E 4101/5101
Lecture 7: Review of econometric models of (causal and stationary) VARs

Ragnar Nymoen

Department of Economics, University of Oslo

4 March 2014

Introduction I

- Main references:
  - H (Hamilton)
    - Ch 9. (Simultaneous equation, in particular 9.4 and 9.5.)
    - Ch 11.6 (VARs and structural econometric models)
  - DM (Davidson and MacKinnon, Ch 13 (note their reference back to Ch 12.2 on Seemingly Unrelated Regressions, known from e.g. E-4160, and the unrestricted reduced form (URF) of the simultaneous equations model (also known from E-4160)
  - A couple of introductory/intermediate books in econometrics which cover *exogeneity* and the *delta method* are mentioned at the end of the slide set.

Introduction II

- As mentioned by Hamilton, vector autoregressions, VARs have become widely adopted in macroeconomics after the so called Sims' critique, in Sims (1980), about "incredible" identifcation assumptions in large scale simulatenous equations models.
- In line with this,Hamilton motivates the VAR by its "convenience for estimation and forecasting": meaning that VARs are easy to estimate and are useful for forecasting.
- The VAR has another very important role as well, as a *statistical model* that underlies *identified structural econometric models*.
- This role is important both for the stationary case and for the case with unit-roots and (potential) co-integration. Three text-books that develop this viewpoint are: Hendry (1955), Johansen (1995) and also Bårdsen and Nymoen (2014).

The bivariate first order VAR I

The simplest example, is a vector autoregressive process (VAR) with two variables and first order dynamics, written in H(amilton) notation as

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix}, \qquad (1)$$

where $\varepsilon_{y,t}$ and $\varepsilon_{x,t}$ are two white-noise timer series (correlated or uncorrelated).

- Notation: Follow H and use y_t for both stochastic variable and realization
- Condition for *weak stationarity*: Neither of the roots (eigenvalues) of the associated characteristic equation are on the unit-circle. Moduli not equal to one.

The bivariate first order VAR II

- Condition for *causal VAR*: Both roots have modulus < 1.
- Harald's Lecture notes 4 and 5 give the necessary background both to specification and estimation of (1), and the generalization to n variables and p lags.

Reference: companion form I

The companion form is well suited for generalizations.
Let $\mathbf{y}_t$ be the $n \times 1$ vector.

$$\mathbf{y}_t = [y_{1t}, y_2, \ldots, y_{nt}]'$$

The VAR of or order p is:

$$\mathbf{y}_t = \boldsymbol{\pi}_1 \mathbf{y}_{t-1} + \boldsymbol{\pi}_2 \mathbf{y}_{t-2} + ... + \boldsymbol{\pi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (2)$$

where $\boldsymbol{\pi}_i$ is a $n \times n$ matrix with coefficients and $\boldsymbol{\varepsilon}_t$ is a vector with white-noise disturbances that may be correlated

Reference: companion form II

Write (2) in companion form:

$$\underbrace{\begin{bmatrix} \mathbf{y}_t \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p+1} \end{bmatrix}}_{\boldsymbol{\xi}_t} {}_{np\times 1} = \underbrace{\begin{bmatrix} \boldsymbol{\pi}_1 & \boldsymbol{\pi}_2 & \cdots & \boldsymbol{\pi}_{p-1} & \boldsymbol{\pi}_p \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix}}_{\mathbf{F}_{np\times np}} \underbrace{\begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{y}_{t-2} \\ \vdots \\ \mathbf{y}_{t-p} \end{bmatrix}}_{\boldsymbol{\xi}_{t-1}} + \underbrace{\begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{\mathbf{v}_t}$$

$$\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t, \quad (3)$$

with the symbols in H equation (10.1.11) p 259 appearing in (3).

