VIII. Dynamic models revisited: VAR

1. Introduction

2. Multivariate autoregression moving average models

3. Dynamic structural models-as a special case

4. Vector autoregression representations (VAR)

5. Vector moving average representations (transfer functions)

6. Impulse response functions

7. Identification

8. Statistical inference

9. A review and applications

10. Exercises

James B. McDonald
Brigham Young University

VIII. Dynamic Econometric Models Revisited: VAR's

1. Introduction

This section will consider multivariate autoregressive moving average models for the vector $Z_t = \begin{pmatrix} Y_t \\ X_t \end{pmatrix}$. This formulation includes the dynamic structural models discussed in section (VI.2), as well as showing their relationship to vector autoregression (VAR) formulations, and lays the foundation for exogeneity tests.

2. Multivariate Autoregressive Moving Average Representation

Let $Z_t = \begin{pmatrix} Y_t \\ X_t \end{pmatrix}$ denote a multivariate autoregressive moving average (MARMA) process which can be written as

$$\boxed{F(L)\, Z_t \; = \; G(L)\, \varepsilon_t} \tag{2.1}$$

or using partitioned matrices as

$$\begin{pmatrix} F_{11}(L) & F_{12}(L) \\ F_{21}(L) & F_{22}(L) \end{pmatrix} \begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} G_{11}(L) & G_{12}(L) \\ G_{21}(1) & G_{22}(1) \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \qquad (2.1)'$$

where Y_t and X_t denote GX1 and KX1 vectors of variables, respectively.

3. Dynamic Structural Equation Representation as a Special Case

Consider the special case of equation (2.1) corresponding to

$$F_{11}(L) = B(L),\ F_{12}(L) = \Gamma(L)$$

$$F_{21}(L) = 0,\ F_{22}(L) = \Phi(L)$$

$$G_{11}(L) = I,\ G_{12}(L) = 0$$

$$G_{21}(L) = 0,\ G_{22}(L) = \theta(L)$$

which yields

$$\boxed{\begin{aligned} B(L)Y_t + \Gamma(L)X_t &= \varepsilon_{t1} \\ \varphi(L)X_t &= \theta(L)\varepsilon_{t2} \end{aligned}}$$

(3.1) is a generalization of the model considered in (VI.2) in which Y_t is endogenous and X_t is exogenous. There are G structural equations with G endogenous variables.

Zellner's condition for dynamic structural econometric models, $F_{21}(L) = 0$ with $G_{21}(L) = 0$ and $G_{12}(L) = 0$, provides the basis for exogeneity tests to be considered later.

Dynamic structural econometric models are sometimes criticized because of “exogeneity” assumptions and the ad hoc nature of some of the “exclusion restrictions”(deleting some variables from the structural equations) used to identify the model. These concerns have lead to the consideration of vector autoregressive (VAR) models. VAR models provide a method for testing exogeneity and forecasting, but are also associated with estimation and identification problems when used for economic analysis. Two tests for exogeneity developed by Granger (Investigating Causal Relations by Econometric Models and Cross -Spectral Methods, Econometrica, 1969) and Sims (Money Income and Causality, AER,1972) are briefly summarized then alternative representations of dynamic econometric models are discussed.

Test 1. Regress Y on lagged values of Y and lagged values of X and then test for the collective explanatory power of the lagged X's. If the lagged X's are not statistically significant, then X is said to fail to "Granger-cause" Y. "Granger-causality" implies the rejection of the collective explanatory power of the X's.

Test 2. Regress Y on current levels of X, past levels of X, and future levels of X. The hypothesis of the coefficients of future X's being equal to zero is consistent with X being "exogenous" to Y for failing to "Granger-cause" Y. Thus if "causality" runs one way, from X to Y, we would expect the coefficients of the "future" X's to all be zero.

Simple F tests can be used for the Granger and Sims tests. Granger-Sims causality tests based on Test 1are built into the STATA package (version 10) and can be performed by typing the command "**vargranger**" on a line following a VAR estimation command **(var** or **svar)** .

This command will be discussed again later.

In previous discussions of univariate time series models (section VII), alternative moving average (MA) and autoregressive (AR) representations were often helpful in model analysis. This will also be true with dynamic econometric models, vector autoregression models, and in exploring the implications of notions of "causality" and "exogeneity."

4. Autoregressive Representation: VAR's.

The multivariate autoregressive representation can be derived from (3.1) and is given by

$$B(L)Y_t + \Gamma(L)X_t = \varepsilon_{t1} \tag{4.1}$$

$$B_0Y_t + B_1Y_{t-1} + \ldots + B_pY_{t-p} + \Gamma(L)X_t = \varepsilon_{t1} \tag{4.2}$$

$$B_0Y_t + B_1Y_{t-1} + \ \ldots + B_pY_{t-p} = \lambda + \varepsilon_{t1} \tag{4.3}$$

where $\lambda = -\Gamma(L)X_t$. Often the "exogenous" variables are "suppressed" in the model and included in the λ term. More formally, the exogenous variables might be explicitly included in the model as in (4.2). Enders (1995, p.295) refers to this form as the **structural vector autoregressive form** (structural VAR form or SVAR). This representation can include current and lagged values of Y_t and X_t in each equation; hence, **OLS estimation of the VAR form would yield biased and inconsistent estimators**.

