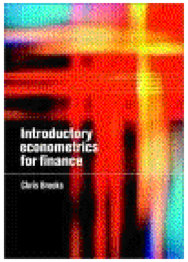
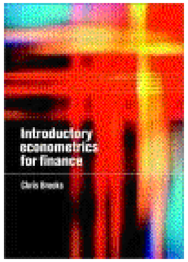


Chapter 6

Multivariate models



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.

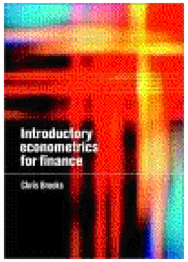
- Simplest case is a bivariate VAR

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

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

where u_{it} is an iid 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

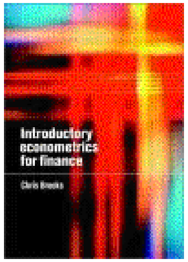
- One important feature of VARs is the compactness with which we can write the notation. For example, consider the case from above where k=1.

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

or even more compactly as

$$\underset{g\times 1}{y_t} = \underset{g\times 1}{\beta_0} + \underset{g\times g}{\beta_1}\ \underset{g\times 1}{y_{t\text{-}1}} + \underset{g\times 1}{u_t}$$

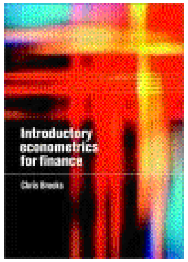


Vector Autoregressive Models: Notation and Concepts (cont'd)

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

$$
\begin{array}{cccccccccccccc}
y_t & = & \beta_0 & + & \beta_1 & y_{t-1} & + & \beta_2 & y_{t-2} & +\ldots+ & \beta_k & y_{t-k} & + & u_t \\
g\times 1 & & g\times 1 & & g\times g & g\times 1 & & g\times g & g\times 1 & & g\times g & g\times 1 & & g\times 1
\end{array}
$$

- 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

- <u>Advantages of VAR Modelling</u>

  - 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.

- <u>Problems with VAR's</u>

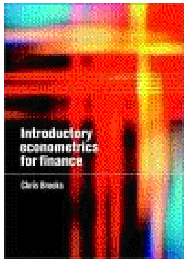
  - VAR's are a-theoretical (as are ARMA models)

  - 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?



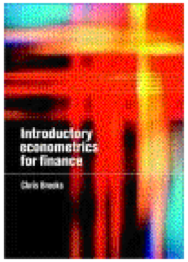
Choosing the Optimal Lag Length for a VAR

- 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
- 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

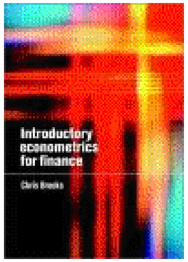
$$LR = T\left[\log\left|\hat{\Sigma}_r\right| - \log\left|\hat{\Sigma}_u\right|\right]$$



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

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

- 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 * 2 * 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^2q restrictions altogether
- Disadvantages: Conducting the LR test is cumbersome and requires a normality assumption for the disturbances.



Information Criteria for VAR Lag Length Selection

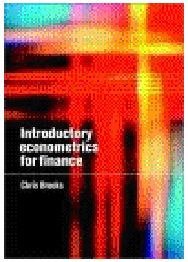
- Multivariate versions of the information criteria are required. These can be defined as:

$$MAIC = \ln\left|\hat{\Sigma}\right| + 2k'/T$$

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

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

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

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

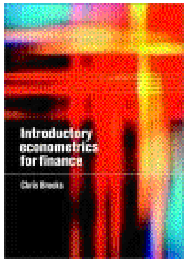
- But what if the equations had a contemporaneous feedback term?

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

- 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

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

or

$$\mathrm{B}\, y_t = \beta_0 + \beta_1\, y_{t\text{-}1} + u_t$$

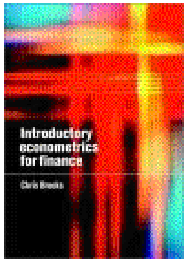
- We can then pre-multiply both sides by B^{-1} to give

$$y_t = \mathrm{B}^{-1}\beta_0 + \mathrm{B}^{-1}\beta_1\, y_{t\text{-}1} + \mathrm{B}^{-1}u_t$$

or

$$y_t = \mathrm{A}_0 + \mathrm{A}_1\, y_{t\text{-}1} + e_t$$

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



Block Significance and Causality Tests

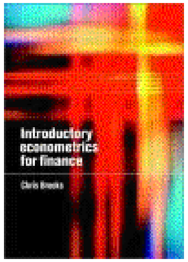
- It is likely that, when a VAR includes many lags of variables, it will be difficult to see which sets of variables have significant effects on each dependent variable and which do not. For illustration, consider the following bivariate VAR(3):

