Eco 5316 Time Series Econometrics

Lecture 18 Structural Vector Autoregression (SVAR) Models

Motivation

- consider again the bivariate VAR(1) model for house price indices in Los Angeles and Riverside MSAs with $\mathbf{y}_t = (\Delta \log p^{LA}_{H,t}, \Delta \log p^{RI}_{H,t})$ and

$$\boldsymbol{B}_0 \mathbf{y}_t = \boldsymbol{c}_0 + \boldsymbol{B}_1 \boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

- to estimate the model, we need to convert it into a reduced form VAR(1), by premultiplying it with $\boldsymbol{B}_0^{-1}$ to obtain

$$\boldsymbol{y}_t = \boldsymbol{c} + \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \boldsymbol{e}_t$$

where $\boldsymbol{c} = \boldsymbol{B}_0^{-1}\boldsymbol{c}_0$, $\boldsymbol{A}_1 = \boldsymbol{B}_0^{-1}\boldsymbol{B}_1$, $\boldsymbol{e}_t = \boldsymbol{B}_0^{-1}\varepsilon_t$, $var(\boldsymbol{e}_t) = \boldsymbol{\Sigma}_e = \boldsymbol{B}_0^{-1}\boldsymbol{\Sigma}_\varepsilon \boldsymbol{B}_0^{-1\prime}$

- the results

	LA	RI
LA(-1)	0.801***	0.676***
	(0.100)	(0.138)
RI(-1)	0.044	0.260**
	(0.083)	(0.114)
const	0.002	−0.002
	(0.001)	(0.002)
Observations	145	145
R^2	0.729	0.648
Adjusted R^2	0.726	0.643
Residual Std. Error (df = 142)	0.015	0.021
F Statistic (df = 2; 142)	191.454***	130.770***
Note:	*p<0.1; **p<0.05; ***p<0.01	

Motivation

- we can estimate any reduced form VAR(p) model

$$\boldsymbol{y}_t = \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{A}_p \boldsymbol{y}_{t-p} + \boldsymbol{e}_t$$

  easily using equation by equation OLS
- if we are only interested in forecasting the reduced form VAR(p) model is all we need
- but to answer some other questions we need to know the coefficients of the original structural VAR(p) model

$$\boldsymbol{B}_0 \boldsymbol{y}_t = \boldsymbol{B}_1 \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{B}_p \boldsymbol{y}_{t-p} + \varepsilon_t$$

- unlike errors $\boldsymbol{e}_t$ in the reduced form model, structural shocks ε_t are not correlated, and have economic interpretation
- to construct impulse response functions (IRFs) and forecast error variance decompositions (FEVDs), $\boldsymbol{B}_0$ is needed to obtain structural innovations from reduced form VAR using $\varepsilon_t = \boldsymbol{B}_0 \boldsymbol{e}_t$

Reduced Form Errors vs Structural Shocks

note that

- IRFs trace out the response of $\boldsymbol{y}_t$ to structural shocks ε_t, not reduced form errors $\boldsymbol{e}_t$
- FEVD gives the fraction of variance of $\boldsymbol{y}_t$ caused by different structural shocks ε_t, not reduced form errors $\boldsymbol{e}_t$
- since $\varepsilon_t = \boldsymbol{B}_0 \boldsymbol{e}_t$ to construct the IRFs and FEVD for a VAR(p) model we need to know $\boldsymbol{B}_0$, in addition to $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_p$

Identification of Structural Shocks

Q: Is it possible to recover $\boldsymbol{c}_0$, $\{\boldsymbol{B}_i\}_{i=0}^{p}$ and $\boldsymbol{\Sigma}_\varepsilon$ from $\boldsymbol{c}$, $\{\boldsymbol{A}_i\}_{i=1}^{p}$ and $\boldsymbol{\Sigma}_e$?

A: Only if we are willing to impose additional restrictions.

Identification of Structural Shocks

Example: bivariate VAR(1)

- reduced form VAR(1) yields estimates of 9 parameters in $\boldsymbol{c}$, $\boldsymbol{A}_1$, $\boldsymbol{\Sigma}_e$

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} a_{1,11} & a_{1,12} \\ a_{1,21} & a_{1,22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} e_{1,t} \\ e_{2,t} \end{pmatrix} \quad \boldsymbol{\Sigma}_e = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

- we are trying to uncover $\boldsymbol{c}_0$, $\boldsymbol{B}_0$, $\boldsymbol{B}_1$, $\boldsymbol{\Sigma}_\varepsilon$ which contain 10 unknown values

$$\begin{pmatrix} 1 & b_{0,12} \\ b_{0,21} & 1 \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} c_{0,1} \\ c_{0,2} \end{pmatrix} + \begin{pmatrix} b_{1,11} & b_{1,12} \\ b_{1,21} & b_{1,22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \quad \boldsymbol{\Sigma}_\epsilon = \begin{pmatrix} \sigma_{\varepsilon_1}^2 & 0 \\ 0 & \sigma_{\varepsilon_2}^2 \end{pmatrix}$$

- one additional restriction on parameters thus needs to be *imposed* in the VAR(1)
- one possible way to do this is the **Choleski decomposition**
  - impose $b_{0,12} = 0$ so that $y_{1,t}$ has contemporaneous effect on $y_{2,t}$, but $y_{2,t}$ does not have a contemporaneous effect on $y_{1,t}$
  - this also means that both $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ have a contemporaneous effect on $y_{2,t}$, but only $\varepsilon_{1,t}$ has an effect on $y_{1,t}$
  - this is how `vars` package constructs IRFs and FEVDs