The structural error terms, ε_{1t}'s, are viewed as being pure innovations or structural shocks which are uncorrelated, both across time and across equations; hence, **Var (ε_{t1}) = Ω is assumed to be a diagonal matrix.** Each structural VAR equation is associated with one innovation.

The **reduced form vector autoregressive representation** (VAR) can be obtained by premultiplying equation (4.2) by B_0^{-1} and then expressing Y in terms of predetermined variables to yield

$$Y_t = -B_0^{-1}B_1Y_{t-1} - \dots - B_0^{-1}B_pY_{t-p} - B_0^{-1}\Gamma(L)X_t + B_0^{-1}\varepsilon_{t1}$$

$$Y_t = \Pi_1Y_{t-1} + \dots + \Pi_pY_{t-p} + \mu + \eta_t \tag{4.5}$$

$$Y_t = \Pi(L)Y_t + \mu + \eta_t \tag{4.6}$$

where $\Pi_i = -B_0^{-1}B_i$, $\mu = B_0^{-1}\lambda = -B_0^{-1}(L)\Gamma(L)X_t$, $\eta_t = B_0^{-1}\varepsilon_{t1}$

and $\Pi(L) = \Pi_1L + \Pi_2L^2 + \dots + \Pi_pL^p$. Note t

μ term. It is important to note that η_t represents the reduced form random disturbances. STATA

and EVIEWS are two fairly "friendly" econometric program which estimate VAR models and allows for the inclusion of X's into the model.

(4.6) is referred to as the **VAR in standard form** by Enders (1995, p. 295). Hamilton (1994, p. 327) notes that (4.6) can be viewed as a **reduced form** of a general dynamic structural model because the right hand side variables are all predetermined. Each equation in the standard or reduced form VAR representation includes one current endogenous variable and possibly lagged values of that and other predetermined variables. (4. 6) can be used for forecasting. **The reduced form VAR's can be estimated using OLS.** Even though there may be correlation between error terms, SURE will not yield more efficient estimators because each equation will normally include the same regressors. The STATA command for estimating the system of VAR equations is as follows:

var y1 y2 ... yg, lags(1 p) (X1 X2 ...Xk)

where corresponding Granger Tests can be performed by using the command

vargranger

The Akaike Information Criterion (AIC) is often used to determine the lag length.

5. Vector Moving Average Representation (transfer functions)

In order to study the dynamics of a VAR, it is useful to consider the vector moving average representation (VMA). Solving equation (4.1) or (4.6) for Y_t yields

$$Y_t = -B^{-1}(L)\Gamma(L)X_t + B^{-1}(L)\varepsilon_{t1} \tag{5.1}$$

$$= (I - \Pi(L))^{-1}\mu + (I - \Pi(L))^{-1}\eta_t$$

$$= \Phi(L)X_t + \Psi(L)\eta_t$$

$$Y_t = \sum_{i=0}^{\infty}\Phi_i X_{t-i} + \sum_{i=0}^{\infty}\Psi_i \eta_{t-i} \tag{5.2}$$

or

$$Y_t = \gamma + u_t \tag{5.3}$$

where

- $$\gamma = \Phi(L)X_t = (I - \Pi(L))^{-1}\mu = (I - \Pi(L))B_0^{-1}\Gamma(L)X_t = -B^{-1}(L)\Gamma(L)X_t$$
- $$u_t = (I - \Pi(L))\eta_t = \sum_{i=1}^{\infty}\Psi_i \eta_{t-i} = \sum_{i=0}^{\infty}\Psi_i B_0^{-1}\varepsilon_{t-i}$$
- γ **Can be interpreted as the equilibrium for Y if the X's remain unchanged**

- The random disturbances in the transfer function representation can be expressed in terms of the reduced form or the structural random disturbances

The **impact, interim, and long run multipliers**, respectively, can be obtained from (5.2) as

$$\frac{dY_t}{dX_t} = \Phi_0 = -B_0^{-1}\Gamma(L=0), \frac{dY_t}{dX_{t-i}} = \Phi_i, \text{ and}$$

$$\sum_{i=0}^{\infty} \Phi_i = \Phi(L=1) = -B^{-1}(L=1)\Gamma(L=1)$$

(5.2) is a vector moving average representation expressed in terms of the structural form VAR innovations ε_t , sin $\eta_t = B_0^{-1}\varepsilon_t$ Equation (5.2) is the **transfer function** discussed in Section VI.2 of the notes.