Reference: companion form III

- Stationarity: All the eigenvalues of the companion matrix **F** have moduli from

$$|\mathbf{F}-\lambda\mathbf{I}| = 0$$

  that are different from unity. See H page 259.
- The VAR is causal if the moduli of all roots are less than 1.
- Note that the number of roots is increasing in both p and n, later under cointegration, we will consider the consequences of allowing a subset of roots to be on the unit circle.

Open VARs with deterministic terms I

- One natural way to obtain a empirically relevant and stastically valid VARs, is to include deterministic terms and non-modelled random explanatory variables:

$$\mathbf{y}_t = \mathbf{c}\mathbf{d}_t + \boldsymbol{\pi}_1\mathbf{y}_{t-1} + \boldsymbol{\pi}_2\mathbf{y}_{t-2} + ... + \boldsymbol{\pi}_p\mathbf{y}_{t-p} + \boldsymbol{\beta}\mathbf{x}_t + \boldsymbol{\varepsilon}_t \quad (4)$$

  where $\mathbf{d}_t$ is a vector with deterministic variables, for example:

$$\mathbf{d}_t = (1, Trend_t, dum_{1t}, \ldots, dum_{Kt})$$

dum_{kt} is a dummy variable representing

- Seasonal dummy
- Dummy for structural break (change in mean of $\mathbf{y}_t$), or outlier
- $\mathbf{x}_t$ is a vector of conditioning economic variables, and $\boldsymbol{\beta}$ is a matrix of coefficients for these variables.

Example data set: Market data for fish I

- Same data set as Angrist et. al. (2000).
- Obtained from: http://people.brandeis.edu/~kgraddy/data.html. Daily Fulton Fish Market Data
- Daily data for price of whiting and quantity sold for a single trader at the Fulton Fish Market in New York
- 111 observations from beginning of Dec 1992 to early May 1993
- In addition to log of price and quantity, the data set contains dummies that may affect demand and supply seperately
- A first "modelling step" is to formulate a VAR that might conform with the assumption of the VAR(p) with Gaussian-errors.

Joint and conditional models I

- Analogous to the standard regression model, the VAR can be interpreted as the joint probability density function (pdf) of the n random variables contained in $\mathbf{y}_t$, written in "equation form".
- Since we have have time series data: the joint pdf if however conditional on the history of the random vector $\mathbf{y}_t$.
- Hence,for $n = 2$ and $p = 1$

$$f(y_{1t}, y_{2t}) = f(y_{1t} \mid y_{2t})f(y_{2t}) = f(y_{1t} \mid y_{2t})f(y_{1t}) \qquad (5)$$

  where the conditioning on y_{1t-1} and y_{2t-1} in the pdfs have been suppressed to save notation.
- (5) immediately suggests two different econometric models:

Joint and conditional models II

1. Joint modelling of the pdf $f(y_{1t}, y_{2t})$ in terms of a simultaneous equations models or a recursive system of equations.
2. Conditional modelling, of y_{1t} given y_2; or of y_{2t} given y_{1t}

- The purpose of the modelling, will (ideally) determine the choice of modelling perspective

Conditional and marginal model I

- A VAR can always be re-parameterized in terms of a conditional model and a marginal model.
- The relevance of the conditional model stems from the fact that it can often, but not always, contain *parameters of interest* (i.e., in terms of economic theory or for the purpose of the analysis).
- Write the VAR (1) as

$$y_t = \mu_{y,t-1} + \varepsilon_{y,t} \tag{6}$$

$$x_t = \mu_{x,t-1} + \varepsilon_{x,t} \tag{7}$$

Conditional and marginal model II

For simplicity: A Gaussian-VAR

$$\begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \mid x_{t-1}, y_{t-1}\right). \tag{8}$$