Identification of Structural Shocks

general case, a VAR(p) model with k variables

- reduced form has $k+pk^2+k(k+1)/2$ parameters
- structural form has $k+(p+1)k^2+k$ parameters
- identification thus requires $k(k-1)/2$ additional restrictions
- Choleski decomposition: set elements of $\boldsymbol{B}_0$ above main diagonal equal zero
- ordering of variables in the VAR(p) model thus matters:
  $y_{i,t}$ is only affected by shocks $\varepsilon_{1,t}, \ldots, \varepsilon_{i,t}$, remaining shocks $\varepsilon_{i+1,t}, \ldots, \varepsilon_{k,t}$ have no contemporaneous effect on $y_{i,t}$ and will only affect $y_{i,t'}$ for $t' > t$ indirectly through their effect on $y_{i+1,t}, \ldots, y_{k,t}$

Identification of Structural Shocks

- Choleski decomposition is one way to uncover $\boldsymbol{B}_0$
- if Choleski decomposition is used, ordering of variables in VAR matters for the IRFs and FEVDs
- the ordering however often does not have direct economic interpretation and is ad hoc
- we will thus look at several alternative ways of introducing restrictions consistent with some economic theory

Choleski Decomposition Approach

- Choleski decomposition: for a positive definite matrix $\boldsymbol{A}$ there exist a lower unitriangular matrix $\boldsymbol{L}$ and a diagonal matrix $\boldsymbol{D}$ such that $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}'$
- since $\varepsilon_t = \boldsymbol{B}_0 \boldsymbol{e}_t$ or equivalently $\boldsymbol{e}_t = \boldsymbol{B}_0^{-1} \varepsilon_t$ we have

$$\boldsymbol{\Sigma}_\varepsilon = \boldsymbol{B}_0 \boldsymbol{\Sigma}_e \boldsymbol{B}_0' \qquad \boldsymbol{\Sigma}_e = \boldsymbol{B}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{B}_0^{-1\prime}$$

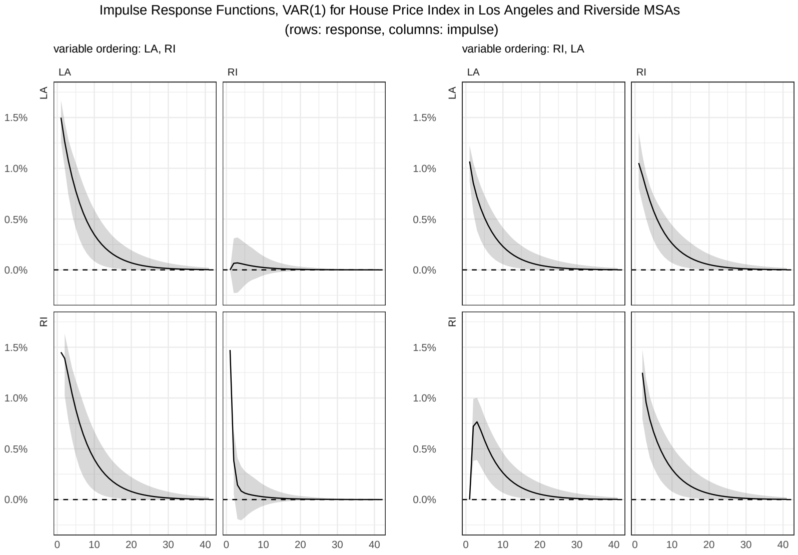
  one particular way to obtain $\boldsymbol{B}_0^{-1}$ is thus to make use of Choleski decomposition and set $\boldsymbol{B}_0^{-1} = \boldsymbol{L}$

Choleski Decomposition Approach

- if Choleski decomposition is used the elements of $\boldsymbol{B}_0$ above main diagonal are equal zero, and ordering of variables in the VAR(p) model matters:
  - $y_{i,t}$ is directly affected only by shocks $\varepsilon_{1,t}, \ldots, \varepsilon_{i,t}$
  - shocks $\varepsilon_{i+1,t}, \ldots, \varepsilon_{k,t}$ have no contemporaneous effect on $y_{i,t}$ and will only affect $y_{i,t'}$ for $t' > t$ indirectly through their effect on $y_{i+1,t}, \ldots, y_{k,t}$
- how much the order of variables matters, and how much the IRFs and FEVD change depends on the magnitude of correlation among elements of $\boldsymbol{e}_t$
- for example, in a bivariate VAR(1)
  - if $corr(e_{1,t}, e_{2,t}) = 0$ then $\varepsilon_{i,t} = e_{i,t}$ so structural shocks are identical to reduced form errors, and ordering does not matter at all
  - if $corr(e_{1,t}, e_{2,t}) = 1$ then there is actually only one structural shock and whichever variable is first determines which structural error its going to be

Choleski Decomposition Approach - Example

$$\boldsymbol{y}_t = (\Delta \log p^{LA}_{H,t}, \Delta \log p^{RI}_{H,t})' \qquad \boldsymbol{y}_t = (\Delta \log p^{RI}_{H,t}, \Delta \log p^{LA}_{H,t})'$$



Structural VARs

- **structural vector autoregressive models (SVAR)**: explicit modeling of *contemporaneous* interdependence between the left-hand side variables
- in some cases there might be a theoretical reason to expect that some variable has no contemporaneous effect on another which would give some guidance for ordering of variables
- to some extent the ordering and resulting Choleski decomposition is ad hoc, and often does not have direct economic interpretation
- it is also not practical to try all possible orderings, since there are $k!$ of them - with $k = 4$ that already means 24 different possibilities

