

Chapter 7

Multivariate Models

Simultaneous Equations Models

- All the models we have looked at thus far have been single equations models of the form

$$y = X\beta + u$$

- All of the variables contained in the X matrix are assumed to be EXOGENOUS.
- y is an ENDOGENOUS variable.

Simultaneous Equations Models (Cont'd)

- An example from economics to illustrate - the demand and supply of a good:

$$Q_{dt} = \alpha + \beta P_t + \gamma S_t + u_t \quad (1)$$

$$Q_{st} = \lambda + \mu P_t + \kappa T_t + v_t \quad (2)$$

$$Q_{dt} = Q_{st} \quad (3)$$

where

Q_{dt} = quantity of new houses demanded at time t

Q_{st} = quantity of new houses supplied (built) at time t

P_t = (average) price of new houses prevailing at time t

S_t = price of a substitute (e.g. older houses)

T_t = some variable embodying the state of housebuilding technology, u_t and v_t are error terms.

Simultaneous Equations Models: The Structural Form

- Assuming that the market always clears, and dropping the time subscripts for simplicity

$$Q = \alpha + \beta P + \gamma S + u \quad (4)$$

$$Q = \lambda + \mu P + \kappa T + v \quad (5)$$

- This is a simultaneous STRUCTURAL FORM of the model.
- The point is that price and quantity are determined simultaneously (price affects quantity and quantity affects price).
- P and Q are endogenous variables, while S and T are exogenous.
- We can obtain REDUCED FORM equations corresponding to (4) and (5) by solving equations (4) and (5) for P and for Q (separately).

Obtaining the Reduced Form

- Solving for Q , Solving for Q

$$\alpha + \beta P + \gamma S + u = \lambda + \mu P + \kappa T + v \quad (6)$$

- Solving for P ,

$$\frac{Q}{\beta} - \frac{\alpha}{\beta} - \frac{\gamma S}{\beta} - \frac{u}{\beta} = \frac{Q}{\mu} - \frac{\lambda}{\mu} - \frac{\kappa T}{\mu} - \frac{v}{\mu} \quad (7)$$

Obtaining the Reduced Form (Cont'd)

- Rearranging (6),

$$\begin{aligned}\beta P - \mu P &= \lambda - \alpha + \kappa T - \gamma S + v - u \\ (\beta - \mu)P &= (\lambda - \alpha) + \kappa T - \gamma S + (v - u) \\ P &= \frac{\lambda - \alpha}{\beta - \mu} + \frac{\kappa}{\beta - \mu} T - \frac{\gamma}{\beta - \mu} S + \frac{v - u}{\beta - \mu}\end{aligned}\quad (8)$$

- Multiplying (7) through by $\beta\mu$,

$$\begin{aligned}\mu Q - \mu\alpha - \mu\gamma S - \mu u &= \beta Q - \beta\lambda - \beta\kappa T - \beta v \\ \mu Q - \beta Q &= \mu\alpha - \beta\lambda - \beta\kappa T + \mu\gamma S + \mu u - \beta v \\ (\mu - \beta)Q &= (\mu\alpha - \beta\lambda) - \beta\kappa T + \mu\gamma S + (\mu u - \beta v) \\ Q &= \frac{\mu\alpha - \beta\lambda}{\mu - \beta} - \frac{\beta\kappa}{\mu - \beta} T + \frac{\mu\gamma}{\mu - \beta} S + \frac{\mu u - \beta v}{\mu - \beta}\end{aligned}\quad (9)$$

Obtaining the Reduced Form (Cont'd)

- (8) and (9) are the reduced form equations for P and Q .

Simultaneous Equations Bias

- But what would happen if we had estimated equations (4) and (5), i.e. the structural form equations, separately using OLS?
- Both equations depend on P . One of the CLRM assumptions was that $E(X'u) = 0$, where X is a matrix containing all the variables on the RHS of the equation.
- It is clear from (8) that P is related to the errors in (4) and (5) - i.e. it is stochastic.
- What would be the consequences for the OLS estimator, $\hat{\beta}$, if we ignore the simultaneity?
- Recall that $\hat{\beta} = (X'X)^{-1}X'y$ and $y = X\beta + u$

Simultaneous Equations Bias (Cont'd)

- So that

$$\hat{\beta} = (X'X)^{-1}X'(X\beta + u)$$

$$\hat{\beta} = (X'X)^{-1}X'X\beta + (X'X)^{-1}X'u$$

$$\hat{\beta} = \beta + (X'X)^{-1}X'u$$

- Taking expectations,

$$E(\hat{\beta}) = E(\beta) + E((X'X)^{-1}X'u)$$

$$E(\hat{\beta}) = \beta + E((X'X)^{-1}X'u)$$

- If the X 's are non-stochastic, $E(X'u) = 0$, which would be the case in a single equation system, so that $E(\hat{\beta}) = \beta$, which is the condition for unbiasedness.

