Università Commerciale Luigi Bocconi

MSc. Finance/CLEFIN
2017/2018 Edition

FINANCIAL ECONOMETRICS AND EMPIRICAL FINANCE - MODULE 2

Mock Question 2 (total 17 points, out of 50 from 3 questions)

Time Advised: 24 minutes (for this question)

Question 2.A (14 points)

Consider a bivariate VAR(2) model for S&P 500 returns and the log changes in the VIX volatility index ($R_t^{S\&P}$ and $\Delta lnVIX_t$). Write:

- The structural, unconstrained VAR(2) that includes contemporaneous effects across the two markets.
- The associated unconstrained reduced-form VAR(2).

Explain through which steps it is possible to transform the structural VAR model into the reduced-form one (algebra is not required, unless it helps you provide an efficient answer). How would/could you estimate the structural VAR? How would/could you estimate the reduced-form model? Explain what are the issues/limitations caused by the transformation of a structural VAR into a reduced-form model.

Debriefing:

Vector Autoregressions: Reduced-Form vs. Structural

- What is the difference? Consider the simple $N = 2$, $p = 1$ case of

$$x_t = \gamma_{10} - b_{12} z_t + \gamma_{11} x_{t-1} + \gamma_{12} z_{t-1} + \epsilon_t^x$$
$$z_t = \gamma_{20} - b_{21} x_t + \gamma_{21} x_{t-1} + \gamma_{22} z_{t-1} + \epsilon_t^z$$

where both x_t and z_t are stationary, ϵ^x_t and ϵ^z_t are uncorrelated white noise processes, also called structural errors

- Using matrices, this VAR(1) model may be re-written as:

Structural VAR
$$\underbrace{\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} x_t \\ z_t \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} \gamma_{10} \\ \gamma_{20} \end{bmatrix}}_{\Gamma_0} + \underbrace{\begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}}_{\Gamma_1} \underbrace{\begin{bmatrix} x_{t-1} \\ z_{t-1} \end{bmatrix}}_{y_{t-1}} + \underbrace{\begin{bmatrix} \epsilon_t^x \\ \epsilon_t^z \end{bmatrix}}_{\epsilon_t}$$
$$By_t = \Gamma_0 + \Gamma_1 y_{t-1} + \epsilon_t$$

- Pre-multiplying both sides by B^{-1} (this will be possible if $b_{12}b_{21} \neq 1$),

$$y_t = \underbrace{B^{-1}\Gamma_0}_{a_0} + \underbrace{B^{-1}\Gamma_1}_{A_1} y_{t-1} + \underbrace{B^{-1}\epsilon_t}_{u_t} = a_0 + A_1 y_{t-1} + u_t$$

Reduced-form VAR

Lecture 4: Multivariate Time Series Analysis– Prof. Guidolin 10

2 Introduction to VAR Analysis

2.1 From Structural to Reduced-Form VARs

Vector Autoregressive Model Var(p): A Vector Autoregressive model of order p is a process that can be represented as

$$\mathbf{y_t} = \mathbf{a_0} + \mathbf{A_1 y_{t-1}} + \mathbf{A_2 y_{t-2}} + ... + \mathbf{A_p y_{t-p}} + \mathbf{u_t} = \mathbf{a_0} + \sum_{j=1}^{p} \mathbf{A_j y_{t-j}} + \mathbf{u_t}$$

where
$\mathbf{y}_t = N \times 1$ vector containing N endogenous variables
$\mathbf{a}_0 = N \times 1$ vector of constants
$\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_p = p\ N \times N$ matrices of autoregressive coefficients $\mathbf{u}_t$
$= N \times 1$ vector of **serially** uncorrelated, white noise disturbances.

Structural VAR or **VAR in primitive form:**

$$y_{1,t} = b_{1,0} - b_{1,2} y_{2,t} + \varphi_{1,1} y_{1,t-1} + \varphi_{1,2} y_{2,t-1} + \epsilon_{1,t}$$
$$y_{2,t} = b_{2,0} - b_{2,1} y_{1,t} + \varphi_{2,1} y_{1,t-1} + \varphi_{2,2} y_{2,t-1} + \epsilon_{2,t}$$

where
$y_{1,t}$ and $y_{2,t}$ are assumed to be stationary
$\epsilon_{1,t}$ and $\epsilon_{2,t}$ are uncorrelated white-noise disturbances with standard deviation σ_1 and σ_2, respectively.