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

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

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

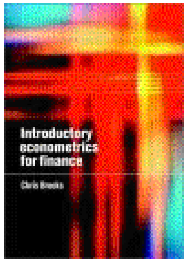
- We might be interested in testing the following hypotheses, and their implied restrictions on the parameter matrices:



Block Significance and Causality Tests (cont'd)

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 $\delta_{22} = 0$

- 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"

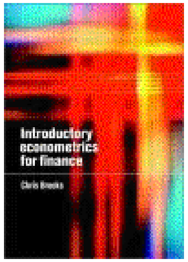


Impulse Responses

- 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 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

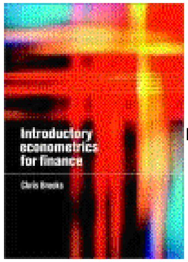
- Variance decompositions offer a slightly different method of examining VAR dynamics. They give the proportion of the movements in the dependent variables that are due to their “own” shocks, versus shocks to the other variables.

- 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,…).

- 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

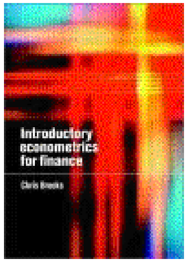
- But for calculating impulse responses and variance decompositions, the ordering of the variables is important.
- The main reason for this is that above, we assumed that the VAR error terms were statistically independent of one another.
- This is generally not true, however. The error terms will typically 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.
- 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)
  - The rate of unemployment
  - Nominal interest rates
  - The spread between long and short term interest rates
  - Unanticipated inflation
  - The dividend yield.

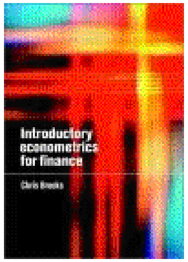
The property index and unemployment are I(1) and hence are differenced.



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

	Lags of Variable					
Dependent 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

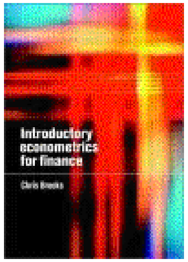


Variance Decompositions for the Property Sector Index Residuals

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

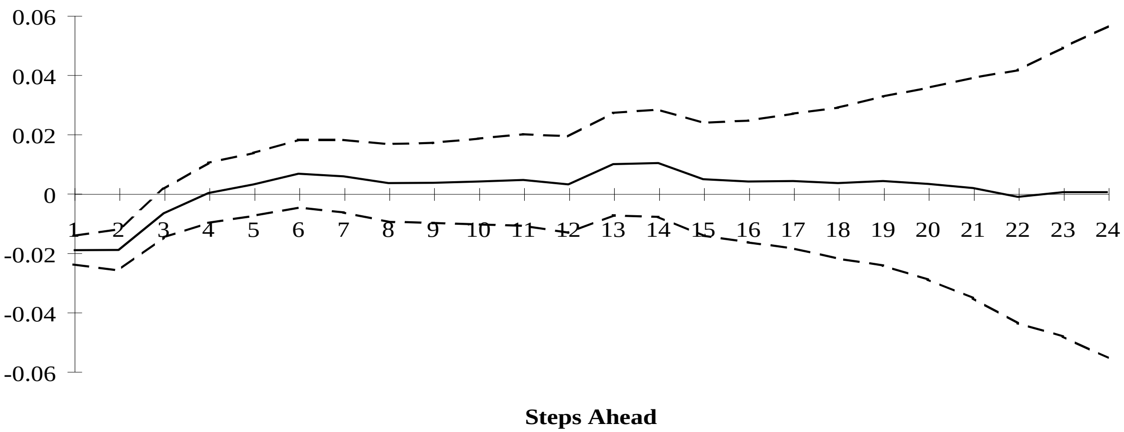
Explained by innovations in

	SIR		DIVY		SPREAD		UNEM		UNINFL		PROPRES	
Months ahead	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



Impulse Responses and Standard Error Bands for Innovations in Dividend Yield and the Treasury Bill Yield

Innovations in Dividend Yields



Innovations in the T-Bill Yield