6. <u>Impulse Response Functions</u>

The impulse response function describes the impact of an innovation or shock on future values of variables in the model. The moving average representation of the VAR in standard form given by (5.2) yields

$$\left(\frac{dY_t}{d\eta_{t-i}}\right) = \Psi_i \qquad (6.1)$$

Since the Ψ_s's are independent of B_0, only the reduced form VAR representation is necessary to estimate (6.1). However, since the structural shocks (ε_t) are related to the reduced form errors (η_t's) by the equation, $\eta_t = B_0^{-1}\varepsilon_t$ and it will be necessary to estimate B_0 in order to "unravel" the impact of the structural innovations or shocks on the time path of Y_t to explore "causal" links between the variables. For cases in which B_0 is known or can be consistently estimated, we can use the result

$$\left(\frac{dY_t}{d\varepsilon_{t-i}}\right)=\left(\frac{dY_t}{d\eta_{t-i}}\right)\left(\frac{d\eta_{t-i}}{d\varepsilon_{t-i}}\right)$$

$$=\Psi_i B_0^{-1} \tag{6.2}$$

Economic analysis and estimation of the impulse response function is conditional on solving the identification problem, i.e. being able to estimate the matrix B_0 from observed data. The matrix of impulse response coefficients in equation (6.2) provides information about how "structural shocks" associated with different structural equaitons will impact the endogenous variables over time.

7. Identification

Recall that the identification problem in econometric models deals with the question as to whether the structural parameters can be determined from the reduced form parameters:

- $B_0\Pi+\Gamma$ (7.1 a-b)
- $\Omega=Var\left(\varepsilon_{t1}\right)=B_0Var\left(\eta_t\right)B_0^{'}=B_0\Sigma B_0^{'}$

In structural dynamic econometric models, the common practice is to impose sufficient restrictions (variable exclusions or to use at least as many instruments as there are endogenous regressors) on B and Γ to enable us to solve BΠ + Γ=0 for B and Γ in terms of the π_{ij} 's . In structural vector autoregressive formulations the identification restrictions may take different forms. However, the problem is the same: can the structural parameters be uniquely determined from the reduced form. Equation (7.1 b) is frequently the focal point of identification of VAR models. If the structure is identified and B_0 can be determined, then the structural impulse response functions can be evaluated using equation (6.2). If economic theory has provided information about ψ_s and B_0, then the impulse response function reflects the implications of economic theory about important "causal" relationships. The impulse response functions are often plotted as a function of "s" to visualize the inter-temporal impact of structural shocks. Evaluatedtypically does not involve imposing (exclusion) restrictions on the structural coefficient matrices to identify the structural model. models, typ, then the impulse response function can be identified as well as the impact multipliers from the transfer function representation. If in a given application the structural model is not identified, the criticism of VARS not having any economic content would be valid. However, if one is only interested in obtaining forecasts, rather than economic analysis, identification may not be important.

Let's consider the identification problem associated with equations (7.1 a-b) in VAR formulations in more detail: Can the [$G + G^2 + GK$] structural VAR parameters

(1) Ω G unknown diagonal elements, the variance of each structural error in each structural equation. The structural shocks are assumed to be independent of other structural shocks.

(2) B_0 G^2 unknown parameters

(3) Γ GK unknown parameters

be recovered from the [G(G+1)/2] + GK reduced form VAR parameters of the variance

(1) $\Sigma = Var(\eta_t)$ $\frac{G(G+1)}{2}$ unknown parameters

(2) Π GK unknown parameters

A <u>necessary condition</u> for identification is that there are at least

$$\left(G^2 + G + GK\right) - \left(\frac{G(G+1)}{2} + GK\right) = \frac{G(G+1)}{2}$$

restrictions imposed on B_0 and on Ω.

Furthermore, if the matrix B_0 is normalized on the G diagonal elements (each structural equation has one dependent variable with a coefficient of "1"), then a necessary identification condition is that there are at least

$$\frac{G(G+1)}{2} - G = \frac{G(G-1)}{2}$$

additional "structural" restrictions on Ω or B_0, for example if the VAR involves three endogenous variables, then (3(3-1)/2=3) additional restrictions would need to be imposed on Ω or B_0.

One approach to the identification problem is what is termed "structural" decomposition which involves determining a matrix A (B_0) which solves equation (7.1b):

$$\Omega = A \Sigma A' \tag{7.2}$$

Recall that the structural covariance matrix (Ω) of the structural shocks is assumed to diagonal

$$\begin{pmatrix} \omega^2_{\varepsilon_1} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \omega^2_{\varepsilon_G} \end{pmatrix} = A \begin{pmatrix} \sigma^2_{\eta_1} & \sigma_{\eta_1\eta_2} & \cdots & \sigma_{\eta_1\eta_G} \\ \sigma_{\eta_2\eta_1} & \sigma^2_{\eta_2} & \cdots & \sigma_{\eta_2\eta_G} \\ \vdots & \vdots & \cdots & \vdots \\ \sigma_{\eta_G\eta_1} & \sigma_{\eta_G\eta_2} & \cdots & \sigma^2_{\eta_G} \end{pmatrix} A'$$

The matrix A can then be used to construct a set of orthogonal "structural" shocks

$$\varepsilon = A\eta$$

such that equation (7.2) , var (ε) = A var (η) A', is satisfied. It is important to remember that (1) the resultant innovations or shocks only have economic meaning if $A=B_0$ and (2) the "decomposition" or "factorization" involved in (7.2) is not unique.