In matrix notation

$$\begin{bmatrix} 1 & \mathrm{b}_{1,2} \\ \mathrm{b}_{2,1} & 1 \end{bmatrix} \begin{bmatrix} \mathrm{y}_{1,t} \\ \mathrm{y}_{2,t} \end{bmatrix} = \begin{bmatrix} \mathrm{b}_{1,0} \\ \mathrm{b}_{2,0} \end{bmatrix} + \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} \\ \varphi_{2,1} & \varphi_{2,2} \end{bmatrix} \begin{bmatrix} \mathrm{y}_{1,t-1} \\ \mathrm{y}_{2,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

or in compact form

$$\mathbf{By_t} = \mathbf{Q_0} + \mathbf{Q_1 y_{t-1}} + \epsilon_t$$

- $y_{1,t}$ depends on its own lag and on both one lag and current value of $y_{2,t}$; $y_{2,t}$ depends on its own lag and on both one lag and current value of $y_{1,t}$.
- It captures contemporaneous feedback effects:
  1. $-b_{1,2}$ measures the contemporaneous effect of a unit change of $y_{2,t}$ on $y_{1,t}$;

5

Vector Autoregressions: Reduced-Form vs. Structural

- The "new" error terms are composites of the two primitive shocks:

$$\begin{bmatrix} u_t^x \\ u_t^z \end{bmatrix} = u_t \equiv B^{-1}\epsilon_t = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \epsilon_t^x \\ \epsilon_t^z \end{bmatrix}$$

$$= \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \epsilon_t^x \\ \epsilon_t^z \end{bmatrix} = \begin{bmatrix} \frac{\epsilon_t^x - b_{12}\epsilon_t^z}{1-b_{12}b_{21}} \\ \frac{\epsilon_t^z - b_{21}\epsilon_t^x}{1-b_{12}b_{21}} \end{bmatrix}$$

- What are the properties of the reduced form errors? Recall that ϵ^x_t and ϵ^z_t were uncorrelated, white noise processes, then:

$$E\begin{bmatrix} u_t^x \\ u_t^z \end{bmatrix} = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} E[\epsilon_t^x - b_{12}\epsilon_t^z] \\ E[\epsilon_t^z - b_{21}\epsilon_t^x] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (\text{as } E[\epsilon_t^x] = E[\epsilon_t^z] = 0)$$

$$Var[u_t^x] = \frac{1}{(1-b_{12}b_{21})^2} Var[\epsilon_t^x - b_{12}\epsilon_t^z]$$

$$= \frac{1}{(1-b_{12}b_{21})^2}\{Var[\epsilon_t^x] + b_{12}^2 Var[\epsilon_t^z] - 2b_{12}Cov[\epsilon_t^x, \epsilon_t^z]\}$$

$$= \frac{\sigma_x^2 + b_{12}^2\sigma_z^2}{(1-b_{12}b_{21})^2} \quad \left(\text{similarly, } Var[u_t^z] = \frac{\sigma_z^2 + b_{21}^2\sigma_x^2}{(1-b_{12}b_{21})^2}\right)$$

Vector Autoregressions: Reduced-Form vs. Structural

- The reduced-form shocks u^x_t and u^z_t will be correlated even though the structural shocks are not:

$$Cov[u_t^x, u_t^z] = E[u_t^x u_t^z] = \frac{1}{(1-b_{12}b_{21})^2} E[(\epsilon_t^x - b_{12}\epsilon_t^z)(\epsilon_t^z - b_{21}\epsilon_t^x)]$$

$$= \frac{1}{(1-b_{12}b_{21})^2} E[\underbrace{\epsilon_t^x\epsilon_t^z}_{=0} - b_{21}(\epsilon_t^x)^2 - b_{12}(\epsilon_t^z)^2 + b_{12}b_{21}\underbrace{\epsilon_t^z\epsilon_t^x}_{=0}]$$

$$= \frac{1}{(1-b_{12}b_{21})^2} E[-b_{21}(\epsilon_t^x)^2 - b_{12}(\epsilon_t^z)^2] = -\frac{b_{21}\sigma_x^2 + b_{12}\sigma_z^2}{(1-b_{12}b_{21})^2}$$

which shows that they are uncorrelated if $b_{12} = b_{21} = 0$

- This is very important: unless the variables are contemporaneously uncorrelated in the structural VAR ($b_{12} = b_{21} = 0$), a reduced-form VAR will generally display correlated shocks
  - If VARs are just extensions of AR models under what conditions will they be stationary? Stay tuned...

