Notes on Structural VAR Modeling

Eric Zivot

May 1, 2000
This version: June 5, 2000

1 The Structural VAR Model for Stationary Data

1.1 VAR Representations

Consider the simple covariance stationary bivariate dynamic simultaneous equations model

$$\begin{aligned} y_{1t} &= \gamma_{10} - b_{12}y_{2t} + \gamma_{11}y_{1t-1} + \gamma_{12}y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= \gamma_{20} - b_{21}y_{1t} + \gamma_{21}y_{1t-1} + \gamma_{22}y_{2t-1} + \varepsilon_{2t} \end{aligned} \tag{1}$$

where

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim i.i.d. \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \right). \tag{2}$$

The sample consists of observations from $t = 1, \ldots, T$ with a &xed initial value $\mathbf{y}_0 = (y_{10}, y_{20})'$. The model (1) is called a *structural VAR* (SVAR) since it is assumed to be derived by some underlying economic theory. The exogenous error terms ε_{1t} and ε_{2t} are independent and are interpreted as *structural innovations*. For example, let y_{1t} denote the log of detrended real GDP and y_{2t} denote the log of detrended nominal money supply. Then realizations of ε_{1t} are interpreted as capturing unexpected shocks to output that are uncorrelated with ε_{2t}, the unexpected shocks to the money supply. In (1), the endogeneity of y_{1t} and y_{2t} is determined by the values of b_{12} and b_{21}.

In matrix form, the model (1) becomes

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

or

$$\mathbf{B}\mathbf{y}_t = \boldsymbol{\gamma}_0 + \boldsymbol{\Gamma}_1\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{3}$$

where

$$E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] = \mathbf{D}$$

and $\mathbf{D}$ is a diagonal matrix with elements σ_1^2 and σ_2^2. In lag operator notation, the SVAR (1) becomes

$$\begin{aligned}\mathbf{B}(L)\mathbf{y}_t &= \boldsymbol{\gamma}_0+\boldsymbol{\varepsilon}_t,\\ \mathbf{B}(L) &= \mathbf{B}-\boldsymbol{\Gamma}_1 L.\end{aligned}$$

The *reduced form* of the SVAR, a standard VAR model, is found by multiplying (3) by $\mathbf{B}^{-1}$, assuming it exists, and solving for $\mathbf{y}_t$ in terms of $\mathbf{y}_{t-1}$ and ε_t :

$$\begin{aligned}\mathbf{y}_t &= \mathbf{B}^{-1}\boldsymbol{\gamma}_0+\mathbf{B}^{-1}\boldsymbol{\Gamma}_1\mathbf{y}_{t-1}+\mathbf{B}^{-1}\boldsymbol{\varepsilon}_t \qquad (4)\\ &= \mathbf{a}_0+\mathbf{A}_1\mathbf{y}_{t-1}+\mathbf{u}_t.\end{aligned}$$

or

$$\begin{aligned}\mathbf{A}(L)\mathbf{y}_t &= \mathbf{a}_0+\mathbf{u}_t,\\ \mathbf{A}(L) &= \mathbf{I}_2-\mathbf{A}_1 L.\end{aligned}$$

Given that

$$\mathbf{B}^{-1} = \frac{1}{\Delta}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix},\ \Delta = \det(\mathbf{B}) = 1-b_{12}b_{21}$$

we have

$$\begin{aligned}\mathbf{a}_0 &= \mathbf{B}^{-1}\boldsymbol{\gamma}_0 = \frac{1}{\Delta}\begin{bmatrix}\gamma_{10}-b_{12}\gamma_{20}\\ \gamma_{20}-b_{21}\gamma_{10}\end{bmatrix} = \begin{bmatrix}a_{10}\\ a_{20}\end{bmatrix},\\ \mathbf{A}_1 &= \mathbf{B}^{-1}\boldsymbol{\Gamma}_1 = \frac{1}{\Delta}\begin{bmatrix}\gamma_{11}-b_{12}\gamma_{21} & \gamma_{12}-b_{12}\gamma_{22}\\ \gamma_{21}-b_{21}\gamma_{11} & \gamma_{22}-b_{21}\gamma_{12}\end{bmatrix} = \begin{bmatrix}a_{11} & a_{12}\\ a_{22} & a_{21}\end{bmatrix},\\ \mathbf{u}_t &= \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t = \frac{1}{\Delta}\begin{bmatrix}\varepsilon_{1t}-b_{12}\varepsilon_{2t}\\ \varepsilon_{2t}-b_{21}\varepsilon_{1t}\end{bmatrix} = \begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix}.\end{aligned}$$