Structural VARs

- instead of using Choleski decomposition, in some cases it is possible to use an economic theory to impose restrictions to achieve identification of parameters of the structural VAR
- **short run restrictions** - restrictions on $\boldsymbol{B}_0$ which captures the contemporaneous relationships of variables - zero (exclusion) restrictions e.g. $b_{0,12} = 0$, symmetry restrictions e.g. $b_{0,12} - b_{0,21} = 0$, other linear restrictions e.g. $b_{0,12} + b_{0,21} = 1, \ldots$
- **long run restrictions** - restrictions on $\boldsymbol{B}_0$ arise by dividing shocks into two groups - those that have a permanent effect on some variables, and those that have no permanent effects on any variable
- **sign restrictions** - restrictions on $\boldsymbol{B}_0$ which imply that IRF for some shock has certain signs at certain horizons, e.g. "$\varepsilon_{j,t}$ does not increase $y_{i,t}$ for s periods"

Short Run Restrictions

- note that Choleski decomposition is essentially a particular way to impose short run restrictions that imposes a recursive structure
- example: if $\boldsymbol{y}_t = (\Delta r_t, \Delta \log p_t, \Delta \log y_t)'$ the Choleski decomposition of $\boldsymbol{\Sigma}_e$ yields a lower diagonal matrix $\boldsymbol{B}_0$ so that in terms of short run restrictions we have $b_{12} = b_{13} = b_{23} = 0$, this also implies the following ordering for contemporaneous causality $r_t \rightarrow p_t \rightarrow y_t$

Short Run Restrictions

- consider a four variable macroeconomic model based on IS-LM framework

$$
\begin{aligned}
Y_t &= P_t^{\alpha_1} e^{\varepsilon_{as,t}} \\
Y_t &= e^{-\alpha_2 (r_t - \Delta \log P_t - \varepsilon_{is,t})} \\
\frac{M_t}{P_t} &= \frac{Y_t^{\alpha_3}}{e^{\alpha_4 r_t}} e^{-\varepsilon_{md,t}} \\
M_t &= M_{t-1} e^{\varepsilon_{ms,t}}
\end{aligned}
$$

  where Y_t is output, M_t money, r_t nominal interest rate, P_t price level

- first equation thus describes the short run aggregate supply, second one IS curve, third one LM curve, and the last one money supply
- let $\boldsymbol{y}_t = (\log Y_t, r_t, \log P_t, \log M_t)'$, take a log of each equation, to obtain

$$
\begin{aligned}
\log Y_t &= \alpha_1 \log P_t + \varepsilon_{as,t} \\
\log Y_t &= -\alpha_2 (r_t - \Delta \log P_t - \varepsilon_{is,t}) \\
\log M_t - \log P_t &= \alpha_3 \log Y_t - \alpha_4 r_t - \varepsilon_{md,t} \\
\log M_t &= \log M_{t-1} + \varepsilon_{ms,t}
\end{aligned}
$$

Short Run Restrictions

- model thus implies following matrix summarizing contemporaneous links

$$\boldsymbol{B}_0 = \begin{bmatrix} 1 & 0 & b_{0,13} & 0 \\ b_{0,21} & 1 & -1 & 0 \\ b_{0,31} & b_{0,32} & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

  where $b_{0,13} = -\alpha_1$, $b_{0,21} = 1/\alpha_2$, $b_{0,31} = \alpha_3$, $b_{0,32} = -\alpha_4$

- note that with four variables 6 restrictions are needed for exact identification, since $k(k-1)/2 = 6$ if $k = 4$
- model is actually *overidentified* - it introduces two additional restrictions, by restricting $b_{0,23} = -1$ and $b_{0,34} = -1$

Overidentifying Restrictions

- some theoretical models can suggest more than $k(k-1)/2$ restrictions
- testing whether these overidentifying restrictions are consistent with data:

1. first, estimate reduced form VAR model, obtain $\boldsymbol{\Sigma}_e$
2. since $\varepsilon_t = \boldsymbol{B}_0 \boldsymbol{e}_t$ and $\boldsymbol{\Sigma}_e = \boldsymbol{B}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{B}_0^{-1\prime}$, given the imposed restrictions use maximum likelihood approach to choose remaining parameters of $\boldsymbol{B}_0$ and $\boldsymbol{\Sigma}_\varepsilon$ in order to maximize the likelihood function

$$-\frac{T}{2} \log |\boldsymbol{B}_0^{-1} \boldsymbol{\Sigma}_\varepsilon \boldsymbol{B}_0^{-1\prime}| - \frac{1}{2} \sum_{t=1}^{T} (\boldsymbol{B}_0 \boldsymbol{e}_t)' \boldsymbol{\Sigma}_\varepsilon^{-1} (\boldsymbol{B}_0 \boldsymbol{e}_t)$$

   and let $\boldsymbol{\Sigma}_R = \boldsymbol{B}_0^{-1} \Sigma_\varepsilon \boldsymbol{B}_0^{-1\prime}$ be the resulting restricted variance matrix
3. denote by R number of overidentifying restrictions i.e. number of restrictions exceeding $k(k-1)/2$, then

$$|\boldsymbol{\Sigma}_R| - |\boldsymbol{\Sigma}_e|$$

   has χ^2 distribution with R degrees of freedom

Overidentifying Restrictions Application - Enders and Holt (2013)