Simultaneous Equations Bias (Cont'd)

- But if the equation is part of a system, then $E(X'u) \neq 0$, in general.
- Conclusion: Application of OLS to structural equations which are part of a simultaneous system will lead to biased coefficient estimates.
- Is the OLS estimator still consistent, even though it is biased?
- No - In fact the estimator is inconsistent as well.
- Hence it would not be possible to estimate equations (4) and (5) validly using OLS.

Avoiding Simultaneous Equations Bias

So What Can We Do?

- Taking equations (8) and (9), we can rewrite them as

$$P = \pi_{10} + \pi_{11}T + \pi_{12}S + \varepsilon_1 \quad (10)$$

$$Q = \pi_{20} + \pi_{21}T + \pi_{22}S + \varepsilon_2 \quad (11)$$

- We CAN estimate equations (10) & (11) using OLS since all the RHS variables are exogenous.
- But ... we probably don't care what the values of the π coefficients are; what we wanted were the original parameters in the structural equations - $\alpha, \beta, \gamma, \lambda, \mu, \kappa$.

Identification of Simultaneous Equations

Can We Retrieve the Original Coefficients from the π 's?

- Short answer: sometimes.
- As well as simultaneity, we sometimes encounter another problem: identification.
- Consider the following demand and supply equations

$$Q = \alpha + \beta P \quad \text{Supply equation} \quad (12)$$

$$Q = \lambda + \mu P \quad \text{Demand equation} \quad (13)$$

We cannot tell which is which!

- Both equations are UNIDENTIFIED or NOT IDENTIFIED, or UNDERIDENTIFIED.
- The problem is that we do not have enough information from the equations to estimate 4 parameters. Notice that we would not have had this problem with equations (4) and (5) since they have different exogenous variables.

What Determines whether an Equation is Identified or not?

- We could have three possible situations:
 1. An equation is unidentified
 - like (12) or (13)
 - we cannot get the structural coefficients from the reduced form estimates
 2. An equation is exactly identified
 - e.g. (4) or (5)
 - can get unique structural form coefficient estimates
 3. An equation is over-identified
 - Example given later
 - More than one set of structural coefficients could be obtained from the reduced form.
- How do we tell if an equation is identified or not?

What Determines whether an Equation is Identified or not? (Cont'd)

- There are two conditions we could look at:
 - The order condition - is a necessary but not sufficient condition for an equation to be identified.
 - The rank condition - is a necessary and sufficient condition for identification. We specify the structural equations in a matrix form and consider the rank of a coefficient matrix.

Statement of the Order Condition (from Ramanathan 1995, pp.666)

- Let G denote the number of structural equations. An equation is just identified if the number of variables excluded from an equation is $G-1$.
- If more than $G-1$ are absent, it is over-identified. If less than $G-1$ are absent, it is not identified.

What Determines whether an Equation is Identified or not? (Cont'd)

Example

- In the following system of equations, the Y 's are endogenous, while the X 's are exogenous. Determine whether each equation is over-, under-, or just-identified.

$$Y_1 = \alpha_0 + \alpha_1 Y_2 + \alpha_3 Y_3 + \alpha_4 X_1 + \alpha_5 X_2 + u_1 \quad (14)$$

$$Y_2 = \beta_0 + \beta_1 Y_3 + \beta_2 X_1 + u_2 \quad (15)$$

$$Y_3 = \gamma_0 + \gamma_1 Y_2 + u_3 \quad (16)$$

Solution

$$G = 3;$$

What Determines whether an Equation is Identified or not? (Cont'd)

If # excluded variables = 2, the eqn is just identified

If # excluded variables > 2, the eqn is over-identified

If # excluded variables < 2, the eqn is not identified

Equation 14: Not identified

Equation 15: Just identified

Equation 16: Over-identified

Tests for Exogeneity

- How do we tell whether variables really need to be treated as endogenous or not?
- Consider again equations (14)-(16). Equation (14) contains Y_2 and Y_3 - but do we really need equations for them?
- We can formally test this using a Hausman test, which is calculated as follows:
 1. Obtain the reduced form equations corresponding to (14)-(16). The reduced forms turn out to be:

$$Y_1 = \pi_{10} + \pi_{11}X_1 + \pi_{12}X_2 + v_1 \quad (17)$$

$$Y_2 = \pi_{20} + \pi_{21}X_1 + v_2 \quad (18)$$

$$Y_3 = \pi_{30} + \pi_{31}X_1 + v_3 \quad (19)$$

Tests for Exogeneity (Cont'd)