The conditional distribution of $\varepsilon_{x,t}$ and $\varepsilon_{y,t}$ is bi-variate normal (with expectation zero and covariance

$$\begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}.$$

By following the steps in *Note to Lecture 7*, we obtain the conditional model for y_t as

$$y_t = \phi_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \tag{9}$$

Conditional and marginal model III

which is an ADL(1,1) model with parameters

$$\phi_1 = \pi_{11} - \frac{\sigma_{xy}}{\sigma_x^2}\pi_{21}, \beta_0 = \frac{\sigma_{xy}}{\sigma_x^2}, \beta_1 = \pi_{12} - \frac{\sigma_{xy}}{\sigma_x^2}\pi_{22}. \quad (10)$$

and a normal disturbance:

$$\varepsilon_t \sim N(0, \sigma^2 \mid x_t, y_{t-1}, x_{t-1}). \quad (11)$$

with variance:

$$\sigma^2 = \sigma_y^2(1 - \rho_{xy}^2). \quad (12)$$

because of the Gaussian assumption. Remember that ρ_{xy}^2 is the squared correlation coefficient:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2}\sqrt{\sigma_x^2}}. \quad (13)$$

Conditional and marginal model IV

The "conditional plus marginal model" of the VAR is completed by the marginal model equation for x_t:

$$x_t = \pi_{21} y_{t-1} + \pi_{22} x_{t-1} + \varepsilon_{x,t} \tag{14}$$

Note that

$$E(\varepsilon_t \mid \varepsilon_{x,t}) = 0 \tag{15}$$

is a consequnce of the assumptions.

- Estimation is by OLS, which is conditional FIML for the case of Gaussian disturbances.
- If our purpose is to test hypotheses about the parameters in the conditional model , we do not need to estimate the marginal model (14).

Conditional and marginal model V

- If the purpose is forecasting, or policy analysis (what happens if there is a change in the marginal model?), the marginal model must also be estimated.
- This allows us to make empirical assessment of:
  - Absence of feed-back (one way Granger causality): $\pi_{21} = 0$ (maintaining $E(\varepsilon_t \mid \varepsilon_{x,t}) = 0$)
  - Invariance or not of the parameters of the conditional expectation function in (9) in the case of structural breaks in the marginal model equation: *Autonomy* and *super-exogeneity*
  - Weak, strong and super exogeneity for conditional equations and sub-systems is reviewed in the slide set to Lecture 5 to ECON 4160 autumn 2013, available form that course page.
  - English textbooks that cover exogeneity are Greene (2012), and Hill, Griffiths and Lim (2012)

Conditional and marginal model VI

- An alternative. in Norwegian. is Ch. 8 in Bårdsen and Nymoen (2014).

Sequential conditioning—the general case I

- The derivation of the ADL$(1,1)$ was by *sequential conditioning*: First, by starting from the VAR, we have already conditioned on the history (y_{t-1} and x_{t-1}). Second, we conditioned on current x_t.
- This approach is general and can be used to derive, from a large joint distribution (many variables and long lags), a conditional ADL(p,p)

$$y_t - \phi_1 y_{t-1} - \ldots - \phi_p y_{t-p} = \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_p x_{t-p} + \varepsilon_t, \tag{16}$$

- By the same token: Can also have $n-1$ different x variables in the conditional model, if the dimension of the VAR is n.

Equilibrium correction model I

The conditional ADL can be re-parameterized as an equilibrium correction model, ECM.
For the first order model (with intercept ϕ_0 added, as in the note to Lecture 7):

$$\Delta y_t = \phi_0 + \beta_0 \Delta x_t + (\phi_1 - 1) y_{t-1} + (\beta_0 + \beta_1) x_{t-1} + \varepsilon_t$$

$$\Delta y_t = \beta_0 \Delta x_t$$

$$+ (\phi_1 - 1) \left\{ y_{t-1} - \underbrace{\frac{\phi_0}{(1-\phi_1)} - \frac{(\beta_0 + \beta_1)}{(1-\phi_1)} x_{t-1}}_{y^*_{t-1}} \right\} + \varepsilon_t$$

Equilibrium correction model II

Define

$$y_t^* = \frac{\phi_0}{(1-\phi_1)} - \frac{(\beta_0+\beta_1)}{(1-\phi_1)} x_{t-1}$$

as the conditional equilibrium path for y_t, then equilibrium-correction is seen to be the obvious name for this parameterization of the model.