The reduced form errors $\mathbf{u}_t$ are linear combinations of the structural errors $\boldsymbol{\varepsilon}_t$ and have covariance matrix

$$\begin{aligned}E[\mathbf{u}_t\mathbf{u}_t'] &= \mathbf{B}^{-1}E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t']\mathbf{B}^{-1\prime}\\ &= \mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1\prime}\\ &= \boldsymbol{\Omega}.\end{aligned}$$

Speci&cally, the elements in $\boldsymbol{\Omega}$ are

$$\begin{bmatrix}\omega_1^2 & \omega_{12}\\ \omega_{12} & \omega_2^2\end{bmatrix} = \frac{1}{\Delta^2}\begin{bmatrix}\sigma_1^2+b_{12}^2\sigma_2^2 & -(b_{21}\sigma_1^2+b_{12}\sigma_2^2)\\ -(b_{21}\sigma_1^2+b_{12}\sigma_2^2) & \sigma_2^2+b_{21}^2\sigma_1^2\end{bmatrix}.$$

Note that Ω is diagonal only if $b_{12}=b_{21}=0$.

1.1.1 Stationarity Conditions

The reduced form VAR (4) is covariance stationary provided the eigenvalues of $\mathbf{A}_1$ have modulus less than 1. The eigenvalues of $\mathbf{A}_1$ satisfy the equation

$$\det\left(\mathbf{I}_2\lambda - \mathbf{A}_1\right) = 0$$

and are equal to the inverses of the roots to the characteristic equation

$$\det(\mathbf{I}_2 - \mathbf{A}_1 z) = 0. \tag{5}$$

Hence, the reduced form VAR is stationary provided the roots of (5) lie outside the complex unit circle. Evaluating the determinant in (5) gives

$$(1 - a_{11}z)(1 - a_{22}z) - a_{12}a_{21}z^2 = 0$$

and the roots can be determined using the quadratic formula.

1.1.2 Identi&cation Issues

Without some restrictions, the parameters in the SVAR are not identi&ed. That is, given values of the reduced form parameters $\mathbf{a}_0, \mathbf{A}_1$ and $\mathbf{\Omega}$, it is not possible to uniquely solve for the structural parameters $\mathbf{B}, \boldsymbol{\gamma}_0, \mathbf{\Gamma}_1$ and $\mathbf{D}$. There are ten structural parameters (eight coefficients and two covariance elements) and only nine reduced form parameters (six coefficients and three covariance elements). Clearly, at least 1 restriction on the parameters of SVAR is required in order to identify all of the structural parameters. Sims (1980) argued that economic theory is not rich enough to suggest proper identi&cation restrictions on the SVAR. Therefore, the best we can do is to estimate the reduced form VAR (4). This bleak view is not generally accepted but there is considerable debate about what constitutes appropriate identifying restrictions. Typical identifying restrictions include

- Zero (exclusion) restrictions on the elements of $\mathbf{B}$; e.g., $b_{12} = 0$.
- Linear restrictions on the elements of $\mathbf{B}$; e.g., $b_{12} + b_{21} = 1$.

1.2 MA Representations

The *moving average* (MA) or Wold representation of the reduced form VAR (4) is found by multiplying both sides of (4) by $\mathbf{A}(L)^{-1} = (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1}$ to give

$$\mathbf{y}_t = \boldsymbol{\mu} + \mathbf{\Psi}(L)\mathbf{u}_t, \tag{6}$$

where

$$\begin{aligned}
\mathbf{\Psi}(L) &= (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1} \\
&= \sum_{k=0}^{\infty} \mathbf{\Psi}_k L^k,\ \ \mathbf{\Psi}_0 = \mathbf{I}_2, \mathbf{\Psi}_k = \mathbf{A}_1^k, \\
\boldsymbol{\mu} &= (\mathbf{I}_2 - \mathbf{A}_1)^{-1}\mathbf{a}_0
\end{aligned}$$

In the Wold representation for $\mathbf{y}_t$, the first matrix in the moving average polynomial $\mathbf{\Psi}(L)$ is $\mathbf{\Psi}_0 = \mathbf{I}_2$. In addition, the error terms $\mathbf{u}_t$ are generally contemporaneously correlated and have covariance matrix $\mathbf{\Omega}$.