- four variable VAR: $\boldsymbol{y}_t = (pe_t, ex_t, r_t, pg_t)'$ where pe_t is log of energy price index deflated by the producer price index, ex_t is real trade weighted exchange rate of U.S. dollar, r_t is 3-month T-bill rate adjusted for inflation, pg_t is log of price index of grain deflated by producer price index
- exact identification requires 6 restrictions: if $k = 4$ then $k(k-1)/2 = 6$
- test whether system with 9 restrictions is consistent with data

$$\boldsymbol{B}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ b_{0,41} & b_{0,42} & b_{0,43} & 1 \end{bmatrix}$$

  results in χ^2 statistic of 13.53 with 3 degrees of freedom and $p = 0.003$, so restrictions are strongly rejected by data
- strongly correlated residuals for exchange and interest rate equation suggest using a modified system where real exchange rate is contemporaneously affected by real interest rate shocks

$$\boldsymbol{B}_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & b_{0,23} & 0 \\ 0 & 0 & 1 & 0 \\ b_{0,41} & b_{0,42} & b_{0,43} & 1 \end{bmatrix}$$

  which yields χ^2 statistic of 4.57 with 2 degrees of freedom and $p = 0.102$,

Long Run Restrictions - Blanchard & Quah (1989)

- in Blanchard and Quah, ε_i shocks are not considered as shocks directly associated with y_i, they instead assert that some shocks have permanent effects and others only temporary effects on some variables
- note: to use Blanchard and Quah technique, at least one variable must be nonstationary, $I(0)$ variables do not have a permanent component
- consider a reduced form VAR $\boldsymbol{A}(L)\boldsymbol{y}_t = \boldsymbol{e}_t$ with vector moving average representation

$$\boldsymbol{y}_t = \boldsymbol{\Psi}(L)\boldsymbol{\varepsilon}_t = \sum_{\ell=0}^{\infty} \boldsymbol{\Psi}_\ell \boldsymbol{\varepsilon}_{t-\ell}$$

- elements of the impulse-response function can be obtained from row i and column j element of $\boldsymbol{\Psi}_\ell$

$$\psi_{\ell,ij} = \frac{\partial y_{t+\ell,i}}{\partial \varepsilon_{t,j}}$$

- cumulative impact up to period ℓ

$$\psi^*_{\ell,ij} = \sum_{s=0}^{\ell} \psi_{s,ij} = \frac{\partial}{\partial \varepsilon_{t,j}} \sum_{s=0}^{\ell} y_{t+s,i}$$

Long Run Restrictions - Blanchard & Quah (1989)

- example: let $\boldsymbol{y}_t = (\Delta \log GDP_t, UR_t)'$, with both variables demeaned, and let $\varepsilon_{1,t}$ represent the technology shocks and $\varepsilon_{2,t}$ the non-technology shocks
- the moving average representation of the bivariate VAR is

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \Psi_{11}(L) & \Psi_{12}(L) \\ \Psi_{21}(L) & \Psi_{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

  and the equation for $\Delta \log GDP_t$ is thus

$$y_{1,t} = \Psi_{11}(L)\varepsilon_{1,t} + \Psi_{12}(L)\varepsilon_{2,t} = \sum_{\ell=0}^{\infty} \psi_{\ell,11} L^{\ell} \varepsilon_{1,t} + \sum_{\ell=0}^{\infty} \psi_{\ell,12} L^{\ell} \varepsilon_{2,t}$$

- long run constraint imposed: non-technology shocks only have a temporary effect on the *level* of GDP and thus

$$\lim_{\ell \to \infty} \frac{\partial \log GDP_{t+\ell}}{\partial \varepsilon_{2,t}} = 0$$

- equivalently, long run cumulative effect of non-technology shocks on the growth rate of GDP so on $\Delta \log GDP_t$ is zero, $\Psi_{12}(1) = 0$

$$\lim_{\ell \to \infty} \frac{\partial \sum_{s=0}^{\ell} \Delta \log GDP_{t+s}}{\partial \varepsilon_{2,t}} = \lim_{\ell \to \infty} \sum_{s=0}^{\ell} \psi_{s,12} = 0$$

Long Run Restrictions - Blanchard & Quah (1989)

- to get more insight how the Blanchard-Quah approach works consider a structural VAR(1)

$$\boldsymbol{B}_0 \boldsymbol{y}_t = \boldsymbol{B}_1 \boldsymbol{y}_{t-1} + \varepsilon_t \qquad \text{where } var(\varepsilon_t) = \Sigma_\varepsilon$$

- we can normalize $\Sigma_\varepsilon = \boldsymbol{I}$ i.e. obtain $\sigma^2_{\varepsilon_i} = 1$ if we divide each equation by σ_{ε_i}; this changes the diagonal elements of $\boldsymbol{B}_0$ from 1 into $1/\sigma_{\varepsilon_i}$
- the associated reduced form VAR(1) is

$$\boldsymbol{y}_t = \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \boldsymbol{e}_t$$

where $\boldsymbol{A}_1 = \boldsymbol{B}_0^{-1} \boldsymbol{B}_1$ and $\boldsymbol{e}_t = \boldsymbol{B}_0^{-1} \varepsilon_t$ and $var(\boldsymbol{e}_t) = \Sigma_e = \boldsymbol{B}_0^{-1} \Sigma_\varepsilon \boldsymbol{B}_0^{-1\prime}$