Equilibrium correction is inherent for stationary variables. For $\beta_0 = \beta_1 = 0$, the solution can be written:

$$y_t = \underbrace{\frac{\phi_0}{1-\phi_1}}_{y^*} + \left\{ y_0 - \underbrace{\frac{\phi_0}{1-\phi_1}}_{y^*} \right\} \phi_1^t + \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i}.$$

Equilibrium correction model III

For $p = 4$ (with intercept ϕ_0 added)

$$y_t - \sum_{i=1}^{4} \phi_i y_{t-i} = \phi_0 + \sum_{i=0}^{4} \beta_i x_{t-i} + \varepsilon_t, \tag{17}$$

one possibility is to put the levels term at the fourth lag

$$\begin{aligned} \Delta y_t = \phi_0 + \sum_{i=1}^{3} \phi_i^{\dagger} \Delta y_{t-i} + \sum_{i=0}^{3} \beta_i^{\dagger} \Delta x_{t-i} \\ + (\phi(1) - 1) y_{t-4} + \beta(1) x_{t-4} + \varepsilon_t \end{aligned} \tag{18}$$

Equilibrium correction model IV

where $\phi(1)$ is $\phi(L)$ with $L = 1, \beta(1)$ is the same for $\beta(L)$.

$$\phi_i^{\dagger} = \sum_{j=1}^{i} \phi_j - 1,\ i = 1, 2, 3, \tag{19}$$

$$\beta_i^{\dagger} = \sum_{j=0}^{i} \beta_j,\ i = 1, 2, 3.$$

Alternatively,put the level-terms at the first-lag

$$\Delta y_t = \phi_0 + \sum_{i=1}^{3} \phi_i^{\ddagger} \Delta y_{t-i} + \sum_{i=0}^{3} \beta_i^{\ddagger} \Delta x_{t-i} \tag{20}$$
$$+ (\phi(1) - 1) y_{t-1} + \beta(1) x_{t-1} + \varepsilon_t$$

Equilibrium correction model V

$$\phi_i^{\ddagger} = -\sum_{j=i+1}^{4} \phi_j,\ i = 1, 2, 3,$$
$$\beta_0^{\ddagger} = \beta_0 \tag{21}$$
$$\beta_i^{\ddagger} = -\sum_{j=i+1}^{4} \beta_j,\ i = 1, 2, 3.$$

In both cases, the long-run multiplier with respect to x_t is

$$K_1 = \frac{\beta(1)}{1 - \phi_1(1)} = \frac{\sum_{j=0}^{4} \beta_j}{(1 - \sum_{j=1}^{4} \phi_j)} \tag{22}$$

and

$$\beta_0^{\dagger} = \beta_0^{\ddagger} = \beta_0, \tag{23}$$

Equilibrium correction model VI

but the other parameters are *not* the same in the two versions.

$$\phi_i^{\dagger} \neq \phi_i^{\ddagger},\ i = 1, 2, 3, \tag{24}$$

$$\beta_i^{\dagger} \neq \beta_i^{\ddagger},\ i = 1, 2, 3. \tag{25}$$

- ECMs are even more flexible than this.
  - The AR lag length and the DL lag length need not be the same.
  - The levels of y and x can be on different lags.
- The long run multiplier K_1 is invariant to the different ways of writing the ECM.
- But the interim multipliers are affected (as illustrated)

Matrix notation for ECM I

Matrix-notation for the ECM with one x_t variable (m lags):

$$\Delta y_t = \mathbf{z}_{1t}'\mathbf{b}_1 + \mathbf{z}_{2t}'\mathbf{b}_2 + \varepsilon_t \tag{26}$$

$$\mathbf{z}_{1t}' = (1, \Delta y_{t-1}, \Delta y_{t-2}, \ldots, \Delta y_{t-p+1}, \Delta x_t, \Delta x_{t-1}, \ldots, \Delta x_{t-m+1}),$$
$$\mathbf{z}_{2t}' = (y_{t-p,} x_{t-m}),$$

$$\mathbf{b}_1 = \begin{pmatrix} \phi_0 \\ \phi_1^\dagger \\ \vdots \\ \phi_{p-1}^\dagger \\ \beta_0^\dagger \\ \vdots \\ \beta_{m-1}^\dagger \end{pmatrix}$$