The *structural moving average* (SMA) representation of $\mathbf{y}_t$ is based on an infinite moving average of the structural innovations $\boldsymbol{\varepsilon}_t$. Substituting $\mathbf{u}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t$ into (6) gives

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu} + \mathbf{\Psi}(L)\mathbf{B}^{-1}\boldsymbol{\varepsilon}_t \\ &= \boldsymbol{\mu} + \mathbf{\Theta}(L)\boldsymbol{\varepsilon}_t \end{aligned} \quad (7)$$

where

$$\begin{aligned} \mathbf{\Theta}(L) &= \sum_{k=0}^{\infty} \mathbf{\Theta}_k L^k \\ &= \mathbf{\Psi}(L)\mathbf{B}^{-1} \\ &= \mathbf{B}^{-1} + \mathbf{\Psi}_1\mathbf{B}^{-1}L + \cdots \end{aligned}$$

That is, $\mathbf{\Theta}_k = \mathbf{\Psi}_k\mathbf{B}^{-1}$ for $k = 0, 1, \ldots$. In particular, notice that $\mathbf{\Theta}_0 = \mathbf{B}^{-1} \neq \mathbf{I}_2$. Note that $\mathbf{\Theta}(L) = \mathbf{B}(L)^{-1} = \mathbf{\Psi}(L)\mathbf{B}^{-1} = (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1}\mathbf{B}^{-1}$.

It is instructive to look at the SMA representation for the bivariate system

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} \\ \theta_{21}^{(1)} & \theta_{22}^{(1)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{bmatrix} + \cdots$$

which illustrates that the elements of the $\mathbf{\Theta}_k$ matrices, $\theta_{ij}^{(k)}$, give the dynamic multipliers or impulse responses of y_{1t} and y_{2t} to changes in ε_{1t} and ε_{2t}.

1.3 Impulse Response Functions

Consider the SMA representation (7) at time $t + s$

$$\begin{bmatrix} y_{1t+s} \\ y_{2t+s} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+s} \\ \varepsilon_{2t+s} \end{bmatrix} + \cdots + \begin{bmatrix} \theta_{11}^{(s)} & \theta_{12}^{(s)} \\ \theta_{21}^{(s)} & \theta_{22}^{(s)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \cdots.$$

The *structural dynamic multipliers* are

$$\begin{aligned} \frac{\partial y_{1t+s}}{\partial \varepsilon_{1t}} &= \theta_{11}^{(s)}, \frac{\partial y_{1t+s}}{\partial \varepsilon_{2t}} = \theta_{12}^{(s)} \\ \frac{\partial y_{2t+s}}{\partial \varepsilon_{1t}} &= \theta_{21}^{(s)}, \frac{\partial y_{2t+s}}{\partial \varepsilon_{2t}} = \theta_{22}^{(s)} \end{aligned} \quad (8)$$

The *structural impulse response functions* (IRFs) are the plots of $\theta_{ij}^{(s)}$ vs. s for $i, j = 1, 2$. These plots summarize how unit impulses of the structural shocks at time t impact the level of y at time $t + s$ for different values of s.

Since y_t is assumed to be covariance stationary we know that

$$\lim_{s\to\infty} \theta_{ij}^{(s)} = 0, \ \ i,j = 1,2 \tag{9}$$

so that no structural shock has a long-run impact on the level of y. The long-run cumulative impact of the structural shocks is captured by the long-run impact matrix

$$\mathbf{\Theta}(1) = \left[\begin{array}{cc} \theta_{11}(1) & \theta_{12}(1) \\ \theta_{21}(1) & \theta_{22}(1) \end{array} \right] = \left[\begin{array}{cc} \sum_{s=0}^{\infty} \theta_{11}^{(s)} & \sum_{s=0}^{\infty} \theta_{12}^{(s)} \\ \sum_{s=0}^{\infty} \theta_{21}^{(s)} & \sum_{s=0}^{\infty} \theta_{22}^{(s)} \end{array} \right]$$

where

$$\mathbf{\Theta}(L) = \left[\begin{array}{cc} \theta_{11}(L) & \theta_{12}(L) \\ \theta_{21}(L) & \theta_{22}(L) \end{array} \right] = \left[\begin{array}{cc} \sum_{s=0}^{\infty} \theta_{11}^{(s)} L^s & \sum_{s=0}^{\infty} \theta_{12}^{(s)} L^s \\ \sum_{s=0}^{\infty} \theta_{21}^{(s)} L^s & \sum_{s=0}^{\infty} \theta_{22}^{(s)} L^s \end{array} \right].$$