- estimating reduced form VAR yields $\boldsymbol{A}_1$ and $\boldsymbol{\Sigma}_e$, the identification problem is then to use these to recover parameters of structural VAR, $\boldsymbol{B}_0$, $\boldsymbol{B}_1$ and $\boldsymbol{\Sigma}_\varepsilon$
- since $\boldsymbol{A}_1 = \boldsymbol{B}_0^{-1} \boldsymbol{B}_1$ once we know $\boldsymbol{B}_0$ we can use $\boldsymbol{A}_1$ to obtain $\boldsymbol{B}_1$ using $\boldsymbol{B}_1 = \boldsymbol{B}_0 \boldsymbol{A}_1$
- the main task is thus to determine the four elements of $\boldsymbol{B}_0$ since we have normalized $\boldsymbol{\Sigma}_\varepsilon = \boldsymbol{I}$

Long Run Restrictions - Blanchard & Quah (1989)

consider a bivariate VAR(1) and denote the elements of $\boldsymbol{B}_0$ and $\boldsymbol{B}_0^{-1}$ as follows

$$\boldsymbol{B}_0 = \begin{bmatrix} b_{0,11} & b_{0,12} \\ b_{0,21} & b_{0,22} \end{bmatrix} \quad \boldsymbol{B}_0^{-1} = \frac{1}{b_{0,11}b_{0,22} - b_{0,21}b_{0,12}} \begin{bmatrix} b_{0,22} & -b_{0,12} \\ -b_{0,21} & b_{0,11} \end{bmatrix} = \begin{bmatrix} \tilde{b}_{0,11} & \tilde{b}_{0,12} \\ \tilde{b}_{0,21} & \tilde{b}_{0,22} \end{bmatrix}$$

so that equation $\Sigma_e = \boldsymbol{B}_0^{-1}\Sigma_\varepsilon \boldsymbol{B}_0^{-1\prime}$ yields with normalization $\sigma^2_{\varepsilon_1} = \sigma^2_{\varepsilon_2} = 1$

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \tilde{b}_{0,11} & \tilde{b}_{0,12} \\ \tilde{b}_{0,21} & \tilde{b}_{0,22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{b}_{0,11} & \tilde{b}_{0,21} \\ \tilde{b}_{0,12} & \tilde{b}_{0,22} \end{bmatrix}$$

or equivalently

$$\begin{aligned} \sigma_1^2 &= \tilde{b}_{0,11}^2 + \tilde{b}_{0,12}^2 \\ \sigma_{12} &= \tilde{b}_{0,11}\tilde{b}_{0,21} + \tilde{b}_{0,12}\tilde{b}_{0,22} \\ \sigma_{12} &= \tilde{b}_{0,11}\tilde{b}_{0,21} + \tilde{b}_{0,12}\tilde{b}_{0,22} \\ \sigma_2^2 &= \tilde{b}_{0,21}^2 + \tilde{b}_{0,22}^2 \end{aligned}$$

so we only have 3 independent equations but 4 unknowns $\{\tilde{b}_{0,11}, \tilde{b}_{0,12}, \tilde{b}_{0,21}, \tilde{b}_{0,22}\}$

in order to get 4 equations in 4 unknowns $\{\tilde{b}_{0,11}, \tilde{b}_{0,12}, \tilde{b}_{0,21}, \tilde{b}_{0,22}\}$ we next impose one long run restriction

Long Run Restrictions - Blanchard & Quah (1989)

- from the reduced form VAR by repeated substitution we have

$$\boldsymbol{y}_t = \boldsymbol{e}_t + \boldsymbol{A}_1 \boldsymbol{e}_{t-1} + \boldsymbol{A}_1^2 \boldsymbol{e}_{t-2} + \ldots = \sum_{\ell=0}^{\infty} \boldsymbol{A}_1^{\ell} \boldsymbol{e}_{t-\ell} = \sum_{\ell=0}^{\infty} \boldsymbol{A}_1^{\ell} L^{\ell} \boldsymbol{e}_t$$

  and using $\boldsymbol{e}_t = \boldsymbol{B}_0^{-1} \varepsilon_t$ we get the vector moving average representation

$$\boldsymbol{y}_t = \boldsymbol{B}_0^{-1} \varepsilon_t + \boldsymbol{A}_1 L \boldsymbol{B}_0^{-1} \varepsilon_t + \boldsymbol{A}_1^2 L^2 \boldsymbol{B}_0^{-1} \varepsilon_t + \ldots = \left(\sum_{\ell=0}^{\infty} \boldsymbol{A}_1^{\ell} L^{\ell} \right) \boldsymbol{B}_0^{-1} \varepsilon_t$$

- let

$$\sum_{\ell=0}^{\infty} \boldsymbol{A}_1^{\ell} L^{\ell} = (\boldsymbol{I} + \boldsymbol{A}_1 L + \boldsymbol{A}_1^2 L^2 + \ldots) = (\boldsymbol{I} - \boldsymbol{A}_1 L)^{-1}$$

  then

$$\boldsymbol{y}_t = (\boldsymbol{I} - \boldsymbol{A}_1 L)^{-1} \boldsymbol{B}_0^{-1} \varepsilon_t$$

- note that this condition can be also obtained from the reduced form VAR(1) by rearranging it first as $(\boldsymbol{I} - \boldsymbol{A}_1 L) \boldsymbol{y}_t = \boldsymbol{e}_t$ from which $\boldsymbol{y}_t = (\boldsymbol{I} - \boldsymbol{A}_1 L)^{-1} \boldsymbol{B}_0^{-1} \varepsilon_t$

Long Run Restrictions - Blanchard & Quah (1989)