Matrix notation for ECM II

$$\mathbf{b}_2 = \begin{pmatrix} (\phi(1) - 1) \\ \beta(1) \end{pmatrix} = \begin{pmatrix} \phi^* \\ \beta^* \end{pmatrix}.$$

(26) is a 1-1 linear transform of the $ARDL(p, m)$ and therefore all results for the OLS estimators also holds for $\hat{\mathbf{b}}_1$and $\hat{\mathbf{b}}_2$.
The estimate of the long-run multiplier K_1 is non-linear

$$\hat{K}_1 = \frac{\widehat{\beta(1)}}{-(\widehat{\phi(1) - 1})} \equiv \frac{\hat{\beta}^*}{-\hat{\phi}^*},$$

Matrix notation for ECM III

However, $\sqrt{\text{Var}[\hat{K}_1]}$can be obtained by the so called Delta-method, Bårdsen (1989):

$$\widehat{\text{Var}[\hat{K}_1]} \approx \left(\frac{1}{-\hat{\phi}^*}\right)^2 \widehat{\text{Var}(\hat{\beta}^*)} + \left(\frac{\hat{\beta}^*}{(-\hat{\phi}^*)^2}\right)^2 \widehat{\text{Var}(\hat{\phi}^*)} \qquad (27)$$
$$+ 2\left(\frac{1}{-\hat{\phi}^*}\right)\left(\frac{\hat{\beta}^*}{(-\hat{\phi}^*)^2}\right)\widehat{\text{Cov}(\hat{\beta}^*, \hat{\phi}^*)}.$$

There are text book presentations in Greene (2012), Hill, Griffiths and Lim (2012) and in Bårdsen and Nymoen (2011)

Matrix notation for ECM IV

This also applies to an ECM with k explanatory variables x_{jt} (with lag order m_j), i.e. the approximate variance of

$$\hat{K}_j = \frac{\widehat{\beta_j(1)}}{-(\widehat{\phi(1) - 1})} \equiv \frac{\hat{\beta}_j^*}{-\hat{\phi}^*},$$

is given by (27), with an obvious change in notation.

- $\sqrt{\text{Var}[\hat{K}_j]}$ is available directly in PcGive ("Dynamic analysis" in test-menu) and in Stata (for example "NLCOM").

The VAR interpreted as a unrestricted reduced form I

Consider again the bivariate case of, x_t and y_t and dynamics of the first order:
The simultaneous equations representation of this dynamic system is

$$\begin{bmatrix} 1 & b_{12,0} \\ b_{21,0} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} b_{11,1} & b_{12,1} \\ b_{21,1} & b_{22,1} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{y,t} \\ \epsilon_{x,t} \end{bmatrix} \tag{28}$$

where $\varepsilon_{y,t}$ and $\varepsilon_{x,t}$ are uncorrelated Gaussian disturbances.
If we start from (28), the VAR in (1) is seen to be a reduced form of the simultaneous equations model:

$$\underbrace{\begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix}}_{\text{in (1)}} = \begin{bmatrix} 1 & b_{12,0} \\ b_{21,0} & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_{11,1} & b_{12,1} \\ b_{21,1} & b_{22,1} \end{bmatrix}$$

The VAR interpreted as a unrestricted reduced form II

and

$$\underbrace{\begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \end{bmatrix}}_{\text{in (1)}} = \begin{bmatrix} 1 & b_{12,0} \\ b_{21,0} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \epsilon_{y,t} \\ \epsilon_{x,t} \end{bmatrix}.$$

- Since the simultaneous equation model (28) is "unrestricted", in fact it is not identified, we call the reduced form of that model the *unrestricted reduced form*, URF .
- The VAR disturbances $\varepsilon_{y,t}$ and $\varepsilon_{x,t}$ are correlated even if (as here) $\epsilon_{y,t}$ and $\epsilon_{x,t}$ are uncorrelated ($b_{12,0} \neq 0$, or $b_{21,0} \neq 0$).