Estimate the reduced form equations (17)-(19) using OLS, and obtain the fitted values, \hat{Y}_1 , \hat{Y}_2 , \hat{Y}_3

2. Run the regression corresponding to equation (14).
3. Run the regression (14) again, but now also including the fitted values \hat{Y}_2^1 , \hat{Y}_3^1 as additional regressors:

$$Y_1 = \alpha_0 + \alpha_1 Y_2 + \alpha_3 Y_3 + \alpha_4 X_1 + \alpha_5 X_2 + \lambda_2 \hat{Y}_2^1 + \lambda_3 \hat{Y}_3^1 + u_1 \quad (20)$$

4. Use an F-test to test the joint restriction that $\lambda_2 = 0$, and $\lambda_3 = 0$. If the null hypothesis is rejected, Y_2 and Y_3 should be treated as endogenous.

Recursive Systems

- Consider the following system of equations:

$$Y_1 = \beta_{10} + \gamma_{11}X_1 + \gamma_{12}X_2 + u_1 \quad (21)$$

$$Y_2 = \beta_{20} + \beta_{21}Y_1 + \gamma_{21}X_1 + \gamma_{22}X_2 + u_2 \quad (22)$$

$$Y_3 = \beta_{30} + \beta_{31}Y_1 + \beta_{32}Y_2 + \gamma_{31}X_1 + \gamma_{32}X_2 + u_3 \quad (23)$$

- Assume that the error terms are not correlated with each other. Can we estimate the equations individually using OLS?
- Equation 21: Contains no endogenous variables, so X_1 and X_2 are not correlated with u_1 . So we can use OLS on (21).

Recursive Systems (Cont'd)

- Equation 22: Contains endogenous Y_1 together with exogenous X_1 and X_2 . We can use OLS on (22) if all the RHS variables in (22) are uncorrelated with that equation's error term. In fact, Y_1 is not correlated with u_2 because there is no Y_2 term in equation (21). So we can use OLS on (22).
- Equation 23: Contains both Y_1 and Y_2 ; we require these to be uncorrelated with u_3 . By similar arguments to the above, equations (21) and (22) do not contain Y_3 , so we can use OLS on (23).
- This is known as a RECURSIVE or TRIANGULAR system. We do not have a simultaneity problem here.
- But in practice not many systems of equations will be recursive...

Indirect Least Squares (ILS)

- Cannot use OLS on structural equations, but we can validly apply it to the reduced form equations.
- If the system is just identified, ILS involves estimating the reduced form equations using OLS, and then using them to substitute back to obtain the structural parameters.
- However, ILS is not used much because
 1. Solving back to get the structural parameters can be tedious.
 2. Most simultaneous equations systems are over-identified.

Estimation of Systems Using Two-Stage Least Squares

- In fact, we can use this technique for just-identified and over-identified systems.
- Two stage least squares (2SLS or TSLS) is done in two stages:

Stage 1:

- Obtain and estimate the reduced form equations using OLS. Save the fitted values for the dependent variables.

Stage 2:

- Estimate the structural equations, but replace any RHS endogenous variables with their stage 1 fitted values.

Estimation of Systems Using Two-Stage Least Squares (Cont'd)

Example: Say equations (14)-(16) are required.

Stage 1:

- Estimate the reduced form equations (17)-(19) individually by OLS and obtain the fitted values, \hat{y}_i .

Stage 2:

Estimation of Systems Using Two-Stage Least Squares (Cont'd)

- Replace the RHS endogenous variables with their stage 1 estimated values:

$$Y_1 = \alpha_0 + \alpha_1 \hat{Y}_2^1 + \alpha_3 \hat{Y}_3^1 + \alpha_4 X_1 + \alpha_5 X_2 + u_1 \quad (24)$$

$$Y_2 = \beta_0 + \beta_1 \hat{Y}_3^1 + \beta_2 X_1 + u_2 \quad (25)$$

$$Y_3 = \gamma_0 + \gamma_1 \hat{Y}_2^1 + u_3 \quad (26)$$

- Now \hat{Y}_2^1 and \hat{Y}_3^1 will not be correlated with u_1 , \hat{Y}_3^1 will not be correlated with u_2 , and \hat{Y}_2^1 will not be correlated with u_3 .
- It is still of concern in the context of simultaneous systems whether the CLRM assumptions are supported by the data.