2. $-b_{2,1}$ measures the contemporaneous effect of a unit change of $y_{1,t}$ on $y_{2,t}$.

- Each contemporaneous variable is correlated with its own error term, therefore the regressors are not uncorrelated with the error terms as required by OLS estimation techniques.
- When $-b_{1,2} \neq 0$, $y_{2,t}$ depends on $y_{1,t}$ and on $\epsilon_{1,t}$ and will be correlated with it. When $-b_{2,1} \neq 0$, $y_{1,t}$ depends on $y_{2,t}$ and on $\epsilon_{2,t}$.
- Contemporaneous terms cannot be used in forecasting.

Reduced-form VAR or VAR in standard form:

$$y_{1,t} = a_{1,0} + a_{1,1}y_{1,t-1} + a_{1,2}y_{2,t-1} + u_{1,t}$$

$$y_{2,t} = a_{2,0} + a_{2,1}y_{1,t-1} + a_{2,2}y_{2,t-1} + u_{2,t}$$

that is obtained from by pre-multiplying both sides of $\mathbf{B}\mathbf{y_t} = \mathbf{Q_0} + \mathbf{Q_1}\mathbf{y_{t-1}} + \epsilon_t$ by $\mathbf{B}^{-1}$

$$\mathbf{y_t} = \mathbf{a_0} + \mathbf{A_1}\mathbf{y_{t-1}} + \mathbf{u_t}$$

where $\mathbf{a_0} = \mathbf{B}^{-1}\mathbf{Q_0}$, $\mathbf{A_1} = \mathbf{B}^{-1}\mathbf{Q_1}$, $\mathbf{u_t} = \mathbf{B_t}^{-1}\epsilon_\mathbf{t}$.

- It does not contain contemporaneous feedback terms.
- It can be estimated equation by equation using OLS.
- $u_{1,t}$ and $u_{2,t}$ are composites of $\epsilon_{1,t}$ and $\epsilon_{2,t}$: in fact

$$\mathbf{u_t} = \mathbf{B}^{-1}\epsilon_\mathbf{t} \quad \text{then}$$

$$u_{1,t} = \frac{\epsilon_{1,t} - b_{1,2}\epsilon_{2,t}}{1 - b_{1,2}b_{2,1}} \quad \text{and} \quad u_{2,t} = \frac{\epsilon_{2,t} - b_{2,1}\epsilon_{1,t}}{1 - b_{1,2}b_{2,1}}$$

Properties (derived by the white noise processes $\epsilon_{1,t}$, $\epsilon_{2,t}$):

1.

$$E[u_{1,t}] = E[\frac{\epsilon_{1,t} - b_{1,2}\epsilon_{2,t}}{1 - b_{1,2}b_{2,1}}] = 0$$

$$E[u_{2,t}] = E[\frac{\epsilon_{2,t} - b_{2,1}\epsilon_{1,t}}{1 - b_{1,2}b_{2,1}}] = 0$$

2.

$$Var[u_{1,t}] = \frac{Var[\epsilon_{1,t} - b_{1,2}\epsilon_{2,t}]}{(1 - b_{1,2}b_{2,1})^2} =$$

$$= \frac{Var[\epsilon_{1,t}] + b_{1,2}^2 Var[\epsilon_{2,t}] - 2b_{1,2}Cov[\epsilon_{1,t}, \epsilon_{2,t}]}{(1 - b_{1,2}b_{2,1})^2} = \frac{\sigma_{\epsilon,1}^2 + b_{1,2}^2\sigma_{\epsilon,2}^2}{(1 - b_{1,2}b_{2,1})^2}$$

$$Var[u_{2,t}] = \frac{\sigma_{\epsilon,2}^2 + b_{2,1}^2\sigma_{\epsilon,1}^2}{(1 - b_{1,2}b_{2,1})^2}$$

constant over time.