The Cholesky decomposition is ***one*** approach to obtaining an orthogonal decomposition (unraveling). The matrix A is selected to be of the form

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ a_{21} & 1 & \dots & 0 \\ a_{31} & a_{32} & 1 & \\ \vdots & & & \vdots \\ a_{G1} & a_{G2} & \dots & 1 \end{pmatrix}$$

Note that there are $\frac{G(G-1)}{2}$ zero restrictions imposed in the *A* matrix. Mathematically,

there is a degree of arbitrariness in where the block of zero's is placed in the matrix *A*. The argument for the "ordering" of the variables is the block diagonal matrix *A* is often based on "timing" of economic interactions, for example monetary policy can respond more rapidly to economic conditions than can the implementation of fiscal policies.

Other decompositions have been considered by Sims (1986) in his discussion of a six-variable VAR macro model, by Beveridge and Nelson (1981), and by Blanchard and Quah (1989) among others. Once again, the reader is reminded that the interpretion of "structural" impulse response functions shocks depend on the selection of *A* , ($\varepsilon = A\eta$), and only have economic interpretations if $A = B_0$.

Markku Lanne and Helmut Lutkepohl (2010, Structural Vector Autoregressions with nonnormal residuals, JBES, 159-168) demonstrate that distributional assumptions can sometimes be used to identify structural shocks.

8. Estimation and Testing Hypotheses

a. Short and long-run VAR models

The STATA command **varbasic** performs a Cholesky decomposition and reports the corresponding "structural" impulse response functions.

varbasic depvar_list , lags(p) exog(var_list)

The STATA command **svar** allows for more flexibility in solving the identification problem and in estimation. These commands are organized according to estimation of short-run and long-run formulations. The format for estimating the short-run formulation is as follows:

svar depvarlist, <u>acon</u>straints lags(p) exog(var_list),

svar depvarlist, <u>bcon</u>straints lags(p) exog(var_list), or

or **svar depvarlist, <u>acon</u>straints <u>bcon</u>straints lags(p) exog(var_list)**

where the acon or bcon commands can be written in different forms (aeq( ) or acns ( )), but allow the imposition of identifying restrictions and correspond to GxG matrices. If either *acon* or *bcon*. is deleted, it is assumed to correspond to an identity matrix. Each of the acon or bcon options be written in different ways and facilitates structural identification and estimation. The purpose of unknown diagonal elements in the b_constraints is to scale the structural innovations to have unit variance. The a_constraints take into account identifying restrictions on the B_0 matrix and also satisify (7.1b),

$$\Omega = Var\left(\varepsilon_{t1}\right) = B_0 Var\left(\eta_t\right) B_0^{'} = B_0 \Sigma B_0^{'}$$

If the B_0 matrix in a bivariate VAR model with one lag is assumed to be of the form

$$B_0 = \begin{pmatrix} -1 & 0 \\ \beta_{21} & -1 \end{pmatrix} ,$$

then the **svar** command could be written as follows

mat A=(1,0\.,1)
mat B=(.,0\0,.)
svar y1 y2, aeq(A) lags(1) exog(x's) or
svar y1 y2, aeq(A) beq(B) lags(1) exog(x's)

where the "." in the matrix command indicates that the corresponding coefficient in the B_0 matrix needs to be estimated and the other parameters are set equal to the indicated values. The ***aeq*** option is a matrix alternative to the **aconstraints** option.

To discuss the estimation of long-run VAR models, it will be helpful to consider VMA or transfer function representation (equation (5.1)) corresponding to long-run adjustments

$$\begin{aligned} Y_t &= \left(I - \Pi\left(L = 1\right)\right)^{-1} \mu + \left(I - \Pi\left(L = 1\right)\right)^{-1} \eta_t \\ &= -B^{-1}\left(L = 1\right)\Gamma\left(L = 1\right) X_t + \left(I - \Pi\left(L = 1\right)\right)^{-1} B_0^{-1} \varepsilon_t \end{aligned}$$

The coefficient matrix of the structural disturbances, shocks, or innovations corresponds to the C matrix in STATA,

$$C = \left(I - \Pi\left(L = 1\right)\right)^{-1} B_0^{-1}$$

The matrix C can be interpreted as the matrix of long-run impact of structural shocks on the endogenous variables. For example, in a bivariate VAR model of the money supply (m) and output (gdp) it might be expected that an unexpected shock to the money supply would not have a long-run impact on output and similarly, an unexpected shock to output would not have a long-run impact on the money supply. The corresponding C matrix would have zeros off-diagonal elements. The corresponding STATA commands could be

mat C=(.,0\0,.)

svar gpd m, lreq(C) or alternatively as

svar d.ln_m d.ln_gdp, lreq(C) to take account of non-stationarity of m and gdp

b. Hypothesis testing

Hypotheses about reduced form (standard form) VAR coefficients,

$$Y_t = \mu + \Pi_1 Y_{t-1} + \ldots + \Pi_p Y_{t-p} + \eta_t,$$

can be tested using the likelihood ratio statistic

$$LR = N (\ln |\Sigma_R| - \ln |\Sigma_{UR}|) \sim \chi^2 \text{ (d.f.)}$$

where Σ_R and Σ_{UR} denote the estimated variance-covariance matrix corresponding to the restricted and unrestricted estimates and df =degrees of freedom. Sim's suggests using

$$(N-c) (\ln |\Sigma_R| - \ln |\Sigma_{UR}|)$$

for small samples where c denotes the number of estimated parameters in the unrestricted model.