Estimation of Systems Using Two-Stage Least Squares (Cont'd)

- If the disturbances in the structural equations are autocorrelated, the 2SLS estimator is not even consistent.
- The standard error estimates also need to be modified compared with their OLS counterparts, but once this has been done, we can use the usual t - and F -tests to test hypotheses about the structural form coefficients.

Instrumental Variables

- Recall that the reason we cannot use OLS directly on the structural equations is that the endogenous variables are correlated with the errors.
- One solution to this would be not to use Y_2 or Y_3 , but rather to use some other variables instead.
- We want these other variables to be (highly) correlated with Y_2 and Y_3 , but not correlated with the errors - they are called INSTRUMENTS.

Instrumental Variables (Cont'd)

- Say we found suitable instruments for Y_2 and Y_3 , z_2 and z_3 respectively. We do not use the instruments directly, but run regressions of the form

$$Y_2 = \lambda_1 + \lambda_2 z_2 + \varepsilon_1 \quad (27)$$

$$Y_3 = \lambda_3 + \lambda_4 z_3 + \varepsilon_2 \quad (28)$$

- Obtain the fitted values from (27) & (28), \hat{Y}_2^1 and \hat{Y}_3^1 , and replace Y_2 and Y_3 with these in the structural equation.
- We do not use the instruments directly in the structural equation.
- It is typical to use more than one instrument per endogenous variable.

Instrumental Variables (Cont'd)

- If the instruments are the variables in the reduced form equations, then IV is equivalent to 2SLS.

What Happens if We Use IV / 2SLS Unnecessarily?

- The coefficient estimates will still be consistent, but will be inefficient compared to those that just used OLS directly.

The Problem With IV

- What are the instruments?
Solution: 2SLS is easier.

Other Estimation Techniques

1. 3SLS - allows for non-zero covariances between the error terms.

Instrumental Variables (Cont'd)

2. LIML - estimating reduced form equations by maximum likelihood
3. FIML - estimating all the equations simultaneously using maximum likelihood

An Example of the Use of 2SLS: Modelling the Bid-Ask Spread and Volume for Options

- George and Longstaff (1993)
- Introduction
 - Is trading activity related to the size of the bid / ask spread?
 - How do spreads vary across options?
- How Might the Option Price / Trading Volume and the Bid / Ask Spread be Related?

Consider 3 possibilities:

1. Market makers equalise spreads across options.
2. The spread might be a constant proportion of the option value.
3. Market makers might equalise marginal costs across options irrespective of trading volume.

Market Making Costs

- The S&P 100 Index has been traded on the CBOE since 1983 on a continuous open-outcry auction basis.
- Transactions take place at the highest bid or the lowest ask.
- Market making is highly competitive.

What Are the Costs Associated with Market Making?

- For every contract (100 options) traded, a CBOE fee of 9c and an Options Clearing Corporation (OCC) fee of 10c is levied on the firm that clears the trade.
- Trading is not continuous.
- Average time between trades in 1989 was approximately 5 minutes.

The Influence of Tick-Size Rules on Spreads

- The CBOE limits the tick size:
 - \$1/8 for options worth \$3 or more
 - \$1/16 for options worth less than \$3
- The spread is likely to depend on trading volume
... but also trading volume is likely to depend on the spread.
- So there will be a simultaneous relationship.

The Data

- All trading days during 1989 are used for observations.
- The average bid & ask prices are calculated for each option during the time 2:00pm – 2:15pm Central Standard time.
- The following are then dropped from the sample for that day:
 1. Any options that do not have bid / ask quotes reported during the 1/4 hour.
 2. Any options with fewer than 10 trades during the day.
- The option price is defined as the average of the bid & the ask.
- We get a total of 2456 observations. This is a pooled regression.

The Models

- For the calls:

$$CBA_i = \alpha_0 + \alpha_1 CDUM_i + \alpha_2 C_i + \alpha_3 CL_i + \alpha_4 T_i + \alpha_5 CR_i + e_i \quad (29)$$

$$CL_i = \gamma_0 + \gamma_1 CBA_i + \gamma_2 T_i + \gamma_3 T_i^2 + \gamma_4 M_i^2 + v_i \quad (30)$$

- And symmetrically for the puts:

$$PBA_i = \beta_0 + \beta_1 PDUM_i + \beta_2 P_i + \beta_3 PL_i + \beta_4 T_i + \beta_5 PR_i + u_i \quad (31)$$

$$PL_i = \delta_0 + \delta_1 PBA_i + \delta_2 T_i + \delta_3 T_i^2 + \delta_4 M_i^2 + w_i \quad (32)$$

The Models (Cont'd)

- where ΔCR_i and ΔPR_i are the squared deltas of the options
- $CDUM_i$ and $PDUM_i$ are dummy variables to allow for the minimum tick size

$$= 0 \quad \text{if } C_i \text{ or } P_i < \$3$$

$$= 1 \quad \text{if } C_i \text{ or } P_i \geq \$3$$

- T^2 allows for a nonlinear relationship between time to maturity and the spread.
- M^2 is used since ATM options have a higher trading volume.
- Aside: are the equations identified?
- Equations (29) & (30) and then separately (31) & (32) are estimated using 2SLS.