1.3.1 Identi&cation issues

In some applications, identi&cation of the parameters of the SVAR is achieved through restrictions on the parameters of the SMA representation (7). For example, suppose that ε_{1t} has no contemporaneous impact on y_{2t}. Then $\theta_{12}^{(0)} = 0$ and so Θ_0 becomes triangular

$$\mathbf{\Theta}_0 = \left[\begin{array}{cc} \theta_{11}^{(0)} & 0 \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{array} \right].$$

Since $\mathbf{\Theta}_0 = \mathbf{B}^{-1}$ we then have

$$\left[\begin{array}{cc} \theta_{11}^{(0)} & 0 \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{array} \right] = \frac{1}{\Delta} \left[\begin{array}{cc} 1 & -b_{12} \\ -b_{21} & 1 \end{array} \right]$$

which implies that $b_{12} = 0$. Hence, assuming $\theta_{12}^{(0)} = 0$ in the SMA representation (7) is equivalent to assuming $b_{12} = 0$ in the SVAR representation (1).

As another example, suppose ε_{1t} has no long-run cumulative impact on y_{2t}. Then $\theta_{12}(1) = \sum_{s=0}^{\infty} \theta_{12}^{(s)} = 0$ and the long-run impact matrix $\mathbf{\Theta}(1)$ becomes triangular:

$$\mathbf{\Theta}(1) = \left[\begin{array}{cc} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{array} \right].$$

This type of long-run restriction places restrictions on the coefficients of the SVAR (1) since

$$\mathbf{\Theta}(1) = \mathbf{\Psi}(1)\mathbf{B}^{-1} = (\mathbf{I}_2 - \mathbf{A}_1)^{-1}\mathbf{B}^{-1}$$

1.3.2 Estimation Issues

In order to compute the structural IRFs, the parameters of the SMA representation (7) need to be estimated. Since $\mathbf{\Theta}(L) = \mathbf{\Psi}(L)\mathbf{B}^{-1}$ and $\mathbf{\Psi}(L) = \mathbf{A}(L)^{-1} = (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1}$ the estimation of the elements in $\mathbf{\Theta}(L)$ can often be broken down into two steps. First, $\mathbf{A}_1$ is estimated from the reduced form VAR (4). Given $\widehat{\mathbf{A}}_1$, the matrices in $\mathbf{\Psi}(L)$ can be estimated using $\widehat{\mathbf{\Psi}}_k = \widehat{\mathbf{A}}_1^k$. Second, $\mathbf{B}$ is estimated from the SVAR (1). Given $\hat{\mathbf{B}}$ and $\hat{\mathbf{\Psi}}_k$ the estimates of $\mathbf{\Theta}_k$, $k = 0, 1, \ldots,$ are given by $\hat{\mathbf{\Theta}}_k = \widehat{\mathbf{\Psi}}_k \hat{\mathbf{B}}^{-1}$.

1.4 Forecast Error Variance Decompositions

The idea behind constructing forecast error variance decompositions is to determine the proportion of the variability of the errors in forecasting y_1 and y_2 at time $t+s$ based on information available at time t that is due to variability in the structural shocks ε_1 and ε_2 between times t and $t+s$. To accomplish this decomposition, we start with the Wold representation for $\mathbf{y}_{t+s}$

$$\mathbf{y}_{t+s}= \boldsymbol{\mu} + \mathbf{u}_{t+s}+\boldsymbol{\Psi}_1\mathbf{u}_{t+s-1}+\cdots + \boldsymbol{\Psi}_{s-1}\mathbf{u}_{t+1}+\boldsymbol{\Psi}_s\mathbf{u}_t+\boldsymbol{\Psi}_{s+1}\mathbf{u}_{t-1} + \cdots .$$