3.

$$Cov[u_{1,t}, u_{2,t}] = \frac{E[(\epsilon_{1,t} - b_{1,2}\epsilon_{2,t})(\epsilon_{2,t} - b_{2,1}\epsilon_{1,t})]}{(1 - b_{1,2}b_{2,1})^2} = \frac{-(b_{2,1}\sigma_{\epsilon,1}^2 + b_{1,2}\sigma_{\epsilon,2}^2)}{(1 - b_{1,2}b_{2,1})^2}$$

- $u_{1,t}$ and $u_{2,t}$ are serially uncorrelated, but are cross-correlated unless $b_{1,2} = b_{2,1} = 0$.

4.

$$\Sigma_u = \begin{bmatrix} \text{Var}[u_{1,t}] & \text{Cov}[u_{1,t}, u_{2,t}] \\ \text{Cov}[u_{1,t}, u_{2,t}] & \text{Var}[u_{2,t}] \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{bmatrix}$$

- In general, it is not possible to identify the structural parameters and errors from the OLS estimates of the parameters and the residuals of the standard form VAR, unless some restrictions are imposed on the primitive system.

Recursive Choleski triangularization: Impose a Choleski decomposition on the covariance matrix of the residuals of the VAR in its standard form, that is the restriction $b_{1,2} = 0$, so that

$$y_{1,t} = b_{1,0} + \varphi_{1,1}y_{1,t-1} + \varphi_{1,2}y_{2,t-1} + \epsilon_{1,t}$$

$$y_{2,t} = b_{2,0} - b_{2,1}y_{1,t} + \varphi_{2,1}y_{1,t-1} + \varphi_{2,2}y_{2,t-1} + \epsilon_{2,t}$$

then

$$u_{1,t} = \epsilon_{1,t} \quad \text{and} \quad u_{2,t} = \epsilon_{2,t} - b_{2,1}\epsilon_{1,t}$$

In matrix form, the restriction $b_{1,2} = 0$ means that $\mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 \\ \text{-b}_{2,1} & 1 \end{bmatrix}$, so that

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \text{-b}_{2,1} & 1 \end{bmatrix} \begin{bmatrix} b_{1,0} \\ b_{2,0} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ \text{-b}_{2,1} & 1 \end{bmatrix} \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} \\ \varphi_{2,1} & \varphi_{2,2} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} +$$

$$+ \begin{bmatrix} 1 & 0 \\ \text{-b}_{2,1} & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

that minimizes

$$S(\beta) = \mathbf{u}'(\mathbf{I}_N \Sigma_u)^{-1}\mathbf{u}$$

- When a reduced-form VAR is unconstrained, the GLS estimator is the same as the OLS estimator, $\hat{\mathbf{B}}$, and therefore an unconstrained VAR can be estimated equation by equation by OLS.

Asymptotic properties of the OLS estimator $\hat{\mathbf{B}}$ (under standard assumptions):

(a) Consistent and asymptotically normally distributed

$$\sqrt{T}vec(\hat{\mathbf{B}} - \mathbf{B}) \xrightarrow{D} N(0, \Sigma_{\hat{\mathbf{B}}}) \quad \text{or} \quad vec(\hat{\mathbf{B}}) \overset{a}{\sim} N(vec(\mathbf{B}), \Sigma_{\hat{\mathbf{B}}}/T)$$

where $\Sigma_{\hat{\mathbf{B}}} = plim(\mathbf{ZZ}'/T)^{-1} \otimes \Sigma_u$.

(b)

$$\hat{\Sigma}_u = \frac{1}{T - Np}\sum_{t=1}^{T} \hat{\mathbf{u}}_t\hat{\mathbf{u}}_t' \quad \text{or} \quad \hat{\Sigma}_u = \frac{1}{T}\sum_{t=1}^{T} \hat{\mathbf{u}}_t\hat{\mathbf{u}}_t'$$

where $\hat{\mathbf{u}}_t = \mathbf{y}_t - \mathbf{BZ}_{t-1}$.

Multivariate ML estimator:

Under the assumptions:

(a) Sample of T observations on $\mathbf{Y}$ and a pre-sample of p initial conditions $y_{-p+1}, y_{-p+2}, ..., y_0$.

(b) Stationary process and Gaussian multivariate white noise innovations.
$\Rightarrow$ $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_T]'$ is jointly normally distributed.

(c) Gaussian multivariate white noise (then innovations at different times are independent).