Results 1

Call bid–ask spread and trading volume regression

$$CBA_i = \alpha_0 + \alpha_1 CDUM_i + \alpha_2 C_i + \alpha_3 CL_i + \alpha_4 T_i + \alpha_5 CR_i + e_i$$

$$CL_i = \gamma_0 + \gamma_1 CBA_i + \gamma_2 T_i + \gamma_3 T_i^2 + \gamma_4 M_i^2 + v_i$$

α_0	α_1	α_2	α_3	α_4	α_5	Adj. R^2
0.08362 (16.80)	0.06114 (8.63)	0.01679 (15.49)	0.00902 (14.01)	-0.00228 (-12.31)	-0.15378 (-12.52)	0.688
γ_0	γ_1	γ_2	γ_3	γ_4		Adj. R^2
-3.8542 (-10.50)	46.592 (30.49)	-0.12412 (-6.01)	0.00406 (14.43)	0.00866 (4.76)		0.618

Note: *t*-ratios in parentheses.

Source: George and Longstaff (1993). Reprinted with the permission of School of Business Administration, University of Washington.

Results 2

Put bid–ask spread and trading volume regression

$$PBA_i = \beta_0 + \beta_1 PDUM_i + \beta_2 P_i + \beta_3 PL_i + \beta_4 T_i + \beta_5 PR_i + u_i$$

$$PL_i = \delta_0 + \delta_1 PBA_i + \delta_2 T_i + \delta_3 T_i^2 + \delta_4 M_i^2 + w_i$$

β_0	β_1	β_2	β_3	β_4	β_5	Adj. R^2
0.05707 (15.19)	0.03258 (5.35)	0.01726 (15.90)	0.00839 (12.56)	-0.00120 (-7.13)	-0.08662 (-7.15)	0.675
δ_0	δ_1	δ_2	δ_3	δ_4		Adj. R^2
-2.8932 (-8.42)	46.460 (34.06)	-0.15151 (-7.74)	0.00339 (12.90)	0.01347 (10.86)		0.517

Note: t -ratios in parentheses.

Source: George and Longstaff (1993). Reprinted with the permission of School of Business Administration, University of Washington.

Comments:

Adjusted $R^2 \approx 0.6$

α_1 and β_1 measure the tick size constraint on the spread

α_2 and β_2 measure the effect of the option price on the spread

α_3 and β_3 measure the effect of trading activity on the spread

α_4 and β_4 measure the effect of time to maturity on the spread

α_5 and β_5 measure the effect of risk on the spread

γ_1 and δ_1 measure the effect of the spread size on trading activity etc.

Calls and Puts as Substitutes

- The paper argues that calls and puts might be viewed as substitutes since they are all written on the same underlying.
- So call trading activity might depend on the put spread and put trading activity might depend on the call spread.
- The results for the other variables are little changed.

Conclusions

- Bid - Ask spread variations between options can be explained by reference to the level of trading activity, deltas, time to maturity etc. There is a 2 way relationship between volume and the spread.
- The authors argue that in the second part of the paper, they did indeed find evidence of substitutability between calls & puts.

Comments

- No diagnostics.
- Why do the CL and PL equations not contain the CR and PR variables?

Conclusions (Cont'd)

- The authors could have tested for endogeneity of CBA and CL.
- Why are the squared terms in maturity and moneyness only in the liquidity regressions?
- Wrong sign on the squared deltas.

Vector Autoregressive Models

- A natural generalisation of autoregressive models popularised by Sims
- A VAR is in a sense a systems regression model i.e. there is more than one dependent variable.
- Simplest case is a bivariate VAR

$$y_{1t} = \beta_{10} + \beta_{11}y_{1t-1} + \cdots + \beta_{1k}y_{1t-k} + \alpha_{11}y_{2t-1} + \cdots + \alpha_{1k}y_{2t-k} + u_{1t}$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{2t-1} + \cdots + \beta_{2k}y_{2t-k} + \alpha_{21}y_{1t-1} + \cdots + \alpha_{2k}y_{1t-k} + u_{2t}$$

Vector Autoregressive Models (Cont'd)

where u_{it} is a white noise disturbance term with $E(u_{it}) = 0$, $(i = 1, 2)$, $E(u_{1t}u_{2t}) = 0$.