- thus let $\boldsymbol{S}$ be the matrix of long run cumulative effects of ε on $\boldsymbol{y}$ given by

$$\boldsymbol{S} = (\boldsymbol{I} + \boldsymbol{A}_1 + \boldsymbol{A}_1^2 + \ldots)\boldsymbol{B}_0^{-1} = (\boldsymbol{I} - \boldsymbol{A}_1)^{-1}\boldsymbol{B}_0^{-1}$$

- in a bivariate framework, Blanchard and Quah impose a constraint that second shock has no cumulative long run effect on first variable, which means that $S_{12} = 0$
- if we denote the elements of matrix $(\boldsymbol{I} - \boldsymbol{A}_1)^{-1}$ as

$$(\boldsymbol{I} - \boldsymbol{A}_1)^{-1} = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & \tilde{a}_{22} \end{bmatrix}$$

  the system of 4 equations in 4 unknowns $\{\tilde{b}_{0,11}, \tilde{b}_{0,12}, \tilde{b}_{0,21}, \tilde{b}_{0,22}\}$ becomes

$$\begin{aligned} \sigma_1^2 &= \tilde{b}_{0,11}^2 + \tilde{b}_{0,12}^2 \\ \sigma_{12} &= \tilde{b}_{0,11}\tilde{b}_{0,21} + \tilde{b}_{0,12}\tilde{b}_{0,22} \\ \sigma_2^2 &= \tilde{b}_{0,21}^2 + \tilde{b}_{0,22}^2 \\ 0 &= \tilde{a}_{11}\tilde{b}_{0,12} + \tilde{a}_{21}\tilde{b}_{0,22} \end{aligned}$$

  where the last condition corresponds to $S_{12} = 0$
- this system can be easily solved given parameters $\{\sigma_1^2, \sigma_2^2, \sigma_{12}, \tilde{a}_{11}, \tilde{a}_{21}\}$ (which come from $\boldsymbol{\Sigma}_e$ and $\boldsymbol{A}_1$)

Long Run Restrictions - Blanchard & Quah (1989)

```r
# obtain data on real GDP and unemployment rate
# construct approximate quarter-over-quarter GDP growth rates
y_tbl <- inner_join(tq_get("GDPC1", get  = "economic.data",
                           from = "1947-01-01", to = "2018-12-31") %>%
                        rename(rGDP = price) %>%
                        mutate(dlrGDP = 100*(log(rGDP) - lag(log(rGDP)))),
                    tq_get("UNRATE", get  = "economic.data",
                           from = "1947-01-01", to = "2018-12-31") %>%
                        rename(UR = price) %>%
                        tq_transmute(select = UR, mutate_fun = to.quarterly) %>%
                        mutate(date = as.Date(date)),
                    by = "date") %>%
    mutate(yearq = as.yearqtr(date)) %>%
    dplyr::select(yearq, dlrGDP, UR)

# Blanchard and Quah use 1950Q2 to 1987Q4 as sample, and demean the data
y_xts <- y_tbl %>%
    filter(yearq >= "1950 Q2", yearq <= "1987 Q4") %>%
    mutate_at(vars(dlrGDP,UR), funs(. - mean(.))) %>%
    tk_xts(select= c("dlrGDP","UR"), date_var = yearq)
```

Long Run Restrictions - Blanchard & Quah (1989)

```
# estimate reduced form VAR
mod_var <- VAR(y_xts, ic = "SC", lag.max = 8)

# Blanchard-Quah long run restriction: row 1 column 2 element of the cumulative effect matrix is 0
mod_svar <- BQ(mod_var)
summary(mod_svar)
```

```
##
## SVAR Estimation Results:
## ========================
##
## Call:
## BQ(x = mod_var)
##
## Type: Blanchard-Quah
## Sample size: 149
## Log Likelihood: -232.816
##
## Estimated contemporaneous impact matrix:
##          dlrGDP      UR
## dlrGDP  0.79328 -0.5393
## UR     -0.03914  0.3787
##
## Estimated identified long run impact matrix:
##         dlrGDP    UR
## dlrGDP  0.5734 0.000
## UR     -2.6987 6.159
##
## Covariance matrix of reduced form residuals (*100):
##        dlrGDP     UR
## dlrGDP  92.01 -23.53
## UR     -23.53  14.50
```

Long Run Restrictions - Blanchard & Quah (1989)

- in the output on the previous slide the contemporaneous impact matrix reported is $\boldsymbol{B}_0^{-1}$, it shows the immediate effect of $\varepsilon_{j,t}$ on $y_{i,t}$ upon impact

```
##
## Estimated contemporaneous impact matrix:
##          dlrGDP      UR
## dlrGDP  0.79328 -0.5393
## UR     -0.03914  0.3787
```

- rows refer to two variables ($\Delta \log GDP_t$, UR_t), and the columns to the two shocks - technology shock $\varepsilon_{1,t}$ and non-technology shock $\varepsilon_{2,t}$
- here on impact a positive one standard deviation technology shock increases GDP by 0.793% and lowers unemployment rate by 0.0391 percentage points
- a negative one standard deviation non-technology shock lowers GDP on impact by 0.539%, increases unemployment rate by 0.378 percentage points

Long Run Restrictions - Blanchard & Quah (1989)

- the long run impact matrix reported shows the cumulative long run impact $\lim_{\ell\to\infty}\sum_{s=0}^{\ell}\psi_{s,ij}=0$

```
##
## Estimated identified long run impact matrix:
##         dlrGDP    UR
## dlrGDP  0.5734 0.000
## UR     -2.6987 6.159
```

- the long run cumulative effect of any non-technology shock on GDP is 0 (this is the long run constraint we imposed)
- the long run cumulative effect of a single positive one standard deviation technology shocks on GDP is to increase it by 0.576%