9. A Review and an Application

(a) A dynamic structural econometric model may be expressed as a special case

$$\begin{array}{c} B(L)\,Y_t + \Gamma(L)\,X_t = \varepsilon_{t1} \\ \varphi(L)\,X_t = \theta(L)\,\varepsilon_{t2} \end{array}$$

A structural vector autoregressive model can be written in the form

$$B_0 Y_t + B_1 Y_{t-1} + \ldots + B_p Y_{t-p} + \Gamma(L) X_t = \varepsilon_{t1} \tag{9.2}$$

Note that the structural VAR generalizes the dynamic structural econometric model.

[illegible] (9.3)

where $\mu = -B_0^{-1}\Gamma(L)X_t$, with impulse response functions given by

$$\left(\frac{dY_t}{d\varepsilon_{t-i}}\right) = \Psi_i B_0^{-1} \tag{9.4}$$

where $\Psi_i = \frac{d^i}{dt^i}(I - \Pi_1 t - \ldots - \Pi_s t^s)^{-1}$ with $t = 0$.

(b) An application: cobweb model

We now consider an application of these approaches to analyzing the Cobweb Model defined in the homework in section (VI)

$Q_t = \beta P_t + \xi Y_t + \delta_d + \varepsilon_{td}$

$Q_t = \delta P_{t-1} + \alpha w_t + \delta_s + \varepsilon_{ts}$.

This dynamic structural model can also be written in the following manner

$$\begin{pmatrix} -1 & \beta \\ -1 & 0 \end{pmatrix}\begin{pmatrix} Q_t \\ P_t \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \delta \end{pmatrix}\begin{pmatrix} Q_{t-1} \\ P_{t-1} \end{pmatrix} + \begin{pmatrix} \alpha & 0 & \xi \\ \gamma & \rho & 0 \end{pmatrix}\begin{pmatrix} 1 \\ W_t \\ Y_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{td} \\ \varepsilon_{ts} \end{pmatrix} = 0$$

Multiplying the dynamic structural equation by the inverse of the leading coefficient matrix yields the reduced form:

$$\begin{pmatrix} Q_t \\ P_t \end{pmatrix} = \begin{pmatrix} 0 & \delta \\ 0 & \dfrac{\delta}{\beta} \end{pmatrix}\begin{pmatrix} Q_{t-1} \\ P_{t-1} \end{pmatrix} + \begin{pmatrix} \gamma & \rho & 0 \\ \dfrac{\gamma - \alpha}{\beta} & \dfrac{\rho}{\beta} & \dfrac{\xi}{\beta} \end{pmatrix}\begin{pmatrix} 1 \\ W_t \\ Y_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{td} \\ \dfrac{\varepsilon_{ts} - \varepsilon_{td}}{\beta} \end{pmatrix}$$

The reduced form includes 7 non-zero coefficients which are functions of 6 structural parameters.

Some practitioners may object to the imposition of "economic" hypotheses (exclusions) on coefficients of variables in the model. Vector Auto Regression (VAR) models attempt to circumvent these assumptions.

The structural VAR Model corresponding to the structural cobweb model could be written as

$$\begin{pmatrix} -1 & \beta_{12}^{0} \\ -1 & \beta_{22}^{0} \end{pmatrix}\begin{pmatrix} Q_t \\ P_t \end{pmatrix} + \begin{pmatrix} \beta_{11}^{1} & \beta_{12}^{1} \\ \beta_{21}^{1} & \beta_{22}^{1} \end{pmatrix}\begin{pmatrix} Q_{t-1} \\ P_{t-1} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \end{pmatrix}\begin{pmatrix} 1 \\ W_t \\ Y_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{td} \\ \varepsilon_{ts} \end{pmatrix} = 0$$

Note that the structural VAR model includes the cobweb model as a special case. The structural VAR includes 12 free parameters compared to 6 parameters for the cobweb model.

The corresponding reduced form VAR for the cobweb model can be written in the form

$$\begin{pmatrix} Q_t \\ P_t \end{pmatrix} = \begin{pmatrix} \pi_{11}^1 & \pi_{12}^1 \\ \pi_{21}^1 & \pi_{22}^1 \end{pmatrix} \begin{pmatrix} Q_{t-1} \\ P_{t-1} \end{pmatrix} + \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{21} & \pi_{23} \end{pmatrix} \begin{pmatrix} 1 \\ W_t \\ Y_t \end{pmatrix} + \begin{pmatrix} \eta_{t1} \\ \eta_{t2} \end{pmatrix}$$

Note that this form is a generalization of the reduced form of the cobweb model. The reduced form VAR involves 10 parameters compared to 7 parameters for the reduced form corresponding to the structure or 6 parameters associated with the reduced form estimates derived from the estimated structure. Thus the reduced form VAR can be looked at as over-fitting the underlying model. The hypotheses associated with the original specification (exclusions and exogeneity) can be collectively tested using the likelihood ratio test. Let $\hat{\Sigma}_{10}$ $\hat{\Sigma}_7, \hat{\Sigma}_6$, denote the variance covariance matrices associated with the unrestricted and two restricted VAR's.

$$\text{(T-10)} \quad \left(\ln|\hat{\Sigma}_7| - \ln|\hat{\Sigma}_{10}|\right)$$

and

$$\text{(T-10)} \quad \left(\ln|\hat{\Sigma}_6| - \ln|\hat{\Sigma}_{10}|\right)$$

can be used to test the hypotheses of interest. These test statistics are asymptotically distributed as $\chi^2(3)$ and $\chi^2(4)$, respectively.