- The analysis could be extended to a VAR(g) model, or so that there are g variables and g equations.

Vector Autoregressive Models: Notation and Concepts

- One important feature of VARs is the compactness with which we can write the notation. For example, consider the case from above where $k=1$.
- We can write this as

$$y_{1t} = \beta_{10} + \beta_{11}y_{1t-1} + \alpha_{11}y_{2t-1} + u_{1t}$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{2t-1} + \alpha_{21}y_{1t-1} + u_{2t}$$

- or

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

Vector Autoregressive Models: Notation and Concepts (Cont'd)

- or even more compactly as

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t$$

$g \times 1 \quad g \times 1 \quad g \times g \quad g \times 1 \quad g \times 1$

- This model can be extended to the case where there are k lags of each variable in each equation:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_k y_{t-k} + u_t$$

$g \times 1 \quad g \times 1 \quad g \times g \quad g \times 1 \quad g \times g \quad g \times 1 \quad g \times 1$

- We can also extend this to the case where the model includes first difference terms and cointegrating relationships (a VECM).

Vector Autoregressive Models Compared with Structural Equations Models

- Advantages of VAR Modelling
 - Do not need to specify which variables are endogenous or exogenous - all are endogenous
 - Allows the value of a variable to depend on more than just its own lags or combinations of white noise terms, so more general than ARMA modelling
 - Provided that there are no contemporaneous terms on the right hand side of the equations, can simply use OLS separately on each equation
 - Forecasts are often better than “traditional structural” models.
- Problems with VAR's
 - VAR's are a-theoretical (as are ARMA models)

Vector Autoregressive Models Compared with Structural Equations Models (Cont'd)

- How do you decide the appropriate lag length?
- So many parameters! If we have g equations for g variables and we have k lags of each of the variables in each equation, we have to estimate $(g + kg^2)$ parameters. e.g. $g=3$, $k=3$, parameters=30
- Do we need to ensure all components of the VAR are stationary?
- How do we interpret the coefficients?

Choosing the Optimal Lag Length for a VAR

- 2 possible approaches: cross-equation restrictions and information criteria

Cross-Equation Restrictions

- In the spirit of (unrestricted) VAR modelling, each equation should have the same lag length
- Suppose that a bivariate VAR(8) estimated using quarterly data has 8 lags of the two variables in each equation, and we want to examine a restriction that the coefficients on lags 5 through 8 are jointly zero. This can be done using a likelihood ratio test

Choosing the Optimal Lag Length for a VAR (Cont'd)

- Denote the variance-covariance matrix of residuals (given by $\hat{u}\hat{u}'/T$), as $\hat{\Sigma}$. The likelihood ratio test for this joint hypothesis is given by

$$LR = T[\log|\hat{\Sigma}_r| - \log|\hat{\Sigma}_u|]$$

where where $|\hat{\Sigma}_r|$ is the variance-covariance matrix of the residuals for the restricted model (with 4 lags), where $|\hat{\Sigma}_u|$ is the variance-covariance matrix of residuals for the unrestricted VAR (with 8 lags), and T is the sample size.

Choosing the Optimal Lag Length for a VAR (Cont'd)

- The test statistic is asymptotically distributed as a χ^2 with degrees of freedom equal to the total number of restrictions. In the VAR case above, we are restricting 4 lags of two variables in each of the two equations = a total of $4 \times 2 \times 2 = 16$ restrictions.
- In the general case where we have a VAR with p equations, and we want to impose the restriction that the last q lags have zero coefficients, there would be $p^2 q$ restrictions altogether
- Disadvantages: Conducting the LR test is cumbersome and requires a normality assumption for the disturbances.

Information Criteria for VAR Lag Length Selection

- Multivariate versions of the information criteria are required. These can be defined as:

$$MAIC = \log|\hat{\Sigma}| + 2k'/T$$

$$MSBIC = \log|\hat{\Sigma}| + \frac{k'}{T} \log(T)$$

$$MHQIC = \log|\hat{\Sigma}| + \frac{2k'}{T} \log(\log(T))$$

where all notation is as above and k is the total number of regressors in all equations, which will be equal to g^2k+g for g equations, each with k lags of the g variables, plus a constant term in each equation. The values of the information criteria are constructed for 0, 1, ... lags (up to some pre-specified maximum \bar{k}).

Does the VAR Include Contemporaneous Terms?