The best linear forecast of $\mathbf{y}_{t+s}$ based on information available at time t is

$$\hat{\mathbf{y}}_{t+s|t}= \boldsymbol{\mu} + \boldsymbol{\Psi}_s\mathbf{u}_t+\boldsymbol{\Psi}_{s+1}\mathbf{u}_{t-1}+\cdots$$

and the forecast error is

$$\mathbf{y}_{t+s}-\hat{\mathbf{y}}_{t+s|t}= \mathbf{u}_{t+s}+\boldsymbol{\Psi}_1\mathbf{u}_{t+s-1}+\cdots + \boldsymbol{\Psi}_{s-1}\mathbf{u}_{t+1}.$$

Next, using $\boldsymbol{\varepsilon}_t= \mathbf{B}^{-1}\mathbf{u}_t$ we may write the forecast error in terms of the structural shocks

$$\begin{aligned}\mathbf{y}_{t+s}-\hat{\mathbf{y}}_{t+s|t} &= \mathbf{B}^{-1}\boldsymbol{\varepsilon}_{t+s}+\boldsymbol{\Psi}_1\mathbf{B}^{-1}\boldsymbol{\varepsilon}_{t+s-1}+\cdots + \boldsymbol{\Psi}_{s-1}\mathbf{B}^{-1}\boldsymbol{\varepsilon}_{t+1} \\ &= \boldsymbol{\Theta}_0\boldsymbol{\varepsilon}_{t+s}+\boldsymbol{\Theta}_1\boldsymbol{\varepsilon}_{t+s-1}+\cdots + \boldsymbol{\Theta}_{s-1}\boldsymbol{\varepsilon}_{t+1}\end{aligned}$$

The forecast errors equation by equation are given by

$$\begin{bmatrix} y_{1t+s}-\hat{y}_{1t+s|t} \\ y_{2t+s}-\hat{y}_{2t+s|t} \end{bmatrix} = \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+s} \\ \varepsilon_{2t+s} \end{bmatrix} + \cdots + \begin{bmatrix} \theta_{11}^{(s-1)} & \theta_{12}^{(s-1)} \\ \theta_{21}^{(s-1)} & \theta_{22}^{(s-1)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+1} \\ \varepsilon_{2t+1} \end{bmatrix}$$

Focusing on the &rst equation, we have

$$\begin{aligned} y_{1t+s}-\hat{y}_{1t+s|t} &= \theta_{11}^{(0)}\varepsilon_{1t+s}+\cdots+\theta_{11}^{(s-1)}\varepsilon_{1t+1} \\ &\quad +\theta_{12}^{(0)}\varepsilon_{2t+s}+\cdots+\theta_{12}^{(s-1)}\varepsilon_{2t+1} \end{aligned} \tag{10}$$

Since it is assumed that $\boldsymbol{\varepsilon}_t \sim i.i.d.\ (0, \mathbf{D})$ where $\mathbf{D}$ is diagonal, the variance of the forecast error in (10) may be decomposed as

$$\begin{aligned} var(y_{1t+s}-\hat{y}_{1t+s|t}) &= \sigma_1^2(s) \\ &= \sigma_1^2\left(\left(\theta_{11}^{(0)}\right)^2+\cdots+\left(\theta_{11}^{(s-1)}\right)^2\right)+\sigma_2^2\left(\left(\theta_{12}^{(0)}\right)^2+\cdots+\left(\theta_{12}^{(s-1)}\right)^2\right). \end{aligned}$$

The proportion of $\sigma_1^2(s)$ due to shocks in ε_1 is then

$$\rho_{1,1}(s) = \frac{\sigma_1^2\left(\left(\theta_{11}^{(0)}\right)^2+\cdots+\left(\theta_{11}^{(s-1)}\right)^2\right)}{\sigma_1^2(s)}$$

and the proportion of $\sigma_1^2(s)$ due to shocks in ε_2 is

$$\rho_{1,2}(s) = \frac{\sigma_2^2\left(\left(\theta_{12}^{(0)}\right)^2 + \cdots + \left(\theta_{12}^{(s-1)}\right)^2\right)}{\sigma_1^2(s)}.$$