(d) Noise error terms are independent with Σ_u, then the covariance matrix of $\mathbf{u}$ is $\Sigma_U = \mathbf{I}_T \otimes \Sigma_u$ and its normal density is
$f_u(\mathbf{u}) = (2\pi)^{-\frac{NT}{2}}|\mathbf{I}_T \otimes \Sigma_u|^{-\frac{1}{2}}exp(-\frac{1}{2}\mathbf{u}'(\mathbf{I}_T \otimes \Sigma_u^{-1})\mathbf{u})$.

(e) $f_y(\mathbf{y}) = (2\pi)^{-\frac{NT}{2}}|\mathbf{I}_T \otimes \Sigma_u|^{-\frac{1}{2}}exp(-\frac{1}{2}(\mathbf{Y} - \mathbf{BZ})'(\mathbf{I}_T \otimes \Sigma_u^{-1})(\mathbf{Y} - \mathbf{BZ}))$.

the ML estimator maximizes

$$\ell(\mathbf{B}, \Sigma_u; \mathbf{Y}, \mathbf{Z}) = ln f_y(\mathbf{Y}) = -\frac{NT}{2}ln(2\pi) - \frac{T}{2}ln|\Sigma_u| - \frac{1}{2}(\mathbf{Y}-\mathbf{BZ})'(\mathbf{I}_T \otimes \Sigma_u^{-1})(\mathbf{Y}-\mathbf{BZ}) =$$

2.4 Estimation of a VAR(p) Model

Multivariate LS estimator:

Starting from

$$\mathbf{Y} = \mathbf{BZ} + \mathbf{U}$$

where $\mathbf{Y} \equiv [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_t]$, $\mathbf{B} \equiv [\mathbf{a}_0, \mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_p]$, $\mathbf{U} \equiv [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_T]$, $\mathbf{Z} \equiv [\mathbf{Z}_0, \mathbf{Z}_1, \mathbf{Z}_{T-1}]$ with $\mathbf{Z}_t \equiv [\mathbf{1}', \mathbf{y}_{t-1}', \mathbf{y}_{t-2}', ..., \mathbf{y}_{t-p+1}']'$. Given that $\mathbf{y} \equiv vec(\mathbf{Y})$, $\beta \equiv vec(\mathbf{B})$ and $\mathbf{u} \equiv vec(\mathbf{U})$ the multivariate LS estimator is

$$\hat{\beta} = ((\mathbf{ZZ}')^{-1} \otimes \Sigma_u)(\mathbf{Z} \otimes \Sigma_u^{-1})\mathbf{y} = ((\mathbf{ZZ}')^{-1}\mathbf{Z} \otimes \mathbf{I}_N)\mathbf{y}$$

$$= \begin{bmatrix} \mathrm{b}_{1,0} \\ \mathrm{b}_{2,0} - b_{1,0}b_{2,1} \end{bmatrix} + \Big] \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} \\ \varphi_{2,1} - b_{2,1}\varphi_{1,1} & \varphi_{2,2} - b_{2,1}\varphi_{1,2} \end{bmatrix} \begin{bmatrix} \mathrm{y}_{1,t-1} \\ \mathrm{y}_{2,t-1} \end{bmatrix} +$$

$$+ \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} - b_{2,1}\epsilon_{1,t} \end{bmatrix}$$

so that

$$a_{1,0} = b_{1,0}, \quad a_{2,0} = b_{2,0} - b_{1,0}b_{2,1}, \quad a_{1,1} = \varphi_{1,1,}, \quad a_{1,2} = \varphi_{1,2},$$

$$a_{2,1} = \varphi_{2,1} - b_{2,1}\varphi_{1,1}, \quad a_{2,2} = \varphi_{2,2} - b_{2,1}\varphi_{1,2}, \quad u_{1,t} = \epsilon_{1,t}, \quad u_{2,t} = \epsilon_{2,t} - b_{2,1}\epsilon_{1,t}$$

It follows that

$$\sigma_1^2 \equiv Var[u_{1,t}] = \sigma_{\epsilon,1}^2$$

$$\sigma_2^2 \equiv Var[u_{2,t}] = \sigma_{\epsilon,2}^2 - b_{2,1}^2\sigma_{\epsilon,1}^2$$

$$Cov[u_{1,t}, u_{2,t}] = -b_{2,1}\sigma_{\epsilon,1}^2$$

- The restriction implies that the observed values of $u_{1,t}$ are completely attributed to pure (structural) shocks to $y_{1,t}$.