- So far, we have assumed the VAR is of the form

$$y_{1t} = \beta_{10} + \beta_{11}y_{1t-1} + \alpha_{11}y_{2t-1} + u_{1t}$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{2t-1} + \alpha_{21}y_{1t-1} + u_{2t}$$

- But what if the equations had a contemporaneous feedback term?

$$y_{1t} = \beta_{10} + \beta_{11}y_{1t-1} + \alpha_{11}y_{2t-1} + \alpha_{12}y_{2t} + u_{1t}$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{2t-1} + \alpha_{21}y_{1t-1} + \alpha_{22}y_{1t} + u_{2t}$$

Does the VAR Include Contemporaneous Terms? (Cont'd)

- We can write this as

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} \\ + \begin{pmatrix} \alpha_{12} & 0 \\ 0 & \alpha_{22} \end{pmatrix} \begin{pmatrix} y_{2t} \\ y_{1t} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

- This VAR is in primitive form.

Primitive versus Standard Form VARs

- We can take the contemporaneous terms over to the LHS and write

$$\begin{pmatrix} 1 & -\alpha_{12} \\ -\alpha_{22} & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix}$$

or

$$Ay_t = \beta_0 + \beta_1 y_{t-1} + u_t$$

- If both sides are pre-multiplied by A^{-1}

$$y_t = A^{-1}\beta_0 + A^{-1}\beta_1 y_{t-1} + A^{-1}u_t$$

or

$$y_t = A_0 + A_1 y_{t-1} + e_t$$

- This is known as a standard form VAR, which we can estimate using OLS.

Block Significance and Causality Tests

- It is likely that, when a VAR includes many lags of variables, it will be difficult to see which sets of variables have significant effects on each dependent variable and which do not. For illustration, consider the following bivariate VAR(3):

$$\begin{aligned} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} &= \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} \\ &+ \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} y_{1t-2} \\ y_{2t-2} \end{pmatrix} \\ &+ \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} \begin{pmatrix} y_{1t-3} \\ y_{2t-3} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \end{aligned}$$

Block Significance and Causality Tests (Cont'd)

- This VAR could be written out to express the individual equations as

$$y_{1t} = \alpha_{10} + \beta_{11}y_{1t-1} + \beta_{12}y_{2t-1} + \gamma_{11}y_{1t-2} + \gamma_{12}y_{2t-2} \\ + \delta_{11}y_{1t-3} + \delta_{12}y_{2t-3} + u_{1t}$$

$$y_{2t} = \alpha_{20} + \beta_{21}y_{1t-1} + \beta_{22}y_{2t-1} + \gamma_{21}y_{1t-2} + \gamma_{22}y_{2t-2} \\ + \delta_{21}y_{1t-3} + \delta_{22}y_{2t-3} + u_{2t}$$

- We might be interested in testing the following hypotheses, and their implied restrictions on the parameter matrices:

	Hypothesis	Implied restriction
1	Lags of y_{1t} do not explain current y_{2t}	$\beta_{21} = 0$ and $\gamma_{21} = 0$ and $\delta_{21} = 0$
2	Lags of y_{1t} do not explain current y_{1t}	$\beta_{11} = 0$ and $\gamma_{11} = 0$ and $\delta_{11} = 0$
3	Lags of y_{2t} do not explain current y_{1t}	$\beta_{12} = 0$ and $\gamma_{12} = 0$ and $\delta_{12} = 0$
4	Lags of y_{2t} do not explain current y_{2t}	$\beta_{22} = 0$ and $\gamma_{22} = 0$ and $\delta_{22} = 0$

Block Significance and Causality Tests (Cont'd)

- Each of these four joint hypotheses can be tested within the F -test framework, since each set of restrictions contains only parameters drawn from one equation.
- These tests could also be referred to as Granger causality tests.
- Granger causality tests seek to answer questions such as “Do changes in y_1 cause changes in y_2 ?” If y_1 causes y_2 , lags of y_1 should be significant in the equation for y_2 . If this is the case, we say that y_1 “Granger-causes” y_2 .
- If y_2 causes y_1 , lags of y_2 should be significant in the equation for y_1 .
- If both sets of lags are significant, there is “bi-directional causality”

Impulse Responses

- VAR models are often difficult to interpret: one solution is to construct the impulse responses and variance decompositions.
- Impulse responses trace out the responsiveness of the dependent variables in the VAR to shocks to the error term. 'A unit shock is applied to each variable and its effects are noted.'
- Consider for example a simple bivariate VAR(1):

$$y_{1t} = \beta_{10} + \beta_{11}y_{1t-1} + \alpha_{11}y_{2t-1} + u_{1t}$$

$$y_{2t} = \beta_{20} + \beta_{21}y_{2t-1} + \alpha_{21}y_{1t-1} + u_{2t}$$

- A change in u_{1t} will immediately change y_1 . It will change y_2 and also y_1 during the next period.
- We can examine how long and to what degree a shock to a given equation has on all of the variables in the system.