If the researcher is only interested in forecasting, the reduced form (structure or

VAR based reduced form) can be used. If the researcher is also interested in economic analysis of the structure, transfer functions or impulse functions will be of interest.

The form for the impulse response functions is particularly simple in the case of a first-order VAR

$$\left(\frac{d\begin{pmatrix} Q_{t+i} \\ P_{t+i} \end{pmatrix}}{d\varepsilon_t} \right) = \Pi_1^i B_0^{-1}$$

.

Estimates of B_0 can be obtained by solving

$$\begin{pmatrix} \omega_1^2 & 0 \\ 0 & \omega_2^2 \end{pmatrix} = \begin{pmatrix} -1 & \beta_{12}^0 \\ -1 & \beta_{22}^0 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ \beta_{12}^0 & \beta_{22}^0 \end{pmatrix}$$

for the structural ω's and β's in terms of the reduced form σ's . There are four unknown structural parameters and three reduced form parameters; hence, there needs to be one restriction (G(G-1)/2 =1) restriction imposed. Let's assume that $\beta_{22} = 0$ as in the original specification, which was motivated by the timing of the economic decisions.

The STATA commands to estimate the reduced form and structural VAR's are as follows:

var q p, lags(1) exog(w f)

mat A =(1,.\1,0)

svar q p, aeq(A) exog(w f)

Other STATA commands which might be of interest are

varirf calculates and analyzing impulse response functions

varstable checks stability conditions of var or svar estimates

varsoc obtain lag-order statistics for a set of VAR's

varwle obtain Wald lag exclusion statistics after var or svar

vargranger performs pariwise Granger causality tests after var or svar

varlmar obtain LM statistics for residual autocorrelation after var or svar

varnorm tests for normally distributed disturbances after var or svar

varfcast compute computes dynamic forecasts of dependent variables after var/svar

EXERCISES

1. Cover, Enders, and Hueng ("Using the Aggregate Demand-Aggregate Supply Model to Identify Structural Demand-side and Supply side Shocks (using a bivariate VAR)," Journal of Money Credit and Banking, 38(2006), 777-790) consider a bivariate VAR

Reduced form VAR: $Z_t = Z_0 + A(L)Z_t + \eta_t \qquad Z_t = \begin{pmatrix} y_t \\ p_t \end{pmatrix} \qquad Var(\eta_t) = \Sigma$

Structural VAR: $\begin{aligned} B_0 Z_t - B_0 A(L) Z_t &= B_0 Z_0 + B_0 \eta_t \\ &= B_0 Z_0 + \varepsilon_t \end{aligned}$ where ε_t

shocks with $Var(\varepsilon_t) = \Omega$. The 2x2 matrix of long-run response multipliers for

Z_t corresponding to the structural shocks can be expressed as (see p.15)

$$C = \left(B_0\left(I - A(L=1)\right)\right)^{-1} = \left(I - A(L=1)\right)^{-1} B_0^{-1} = \left(I - A(L=1)\right)^{-1} D \quad ***$$

Identification and estimation of the four parameters in the D matrix (hence in the B_0 = D^{-1} matrix) requires at least four restrictions, which were chosen to be satisfied by

- $\Omega = I_2$
- Demand shock have no permanent impact on output, i.e. the element in the first row and second column of *** is zero. This is known as the Blanchard-Quah long-run neutrality restriction. Show that this is equivalent to $d_{12}\left(1 - a_{22}\left(L=1\right)\right) + d_{22}a_{12}\left(L=1\right) = 0$

2. Consider the Kmenta-Smith(RESTAT, (August 1973, 299-307)) structural form defined by

$$C_t = \gamma_{10} + \beta_{12}Y_t + \beta_{15}L_t + \beta_{1\,11}C_{t-1} + \varepsilon_{t1}$$
$$I_t^d = \gamma_{20} + \beta_{25}Y_t + \gamma_{25}(S_{t-1} - S_{t-2}) + \gamma_{2\,12}t + \gamma_{27}I_{t-1}^{d} + \varepsilon_{t2}$$
$$I_t^r = \gamma_{30} + \beta_{33}r + \gamma_{35}(S_{t-1} - S_{t-2}) + \gamma_{3\,12}t + \gamma_{38}I^r_{t-1} + \varepsilon_{t3}$$
$$I_t^i = \gamma_{50} + \beta_{43}r + \gamma_{45}\,(S_{t-1} - S_{t-2}) + \gamma_{4\,12}t + \gamma_{49}I_{t-1}^{i} + \varepsilon_{t4}$$
$$\gamma_t = \gamma_{50} + \beta_{52}Y_t + \gamma_{53}M_t + \gamma_{54}M_{t-1} + \varepsilon_{t5}$$
$$Y_t = C_t + I_t^d + I_t^i + G_t$$
$$S_t = Y_t - I_t^i$$
$$L_t = M_t + R_t$$

where

- Y = gross national product ($ bill.)
- C = consumption expenditures ($ bill.)
- I^d = producer's outlays on durable plant and equipment ($ bill.)
- I^r = residential construction ($ bill.)
- I^i = investment in inventories ($ bill.)
- G = government purchases of goods and services plus net foreign investment ($ bill.)
- S = final sales of goods and services ($ bill.)
- t = time in quarters (first quarter of 1954 = 0)
- r = yield on all corporate bonds (%)