Question 2.B (2 points)

Suppose that the bivariate structural VAR(2) is to be exactly identified by imposing either of the two possible Choleski triangularization schemes:

$$\boldsymbol{B}' = \begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{B}'' = \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}$$

Carefully explain the implications and differences in economic interpretations of the estimated, corresponding reduced-form model deriving from imposing the restriction in $\boldsymbol{B}'$ instead of $\boldsymbol{B}''$. How does your answer change when the restriction

$$\boldsymbol{B}''' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is imposed instead?

Debriefing:

Trivially, $\boldsymbol{B}'''$ implies that the original structural model is in reduced form or, alternatively, the model has been over-identified by removing all contemporaneous effects between variables.
$\boldsymbol{B}'$ implies that S&P 500 returns are ordered before VIX log-changes, so that any $u_t^{S\&P}$ shock to the S&P 500 is structural and primitive, while the u_t^{VIX} shocks are correlated with both structural shocks to S&P 500 and the VIX.
On the opposite, $\boldsymbol{B}''$ implies that log changes in VIX are order before S&P 500, so that any u_t^{VIX} shock to the VIX is structural and primitive, while the $u_t^{S\&P}$ shocks are correlated with both structural shocks to S&P 500 and the VIX.

Identifying Structural from Reduced-Form VARs

- In a sense, shocks to z_t are more primitive, enjoy a higher rank, and move the system also through a contemporaneous impact on x_t
- The VAR(1) now acquires a triangular structure:

$$x_t = \gamma_{10} - b_{12} z_t + \gamma_{11} x_{t-1} + \gamma_{12} z_{t-1} + \epsilon_t^x$$
$$z_t = \gamma_{20} + \gamma_{21} x_{t-1} + \gamma_{22} z_{t-1} + \epsilon_t^z$$
$$y_t = \underbrace{B^{-1}\Gamma_0}_{\mathbf{a}_0} + \underbrace{B^{-1}\Gamma_1}_{A_1} y_{t-1} + \underbrace{B^{-1}\epsilon_t}_{u_t} = \mathbf{a}_0 + A_1 y_{t-1} + u_t \qquad \text{with } B \equiv \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}$$

- This corresponds to imposing a Choleski decomposition on the covariance matrix of the residuals of the VAR in its reduced form
- Indeed, now we can re-write the relationship between the pure shocks (from the structural VAR) and the regression residuals as

$$\begin{bmatrix} u_t^x \\ u_t^z \end{bmatrix} = u_t \equiv B^{-1}\epsilon_t = \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \epsilon_t^x \\ \epsilon_t^z \end{bmatrix} = \begin{bmatrix} \epsilon_t^x - b_{12}\epsilon_t^z \\ \epsilon_t^z \end{bmatrix}$$

Lecture 4: Multivariate Time Series Analysis– Prof. Guidolin 15

- In a N-variate VAR, we need to impose $(N^2 - N)/2$ in order to retrieve the N structural shocks from the residual of the OLS estimate.

Example for a VAR(1) with three endogenous variables:
We need to impose $(3^2 - 3)/2 = 3$ restrictions that is equivalent to pre-multiplying the structural VAR by the lower triangular matrix

$$\mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -b_{1,2} & 1 & 0 \\ -b_{1,3} & -b_{2,3} & 1 \end{bmatrix}$$

so that

$$\mathbf{u}_t = \mathbf{B}^{-1}\epsilon_t = \begin{bmatrix} 1 & 0 & 0 \\ -b_{2,1} & 1 & 0 \\ -b_{3,1} & -b_{3,2} & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \epsilon_{3,t} \end{bmatrix} = \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} - b_{2,1}\epsilon_{1,t} \\ \epsilon_{3,t} - b_{3,1}\epsilon_{1,t} - b_{3,2}\epsilon_{2,t} \end{bmatrix}$$

- There are as many Choleski decompositions as all the possible orderings of the variables. Therefore, when we apply a Choleski triangular identification scheme to a VAR model we are introducing a number of (potentially arbitrary) assumptions on the contemporaneous relationships among the variables.