Variance Decompositions

- Variance decompositions offer a slightly different method of examining VAR dynamics. They give the proportion of the movements in the dependent variables that are due to their “own” shocks, versus shocks to the other variables.
- This is done by determining how much of the s -step ahead forecast error variance for each variable is explained innovations to each explanatory variable ($s= 1,2,\dots$).
- The variance decomposition gives information about the relative importance of each shock to the variables in the VAR.

Impulse Responses and Variance Decompositions: The Ordering of the Variables

- But for calculating impulse responses and variance decompositions, the ordering of the variables is important.
- The main reason for this is that above, we assumed that the VAR error terms were statistically independent of one another.
- This is generally not true, however. The error terms will typically be correlated to some degree.
- Therefore, the notion of examining the effect of the innovations separately has little meaning, since they have a common component.

Impulse Responses and Variance Decompositions: The Ordering of the Variables (Cont'd)

- What is done is to “orthogonalise” the innovations.
- In the bivariate VAR, this problem would be approached by attributing all of the effect of the common component to the first of the two variables in the VAR.
- In the general case where there are more variables, the situation is more complex but the interpretation is the same.

An Example of the use of VAR Models: The Interaction between Property Returns and the Macroeconomy.

- Brooks and Tsolacos (1999) employ a VAR methodology for investigating the interaction between the UK property market and various macroeconomic variables.
- Monthly data are used for the period December 1985 to January 1998.
- It is assumed that stock returns are related to macroeconomic and business conditions.
- The variables included in the VAR are
 - FTSE Property Total Return Index (with general stock market effects removed)

An Example of the use of VAR Models: The Interaction between Property Returns and the Macroeconomy. (Cont'd)

- The rate of unemployment
- Nominal interest rates
- The spread between long and short term interest rates
- Unanticipated inflation
- The dividend yield.

The property index and unemployment are $I(1)$ and hence are differenced.

Marginal Significance Levels associated with Joint F -tests that all 14 Lags have not Explanatory Power for that particular Equation in the VAR

- Multivariate AIC selected 14 lags of each variable in the VAR

Dependent variable	Lags of variable					
	SIR	DIVY	SPREAD	UNEM	UNINFL	PROPRES
SIR	0.0000	0.0091	0.0242	0.0327	0.2126	0.0000
DIVY	0.5025	0.0000	0.6212	0.4217	0.5654	0.4033
SPREAD	0.2779	0.1328	0.0000	0.4372	0.6563	0.0007
UNEM	0.3410	0.3026	0.1151	0.0000	0.0758	0.2765
UNINFL	0.3057	0.5146	0.3420	0.4793	0.0004	0.3885
PROPRES	0.5537	0.1614	0.5537	0.8922	0.7222	0.0000

The test is that all 14 lags have no explanatory power for that particular equation in the VAR.

Source: Brooks and Tsolacos (1999).

Variance Decompositions for the Property Sector Index Residuals

- Ordering for Variance Decompositions and Impulse Responses:
 - Order I: PROPRES, DIVY, UNINFL, UNEM, SPREAD, SIR
 - Order II: SIR, SPREAD, UNEM, UNINFL, DIVY, PROPRES.

Months ahead	<i>Explained by innovations in</i>											
	<u>SIR</u>		<u>DIVY</u>		<u>SPREAD</u>		<u>UNEM</u>		<u>UNINFL</u>		<u>PROPRES</u>	
	I	II	I	II	I	II	I	II	I	II	I	II
1	0.0	0.8	0.0	38.2	0.0	9.1	0.0	0.7	0.0	0.2	100.0	51.0
2	0.2	0.8	0.2	35.1	0.2	12.3	0.4	1.4	1.6	2.9	97.5	47.5
3	3.8	2.5	0.4	29.4	0.2	17.8	1.0	1.5	2.3	3.0	92.3	45.8
4	3.7	2.1	5.3	22.3	1.4	18.5	1.6	1.1	4.8	4.4	83.3	51.5
12	2.8	3.1	15.5	8.7	15.3	19.5	3.3	5.1	17.0	13.5	46.1	50.0
24	8.2	6.3	6.8	3.9	38.0	36.2	5.5	14.7	18.1	16.9	23.4	22.0

Source: Brooks and Tsolacos (1999).

Impulse Responses and Standard Error Bands for Innovations in Dividend Yield and Unexpected Inflation