The VAR interpreted as a unrestricted reduced form III

- This removes the interpretability of the impulse responses since for example
$$\frac{\partial y_{t+s}}{\partial \varepsilon_{y,t}}$$
is in general *not* interpretable as the effect on y_{t+s} of a shock to y in period t.
- This shows that the unrestricted VAR is *not a structural model*.
- It would not help to estimate (28) in this case, since neither equation is identified on the order condition.
- In fact, the under-identification of the simultaneous equation model (28) and the non-structural VAR is one and the same thing: It is all about *lack of identification*.

The VAR interpreted as a unrestricted reduced form IV

- The VAR has unidentified impulse-responses
- The simultaneous equations model (28) is not identified on the rank and order conditions for a simultaneous equation model.

SVARS and identified systems of equations model (SEM) I

- One popular way to identify the impulse-response, thus obtaining a *structural* VAR, a SVAR, is to replace $\varepsilon_{y,t}$ and $\varepsilon_{x,t}$ by a pair of uncorrelated disturbances (orthogonalized innovations in Hamilton's terminology).
  - This is called the Cholesky decomposition/factorization (the theorem is in H Ch 4.4 p 91-91), while the application to our case is on page 320 (orthogonalization) and 327-330 (identification of impulse-responses)
  - The Cholesky factorization is equivalent to choosing a *recursive* system of equations model.
- So could estimate the *recursive* system of equations in the first place.

SVARS and identified systems of equations model (SEM) II

  - Note: although clearly resctrictive the recursive system is not identified on the rank/order condition, but this is because the rank/order condition does not assume anything about the correlation between the structural disturbances ϵ_{xt} and ϵ_{yt}.
- Other ways to identify the VAR impulse-responses involves more complicated operations on the VAR covariance matrix. See the last sections of Hamilton's Ch 11.
- Such restrictions are also equivalent to restrictions on the contemporaneous and/or lagged coefficients of the SEM representation of the system.

The VAR system and dynamic SEM I

- We now switch attention back to dynamic systems and dynamic models of the system such as a the bivariate bivariate open-VAR (also called VARX)

$$\underbrace{\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}}_{\mathbf{y}_t} = \underbrace{\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix}}_{\mathbf{\Pi}_1} \underbrace{\begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix}}_{\mathbf{y}_{t-1}} + \underbrace{\begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}}_{\mathbf{\Gamma}_1} \underbrace{\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix}}_{\mathbf{x}_t} + \underbrace{\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}}_{\boldsymbol{\varepsilon}_t} \tag{29}$$

The VAR system and dynamic SEM II

Generalizations to higher dimensions (more variables) and longer lags are unproblematic, as long as $\mathbf{y}_t$ and $\mathbf{z}_t$ are covariance stationary:

$$\mathbf{y}_t = \sum_{i=0}^{p} \mathbf{\Pi}_i \mathbf{y}_{t-i} + \sum_{i=0}^{q} \mathbf{\Gamma}_i \mathbf{x}_{t-i} + \boldsymbol{\varepsilon}_t \tag{30}$$

$$\boldsymbol{\varepsilon}_t = IIN(\mathbf{0}, \mathbf{\Sigma}) \tag{31}$$

The VAR system and dynamic SEM III

- As we have seen, dynamic linear structural models of (30)-(31) can be obtained by pre-multiplying (30) by a non-singular matrix $\mathbf{B}$:

$$\mathbf{B}\mathbf{y}_t = \sum_{i=0}^{p} \mathbf{B}\mathbf{\Pi}_i \mathbf{y}_{t-i} + \sum_{i=0}^{q} \mathbf{B}\mathbf{\Gamma}_i \mathbf{z}_{t-i} + \mathbf{B}\boldsymbol{\varepsilon}_t \qquad (32)$$

  so that the structural coefficients are in $\mathbf{B}$, $\mathbf{B\Pi}_i$, $\mathbf{B\Gamma}_i$ and the vector of structural disturbances are $\boldsymbol{\epsilon}_t = \mathbf{B}\boldsymbol{\varepsilon}_t$ with $E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t^{'}) = \mathbf{\Omega}$.
- Assume *just identification*, or *over-identification* of (32)
- Remember from ECON 4160 or equivalent: as long as the covariance matrix $\mathbf{\Omega}$ is unrestricted, checking the rank condition for identification is sufficient.

FIML estimation of an exactely identified SEM I

- Since the model is exactly identified, full information maximum likelihood (FIML) estimation can be obtained by:
  - First, maximize the likelihood function of the VAR: Under the Gaussian assumption: OLS equation by equation.
  - Second, Use the unique mapping from reduced form parameters to the (structural) parameters of the simultaneous equations to derive the FIML estimators of the structural parameters
- Some authors call this the *indirect maximum likelihood estimator*
- Note that the essential premise here is that there is no difference between the *Unrestricted Reduced Form* (URF), the VAR, and the *Restricted Reduced Form* (RRF) which is the solution of the over-identified SEM.

FIML estimation of an over-identified SEM I

- In the case of over-identification, the maximum likelihood estimator of the RRF parameters $\mathbf{\Sigma}$, $\mathbf{\Pi}_i$ and $\mathbf{\Gamma}_i$ become complicated non-linear functions of the structural parameters.
- Trying to solve that highly non-linear maximization problem slowed down the progress of econometrics, because FIML seemed to complicated to apply with existing facilities.
- However, that point is now behind us, and FIML is easy to do in practice, using for example PcGive or Stata.

FIML estimation of an over-identified SEM II

- Thefore, over-identifying restrictions can be tested by comparing the URF likelihood with the maximized likehood of the RRF. The LR test-statistic

  $$-2(L_{RRF} - L_{URF})$$

  is Chi-squared distributed with degrees of freedom equal to the degree of over-identification, which is part of the PcGive output for example

2SLS and 3SLS I

- 2SLS estimation ignores any off-diagonal elements of the SEM disturbance covariance matrix $\mathbf{\Omega}$ and is sometimes called a Limited Information Maximum Likelihood Estimator of the structural parameters.
- 2SLS estimation consistent for $\mathbf{B}$, $\mathbf{B\Pi}_i$, $\mathbf{B\Gamma}_i$, but inefficient if $\mathbf{\Omega}$ has non-zero off diagonal elements.
- 3SLS stems from the era when FIML could not be done in practice. Today it is sometimes used to get starting values for FIML

Looking ahead I

- Next week: *Seminar* exercises about modelling of the causal VAR
- Lecture: Econometric models of *non-causal* VARs.
- Main example will be structural equations with $t+1$ dated expectation variables
- New Keynesian Phillips curves and DSGE models the economic example

References

Angrist, J., K. Graddy and G. W. Ibens (2000): The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with the Application to the Demand for Fish, *Review of Economic Studies*, 67, 499-527

Bårdsen, G.(1989): "Estimation of Long-Run Coefficients in Error-Correction Models" *Oxford Bulletin of Economics and Statistics*, 51, 345-350.

Bårdsen, G og Nymoen (2014) Videregående emner i økonometri

Greene, W.H. (2012) *Econometric Analysis*. 7th Edition. Pearson International

Hendry, D. F. (1995) *Dynamic Econometrics*, Oxford University Press

Hill, R.C., W.G. Griffiths and G. C. Lim (2012). Principles of Econometrics. 4th Edition. Wiley.

Johansen, S. (1995) *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press