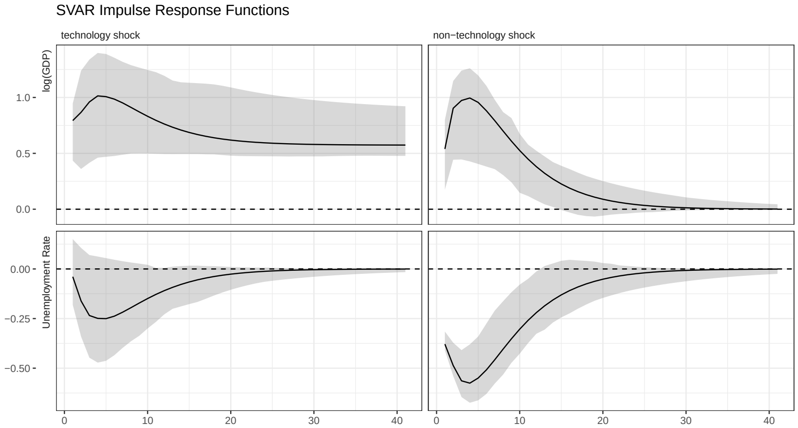
Long Run Restrictions - Blanchard & Quah (1989)

```r
# standard non-cumulative IRFs
svar_irf <- irf(mod_svar, n.ahead = 40, ci = .9)
# cumulative svar_irfs
svar_irf_c <- irf(mod_svar, n.ahead = 40, ci = .9, cumulative = TRUE)

# arrange IRF data into a tibble to be used with ggplot
svar_irf_tbl <-
    bind_rows(# standard IRFs for UR
              svar_irf %>%
                  keep(names(.) %in% c("irf", "Lower", "Upper")) %>%
                  modify_depth(2, as_tibble) %>%
                  modify_depth(1, bind_rows, .id = "impulse") %>%
                  map_df(bind_rows, .id = "key") %>%
                  dplyr::select(-dlrGDP) %>%
                  gather(response, value, -key, -impulse),
              # cumulative IRFs for GDP
              svar_irf_c %>%
                  keep(names(.) %in% c("irf", "Lower", "Upper")) %>%
                  modify_depth(2, as_tibble) %>%
                  modify_depth(1, bind_rows, .id = "impulse") %>%
                  map_df(bind_rows, .id = "key") %>%
                  dplyr::select(-UR) %>%
                  gather(response, value, -key, -impulse)) %>%
    group_by(key, impulse, response) %>%
    mutate(lag = row_number()) %>%
    ungroup() %>%
    # change signs for the non-technology shock IRFs so that they show effects of a positive shock
    mutate(value = if_else(impulse == "UR", -value, value)) %>%
    spread(key, value)
```

Long Run Restrictions - Blanchard & Quah (1989)

```r
# plot IRFs using ggplot
svar_irf_tbl %>%
    mutate(impulse_label = case_when(impulse == "dlrGDP" ~ 1,
                                     impulse == "UR"     ~ 2) %>%
                           factor(labels = c("technology shock", "non-technology shock")),
           response_label = case_when(response == "dlrGDP" ~ "log(GDP)",
                                      response == "UR" ~ "Unemployment Rate") ) %>%
    ggplot(aes(x = lag, y = irf)) +
        geom_ribbon(aes(x = lag, ymin = Lower, ymax = Upper), fill = "gray50", alpha = .3) +
        geom_line() +
        geom_hline(yintercept = 0, linetype = "dashed") +
        labs(x = "", y = "", title = "SVAR Impulse Response Functions") +
        facet_grid(response_label ~ impulse_label, switch = "y", scales = "free_y")
```

Long Run Restrictions - Blanchard & Quah (1989)



Long Run Restrictions - Blanchard & Quah (1989)

- the peak effect for both shocks occurs 3 quarters after the shock hits the economy
- in case of a positive one standard deviation shock to technology, at the peak GDP increases by about 1% and unemployment rate falls by roughly 0.25 percentage points
- in case of a positive one standard deviation non-technology shock, at the peak GDP increases by about 1% and unemployment rate falls by roughly 0.57 percentage points

Long Run Restrictions - Blanchard & Quah (1989)

```
# construct longer cumulative IRFs, and keep non-technology shocks as negative one
svar_irf_c_longer <- irf(mod_svar, n.ahead = 100, cumulative = TRUE, boot = FALSE)
```

note that by construction the contemporaneous impact matrix from `summary(mod_svar)`

```
##
## Estimated contemporaneous impact matrix:
##          dlrGDP      UR
## dlrGDP  0.79328 -0.5393
## UR     -0.03914  0.3787
```

is identical to the elements of the IRFs for period 0 (impact period)

```
svar_irf_c_longer$irf[[1]][1,]
```

```
##      dlrGDP          UR
##  0.79328160 -0.03914466
```

```
svar_irf_c_longer$irf[[2]][1,]
```

```
##     dlrGDP         UR
## -0.5392647  0.3787316
```

Long Run Restrictions - Blanchard & Quah (1989)

also note that the long run impact matrix from `summary(mod_svar)`

```
##
## Estimated identified long run impact matrix:
##         dlrGDP    UR
## dlrGDP  0.5734 0.000
## UR     -2.6987 6.159
```

is essentially the same as the elements of the IRFs for period 100

```
svar_irf_c_longer$irf[[1]][101,]
```