M = money supply, i.e., demand deposits plus currency outside banks ($ bill.)
R = time deposits in commercial banks ($ bill.)
L = money supply plus time deposits in commercial banks ($ bill.)

a. Write the form of the reduced form representations.
b. Investigate the form of the associated transfer functions.
c. Compare the form of the transfer functions associated with those used by Maloney and Ireland.
d. Discuss how you could compare restricted and unrestricted estimated transfer functions, also see exercise 2 in Section V)

a. Write out a structural and reduced form VAR representation of the Kmenta-Smith model.

b. Discuss how, given data, you could use the VAR representation to test the restrictions imposed in the Kmenta-Smith formulation.

c. What is the relationship between the coefficients in the transfer functions and the impulse response functions.

3. Samuelson proposed a model known as the multiplier-accelerator model which has the potential to generate a business cycle. This model is defined by

$$y_t = c_t + I_t + G_t$$

$$c_t = \gamma y_{t-1} \qquad 0<\gamma<1$$
$$I_t = \alpha (c_t - c_{t-1}) \qquad 0<\alpha$$

a. Write out the forms for the
   1. structural and
   2. reduced form

   vector autoregressive models.

b. How can the exogeneity assumptions be tested?

c. Describe how, given data, you could use the VAR representation to test the validity of the economic hypothesis imposed by the Multiplier-Accelerator model.

4. Given the relationships

$$\text{(E.1)} \quad F(L)\, Z_t = G(L)\, \varepsilon_t \qquad \text{MARMA}$$
$$\text{(E.3)} \quad Z_t = C(L)\varepsilon_t \qquad \text{Multivariate MA}$$
$$\text{(E.4)} \quad A(L)Z_t = \varepsilon_t \qquad \text{Multivariate AR}$$
$$C(L) = F^{-1}(L)\, G(L)$$
$$A(L) = C^{-1}(L) = G^{-1}(L)F(L)$$

Investigate the relationship between the conditions:

Zellner:	$F_{21}(L) \equiv 0, G_{21}(L) \equiv 0, G_{12}(L) \equiv 0$
Granger (MA):	$C_{21}(L) \equiv 0$
AR:	$A_{21}(L) \equiv 0$

Hint: Recall that the inverse of a partitioned matrix is given by

$$D^{-1} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}^{-1} = \begin{pmatrix} D^{11} & D^{12} \\ D^{21} & D^{22} \end{pmatrix}$$

$$D^{11} = (D_{11} - D_{12}D_{22}^{-1}D_{21})^{-1}$$

$$D^{12} = -D^{11}D_{12}D_{22}^{-1}$$

$$D^{21} = -D_{22}^{-1}D_{21}D^{11}$$

$$D^{22} = (D_{22} - D_{21}D_{11}^{-1}D_{12})^{-1}$$

Appendix
Exogeneity tests: more details

1. Introduction

An important objective of research in economics is to determine relationships between variables and explain which variables are exogenous and which are endogenous. Specification of causal relationships is usually viewed as being the domain of the theoretician and not determined statistically. A number of statistical tests have been proposed which claim to test for causality or exogenity. There has been considerable discussion as to whether the proposed tests do indeed check what is alleged. One of the key issues in the discussion centers around what is meant by causality. The paper by Zellner (1979, Carnegie Rochester Conference series on Public Policy) summarizes the major positions in this debate.

The purpose of this section is to outline the suggested tests and associated definitions. Before outlining Granger's test (1969, Econometrica), some notation needs to be discussed.

Let A_t be a stationary stochastic process, let $\overline{A}_t$ represent the set of past values $\{A_{t-j}, j=1, 2, \ldots, \infty\}$, and $\overline{\overline{A}}_t$ represent the set of past and present values $\{A_{t-j}, j=0, 1, \ldots, \infty\}$. Further, let $\overline{A}_t(k)$ represent the set $\{A_{t-j}, j=k+1, \ldots, \infty\}$.

Denote the optimum, unbiased, least squares predictor of A_t that uses the set of values of B_t by $P_t(A|B)$. Thus, for instance, $P_t(X|\overline{X})$ will be the optimum predictor of X_t using only past X_t. The prediction error will be denoted by $\varepsilon_t(A|B) = A_t - P_t(A|B)$. Let $\sigma^2(A|B)$ be the variance of $\varepsilon_t(A|B)$. Let U_t denote all the information in the universe accumulated as of time t, and let $U_t - Y$ denote all this information other than the series Y_t.