Recursive Choleski Identification

- In fact, this method is quite general and extends well beyond this VAR(1), N = 2 example: in a N-variable VAR(p), **B** is a NxN matrix because there are N residuals and N structural shocks
- Exact identification requires $(N^2 - N)/2$ restrictions placed on the relationship btw. regression residuals and structural innovations
- Because a Choleski decomposition is triangular, it forces exactly $(N^2 - N)/2$ values of the **B** matrix to equal zero
  - Because with N = 2, $(2^2 - 2)/2 = 1$, you see that $b_{21} = 0$ was sufficient in our example
- There are as many Choleski decompositions as all the possible orderings of the variables, a combinatorial factor of N
  - A Choleski identification scheme results in a specific ordering, we are introducing a number of (potentially arbitrary) assumptions on the contemporaneous relationships among variables
  - Choleski decompositions are deliberate in the restrictions they place but tend not to be based on theoretical assumptions

Lecture 4: Multivariate Time Series Analysis– Prof. Guidolin 18

Question 2.C (1 point)

Suppose that the estimation of a constrained, reduced-form VAR(2) has provided the following ML estimates of the conditional mean function and of the covariance matrix of the reduced-form shocks (p-values are in parentheses):

$$\begin{cases} R_t^{S\&P} = \underset{(0.044)}{0.006} + \underset{(0.093)}{0.053} R_{t-1}^{S\&P} - \underset{(0.003)}{0.473} \Delta lnVIX_{t-1} + \underset{(0.045)}{0.113} \Delta lnVIX_{t-2} + u_t^{S\&P} \\ \Delta lnVIX_t = -\underset{(0.149)}{0.194} - \underset{(0.024)}{0.375} R_{t-1}^{S\&P} + \underset{(0.050)}{0.094} R_{t-2}^{S\&P} + \underset{(0.000)}{0.804} \Delta lnVIX_{t-1} + u_t^{VIX} \end{cases}$$

$$Var\left(\begin{bmatrix} u_t^{S\&P} \\ u_t^{VIX} \end{bmatrix}\right) = \begin{bmatrix} \underset{(0.000)}{0.008} & \underset{(0.007)}{-0.016} \\ \underset{(0.007)}{-0.016} & \underset{(0.000)}{0.014} \end{bmatrix}$$

You would like to recover the original structural parameters, including the contemporaneous, average impact of both VIX changes on S&P 500 returns and vice-versa. Is there a chance that this may be possible even though you are *not* ready to impose a Choleski ordering on the two variables?

Debriefing: One cannot say for sure but the evidence shown has two implications:

- The estimated reduced-form VAR(2) carries restrictions and in fact estimation has been properly performed by MLE applied to the bivariate system.
- There are two restrictions that have been imposed, setting the coefficients of $R_{t-2}^{S\&P}$ to zero in the first equation and the coefficient of $\Delta lnVIX_{t-2}$ to zero in the second equation; indeed note that ML estimation has been performed, because it is likely that the reduced-

form VAR will include restrictions.

Now, we can only speculate that such restrictions derive from restrictions that have been imposed on the matrix Γ_2 in the structural representation of the model,

$$\boldsymbol{B}\begin{bmatrix} R_t^{S\&P} \\ \Delta lnVIX_t \end{bmatrix} = \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_1 \begin{bmatrix} R_{t-1}^{S\&P} \\ \Delta lnVIX_{t-1} \end{bmatrix} + \boldsymbol{\Gamma}_2 \begin{bmatrix} R_{t-2}^{S\&P} \\ \Delta lnVIX_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_t^{S\&P} \\ \epsilon_t^{VIX} \end{bmatrix},$$

in the sense that

$$\boldsymbol{\Gamma}_2 = \begin{bmatrix} 0 & \gamma_{12}^2 \\ \gamma_{21}^2 & 0 \end{bmatrix}.$$

However, we know that the exact identification of a bivariate structural VAR requires imposing $(2^2 - 2)/2 = 1$ restriction and that such constraints do not have to be imposed necessarily on the matrix of contemporaneous effects $\boldsymbol{B}$**.** Because two such restrictions seem to have been imposed on $\boldsymbol{\Gamma}_2$, yes, there is a chance for the structural model—in particular for the two coefficients measuring the contemporaneous, average impact of both VIX changes on S&P 500 returns and vice-versa—to be identified (probably, over-identified), even though no Choleski triangularization has affected $\boldsymbol{B}$ (in fact, no restrictions at all have been imposed).