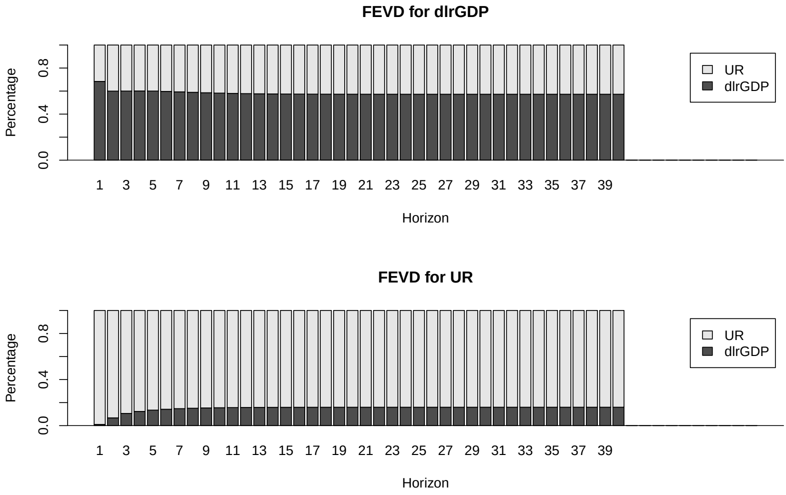
```
##     dlrGDP         UR
##  0.5733992 -2.6986514
```

```
svar_irf_c_longer$irf[[2]][101,]
```

```
##        dlrGDP            UR
## -8.760994e-09  6.159035e+00
```

Long Run Restrictions - Blanchard & Quah (1989)

```
mod_svar %>% fevd(n.ahead=40) %>% plot(addbars = 10)
```



Long Run Restrictions - Other Examples

other examples of long-run neutrality where changes in nominal variables have no effect on real economic variables in the long-run:

- permanent change in nominal money stock has no long-run effect on the level of real output
- permanent change in the rate of inflation has no long-run effect on unemployment (vertical Phillips curve)
- permanent change in the rate of inflation has no long-run effect on real interest rates (long-run Fisher relationship).

Sign Restrictions

- **price puzzle**: in a VAR with $\boldsymbol{y}_t = (\log GFP_t, \log p_t^{GDP}, r_t)$ after monetary tightening prices *go up* which is completely counter intuitive according to the standard transmission mechanism
- Sims (1992): (i) interest rate not the only instrument and (ii) prices appear to rise because the VAR model does not include information about future inflation that is available to Fed
- Uhlig (2005): study monetary policy shocks using restrictions which are implied by several theoretical economic models - a contractionary monetary policy shock does not
  - reduce short term interest rate for x periods after the shock
  - increase prices for x periods after the shock
  - increase monetary aggregates (reserves) for x periods after the shock

Sign Restrictions

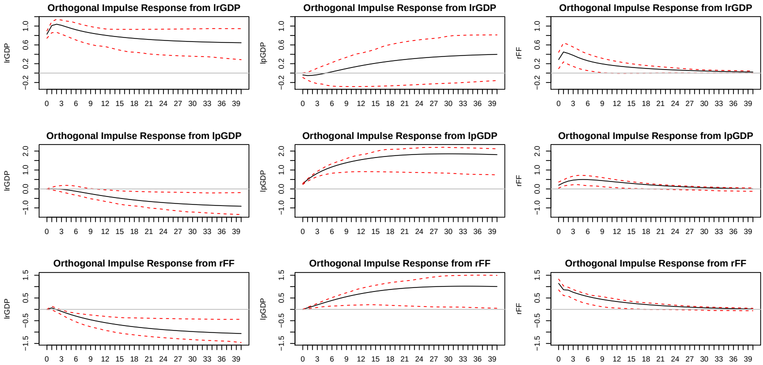
- consider a simple VAR with GDP growth rate, inflation rate, and nominal interest rate so that $\boldsymbol{y}_t = (\Delta logGDP_t, \Delta \log p_t^{GDP}, FF_t)'$

```
library(Quandl)
Quandl.api_key('DLk9RQrfTVkD4UTKc7op')

rGDP <- Quandl("FRED/GDPC1", type="zoo")
pGDP <- Quandl("FRED/GDPDEF", type="zoo")
rFF <- Quandl("FRED/FEDFUNDS", collapse="quarterly", type="zoo")

lrGDP <- log(rGDP)*100
lpGDP <- log(pGDP)*100

y <- cbind(lrGDP, lpGDP, rFF)
vlabels <- c("log(GDP)","GDP deflator","FF rate")

y <- na.trim(y)
y <- window(y, end="2007 Q4")
```

Sign Restrictions

- IRFs based on Choleski decomposition - increase in nominal interest rate is associated with price increase in future

```
library(vars)
mod_var <- VAR(y, ic = "SC", lag.max = 16, type = "none")
svar_irf <- irf(mod_var, n.ahead = 40, ci = .9)

par(mfrow = c(3,3), cex = 0.6, mar = c(4,4,2,1))
plot(svar_irf, plot.type="single", ask = FALSE)
```



Sign Restrictions

- IRF based on sign restriction that a contractionary monetary policy increases nominal interest rate and decreases prices for at least 4 quarters
- sign restrictions are only weak restrictions on $\boldsymbol{B}_0$, in the SVAR model there is a lot of uncertainty regarding the response of GDP to an increase in nominal interest rate

```r
library(VARsignR)
constr <- c(+3,-2)
mod_svar <- uhlig.reject(as.ts(y), nlags=2, constant=FALSE, steps=40, constrained=constr)

irfplot(mod_svar$IRFS, type="median", labels=vlabels)
```