of some equation system
 - Two-variable VAR for money and the price level

$$\begin{aligned}m_t &= \phi_{10} + \phi_{11}m_{t-1} + \phi_{12}p_{t-1} + a_{1t} \\ p_t &= \phi_{20} + \phi_{21}m_{t-1} + \phi_{22}p_{t-1} + a_{2t}\end{aligned}\tag{3}$$

- This has no simultaneous determination
- Suppose economic theory suggests

$$\begin{aligned}m_t &= \alpha_m + \beta_{1,11}m_{t-1} + \beta_{1,20}p_t + \beta_{1,21}p_{t-1} + \varepsilon_t^m \\ p_t &= \alpha_p + \beta_{2,10}m_t + \beta_{2,11}m_{t-1} + \beta_{2,21}p_{t-1} + \varepsilon_t^p\end{aligned}$$

- $\{\varepsilon_t^m\}$ and $\{\varepsilon_t^p\}$ are zero mean, constant variance, serially uncorrelated processes
- $\beta_{1,20}$ may reflect effect of prices on the nominal quantity of money demanded
- $\beta_{2,10}$ may reflect effect of supply of money on prices
- Can solve for reduced form and it will look exactly like the VAR reduced form (3)

- Inclusion of contemporary variable in only one of two equations (**recursive system**)

$$m_t = b_{10} + b_{11}m_{t-1} + b^*p_t + b_{12}p_{t-1} + e_t^m$$

$$p_t = \phi_{20} + \phi_{21}m_{t-1} + \phi_{22}p_{t-1} + e_t^p$$

- Estimation of both equations by OLS is consistent if the errors in the two equations are uncorrelated
- Can compute from OLS estimation of VAR using Cholesky decomposition
- How do these coefficients relate to **simultaneous system**

$$m_t = \alpha_m + \beta_{1,11}m_{t-1} + \beta_{1,20}p_t + \beta_{1,21}p_{t-1} + \varepsilon_t^m$$

$$p_t = \alpha_p + \beta_{2,10}m_t + \beta_{2,11}m_{t-1} + \beta_{2,21}p_{t-1} + \varepsilon_t^p$$

or to reduced form **VAR**

$$m_t = \phi_{10} + \phi_{11}m_{t-1} + \phi_{12}p_{t-1} + \varepsilon_t^m$$

$$p_t = \phi_{20} + \phi_{21}m_{t-1} + \phi_{22}p_{t-1} + \varepsilon_t^p$$

Relationship between recursive system, simultaneous system and VAR

- Clearly recursive system, simultaneous system and VAR do not represent the same things
- VAR is a reduced form of simultaneous system
 - Coefficients in VAR are functions of the coefficients in the simultaneous system
 - Error terms in VAR are a combination of the error terms in the simultaneous system
- Recursive system is a version of the VAR in which all of the correlation between ε_t^m and ε_t^p in the VAR is reflected in the coefficient of p_t in the m_t equation in the recursive system
- There is no necessary reason that the coefficients in the simultaneous system should equal the coefficients in the recursive system
 - The coefficients will be equal if the behavioral simultaneous system is recursive
 - The coefficient of m_t in the p_t equation is zero
 - And the errors are uncorrelated

Granger causality tests and VARs

- Relationship of money growth and inflation

$$\begin{aligned}m_t &= \phi_{10} + \phi_{11}m_{t-1} + \phi_{12}p_{t-1} + \varepsilon_t^m \\ p_t &= \phi_{20} + \phi_{21}m_{t-1} + \phi_{22}p_{t-1} + \varepsilon_t^p\end{aligned}\quad (4)$$

- Does money “cause” inflation?
- One version of “cause”: Granger causality
 - Money causes inflation if and only if the equations for money and inflation (4) **cannot** be rewritten

$$\begin{aligned}m_t &= \phi_{10} + \phi_{11}m_{t-1} + \phi_{12}p_{t-1} + \varepsilon_t^m \\ p_t &= \phi_{20} + \phi_{22}p_{t-1} + \varepsilon_t^p\end{aligned}\quad (5)$$

- If money helps to predict inflation, then money “causes” inflation, $\phi_{21} \neq 0$, equations (4)
- If money does not help to predict inflation, then money does “not cause” inflation, $\phi_{21} = 0$, equations (4) reduce to equations (5)

- Use F-tests

$$\begin{aligned}m_t &= \phi_{10} + \phi_{11}m_{t-1} + \phi_{12}p_{t-1} + \varepsilon_t^m \\p_t &= \phi_{20} + \phi_{21}m_{t-1} + \phi_{21,2}m_{t-2} + \phi_{22}p_{t-1} + \varepsilon_t^p\end{aligned}\quad (6)$$

- Test whether both coefficients of lagged money are statistically significant to test whether money helps to predict inflation
 - $\phi_{21}m_{t-1} = \phi_{21,2} = 0$
 - F-test on both coefficients jointly

Granger causality and test on more than one equation at a time

- Joint test across equations on

$$m_t = \phi_{10} + \phi_{11}m_{t-1} + \phi_{12}p_{t-1} + \phi_{13}y_{t-1} + \varepsilon_t^m$$

$$p_t = \phi_{20} + \phi_{21}m_{t-1} + \phi_{22}p_{t-1} + \phi_{23}y_{t-1} + \varepsilon_t^p$$

$$y_t = \phi_{30} + \phi_{31}m_{t-1} + \phi_{32}p_{t-1} + \phi_{33}y_{t-1} + \varepsilon_t^y$$

- Want to test whether money helps to predict p and y
 - Test whether coefficients of lagged money are statistically significant, $\phi_{21} = \phi_{31} = 0$
 - F-test or χ^2 on both coefficients jointly (uses determinant of covariance matrix instead of sum of squared residuals)

- Granger causality is related to exogeneity
 - For many practical econometric purposes, the issue is whether the variable can be included in the right-hand side of an equation and one can ignore how it is determined
 - Generally can ignore variable's determination if the covariance of the variable with the error term is zero

Example of Granger causality and structural equations

- Relationship between money and inflation – Suppose economic theory suggests

$$\begin{aligned}m_t &= \alpha_m + \beta_{1,11}m_{t-1} + \beta_{1,20}p_t + \beta_{1,21}p_{t-1} + \varepsilon_t^m \\p_t &= \alpha_p + \beta_{2,10}m_t + \beta_{2,11}m_{t-1} + \beta_{2,21}p_{t-1} + \varepsilon_t^p\end{aligned}$$

- Two-variable VAR for money and the price level

$$\begin{aligned}m_t &= \phi_{10} + \phi_{11}m_{t-1} + \phi_{12}p_{t-1} + a_{1t} \\p_t &= \phi_{20} + \phi_{21}m_{t-1} + \phi_{22}p_{t-1} + a_{2t}\end{aligned}$$

- If money helps to predict inflation but inflation does not help to predict money – $\phi_{12} = 0$ and money Granger causes inflation
 - Then can write

$$\begin{aligned}m_t &= a_m + b_{1,11}m_{t-1} + \varepsilon_t^m \\p_t &= a_p + b_{2,10}m_t + b_{2,11}m_{t-1} + b_{2,21}p_{t-1} + \varepsilon_t^p\end{aligned}$$

- It is possible to write structural equations with money exogenous relative to inflation

Impulse responses and Variance decompositions

- Impulse responses trace out the effect of a shock to one variable, ε_0^m , in one period, e.g. period 0, on all the variables
- Variance decompositions show how much of the variance of each variable is due to shocks to each variable, e.g. ε_0^m and ε_0^p