We then have the following definition.

Definition: Causality

If $\sigma^2(Y|\overline{U}) < \sigma^2(Y|\overline{U}-\overline{X})$, then X is said to cause Y, denoted by $X \rightarrow Y$; t
is "causing" Y if we are better able to predict Y_t using all available information, including past values of X and Y, rather than if all information other than X had been used.

Definition: Feedback

If $\sigma^2(X|\overline{U}) < \sigma^2(X|\overline{U}-\overline{Y})$ and

$\sigma^2(Y|\overline{U}) < \sigma^2(Y|\overline{U}-\overline{X})$, then

we say that feedback is occurring, which is denoted $Y \leftrightarrow X$, that is, feedback is said to occur when X is "causing" Y and Y is "causing" X.

Definition: Instantaneous Causality

If $\sigma^2(Y|\overline{U},\overline{\overline{X}}) < \sigma^2(Y|\overline{U})$, we say that instantaneous causality between X a
occurring. In other words, the current value of Y_t is better "predicted" if the present value of X_t is included in the "prediction" than if it is not.

There are a number of impediments to operational interpretations of the previous definitions. The feasibility of optimal minimum variance forecasts may be questionable, and the notion of using all information in the universe is untractable. The relationship between these definitions and the philosophical notion of causality is an important question.

We now consider a representation of a dynamic econometric model which facilitates a discussion of tests proposed by Granger (1969, Econometrica) and Sims (1972, AER). The statistical tests will then be defined and an application considered.

2. Operational Definition of Tests

Zellner's definition of X_t being exogenous is that

$F_{21}(L) = 0$ with $G_{12}(L) = 0$ and $G_{21}(L) = 0$

This implies $C_{21}(L) = 0$ is the moving average representation

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} C_{11}(L) & C_{12}(L) \\ 0 & C_{22}(L) \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

which implies

$$\boxed{Y_t = C_{11}(L)\varepsilon_{1t} + C_{12}(L)C_{22}^{-1}(L)X_t}$$

Thus Y_t depends on current and lagged X_t's. Sim's proposes regressing

Y_t on current, lagged, and future X_t's and performing a joint hypothesis test on the coefficients of future X's being zero.

Failure to reject this hypothesis is consistent with X_t being exogenous to Y_t and not the other way around. Rejecting the null hypothesis raises questions about Y_t being exogenous. Two problems associated with implementing this test are observed: (1) how many lags should be used and (2) the random disturbance is $C_{11}(L)\varepsilon_{1t}$ which will likely exhibit autocorrelation and raise questions about the validity of the usual F-statistic in performing such tests. Reference: Hamilton (1995, p. 304).

A <u>second method</u> of testing for causality can be motivated from the multivariate autoregressive representation of the model

$$\begin{pmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{pmatrix} \begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$A_{21}(L) = 0$ if and only if $C_{21}(L) = 0$ (Granger Causality). Zellner's definition

implies $C_{21}(L) = 0$ which is equivalent to $A_{21}(L) = 0$. (A good exercise). Hence,

$$A_{11}(L)Y_t + A_{12}(L)X_t = \varepsilon_{1t}$$
$$A_{22}(L)X_t = \varepsilon_{2t}$$
$$A_{11}(L) = A_0 + A_1L + ...$$

Premultiplying by A_0^{-1} and solving for Y_t yields

$$\boxed{Y_t = A_0^{-1}A_1Y_{t-1} + A_0^{-1}A_2Y_{t-2} + ... + A_0^{-1}A_{12}(L)X_t + A_0^{-1}\varepsilon_{1t}}$$

This form suggests a second test*

> <u>Regress Y on lagged Y and current and lagged X's. If the coefficients on the X's are significantly different from zero, then the X's "cause" Y</u>. If not, then X does not "cause" Y. A Chow test is asymptotically distributed as an F. Hamilton (1995, p. 305)

*This test is probably most commonly applied by regressing Y on lagged Y and lagged X's.

3. <u>An Application of Causality Tests</u>

A widely used model in monetary analysis is given by

$$Y_t = \sum_{j=0}^{\infty} h_j m_{t-j} + \eta_t \quad \text{where}$$

y_t = log of nominal income
m_t = log of money
$E(\eta_t m_{t-j}) = 0$, j = 0, 1, 2, . . .

The h_j's have been interpreted as dynamic multipliers (Anderson and Jordan, 1968, <u>Federal Reserve Bank of St. Louis Review</u>).

The two major criticisms of this model and its interpretation are that this equation is not the final form or transfer function because many other variables have an impact, and secondly, m_t need not be exogenous because the monetary authority considers the behavior of lagged y_t's in the determination of m_t.

Sims (1972, <u>AER</u>) considered the model

$$Y_t = \sum_{j=-\infty}^{\infty} \delta_j m_{t-j} + \varepsilon_t$$

and tested the joint hypothesis that the coefficients of future m_t's were equal to zero,

Ho: $\delta_{-1} = \delta_{-2} = ... = 0$

Sim's conclusion was that he could not reject the hypothesis with a high degree of confidence. Sim's paper and statistical results have generated considerable discussion in the literature. Some of the limitations were discussed in the previous notes.