Using similar computations, the forecast error variance decompositions (FEVDs) for y_{2t+s} are

$$\begin{aligned}\rho_{2,1}(s) &= \frac{\sigma_1^2\left(\left(\theta_{21}^{(0)}\right)^2 + \cdots + \left(\theta_{21}^{(s-1)}\right)^2\right)}{\sigma_2^2(s)},\\ \rho_{2,2}(s) &= \frac{\sigma_2^2\left(\left(\theta_{22}^{(0)}\right)^2 + \cdots + \left(\theta_{22}^{(s-1)}\right)^2\right)}{\sigma_2^2(s)},\end{aligned}$$

where

$$\begin{aligned} var(y_{2t+s} - \hat{y}_{2t+s|t}) &= \sigma_2^2(s) \\ &= \sigma_1^2\left(\left(\theta_{21}^{(0)}\right)^2 + \cdots + \left(\theta_{21}^{(s-1)}\right)^2\right) + \sigma_2^2\left(\left(\theta_{22}^{(0)}\right)^2 + \cdots + \left(\theta_{22}^{(s-1)}\right)^2\right).\end{aligned}$$

1.5 Identi&cation Using Recursive Causal Orderings

Consider the bivariate SVAR (1). From the previous discussion of identi&cation, we know that we need at least one restriction on the parameters of (1) for identi&cation. Suppose $b_{12} = 0$ so that $\mathbf{B}$ is lower triangular. That is,

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix}$$

and

$$\mathbf{B}^{-1} = \mathbf{\Theta}_0 = \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix}.$$

This assumptions imposes the restriction that the value y_{2t} does not have a contemporaneous effect on y_{1t}. Since $b_{21} \neq 0$ a priori we allow for the possibility that y_{1t} has a contemporaneous effect on y_{2t}. Further, under this assumption the reduced form VAR errors $\mathbf{u}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t$ become

$$\mathbf{u}_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix}\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} - b_{21}\varepsilon_{1t} \end{bmatrix}.$$

The restriction $b_{12} = 0$ is sufficient to just identify b_{21} and, hence, just identify $\mathbf{B}$. To establish this result, we show how b_{21} can be uniquely identi&ed from the elements of the reduced form covariance matrix $\mathbf{\Omega}$. Since $\mathbf{\Omega} = \mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1\prime}$ and $\mathbf{B}$ is lower triangular, we have

$$\begin{aligned}
\begin{bmatrix} \omega_1^2 & \omega_{12} \\ \omega_{12} & \omega_2^2 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} 1 & -b_{21} \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1^2 & -b_{21}\sigma_1^2 \\ -b_{21}\sigma_1^2 & \sigma_2^2 + b_{21}^2\sigma_1^2 \end{bmatrix}.
\end{aligned}$$

Then, we can solve for b_{21} via

$$b_{21} = -\frac{\omega_{12}}{\omega_1^2} = \rho\frac{\omega_2}{\omega_1},$$

where $\rho = \omega_{12}/\omega_1\omega_2$ is the correlation between u_1 and u_2. Notice that $b_{21} \neq 0$ provided $\rho \neq 0$.

1.5.1 Estimation Procedure

The SMA representation of the SVAR based on a recursive causal ordering may be estimated using the following procedure:

- Estimate the reduced form VAR by OLS equation by equation:

$$\begin{aligned}
\mathbf{y}_t &= \hat{\mathbf{a}}_0 + \widehat{\mathbf{A}}_1 \mathbf{y}_{t-1} + \hat{\mathbf{u}}_t \\
\widehat{\mathbf{\Omega}} &= \frac{1}{T}\sum_{t=1}^{T} \hat{\mathbf{u}}_t \hat{\mathbf{u}}_t'
\end{aligned}$$

- Estimate b_{21} and $\mathbf{B}$ from $\widehat{\mathbf{\Omega}}$:

$$\begin{aligned}
\hat{b}_{21} &= -\frac{\widehat{\omega}_{12}}{\widehat{\omega}_1^2}, \\
\widehat{\mathbf{B}} &= \begin{bmatrix} 1 & 0 \\ \hat{b}_{21} & 1 \end{bmatrix}.
\end{aligned}$$

- Estimate SMA from estimates of $\mathbf{a}_0, \mathbf{A}_1$ and $\mathbf{B}$:

$$\begin{aligned}
\mathbf{y}_t &= \widehat{\boldsymbol{\mu}} + \widehat{\mathbf{\Theta}}(L)\widehat{\varepsilon}_t \\
\widehat{\boldsymbol{\mu}} &= \hat{\mathbf{a}}_0(\mathbf{I}_2 - \widehat{\mathbf{A}}_1)^{-1} \\
\widehat{\mathbf{\Theta}}_k &= \widehat{\mathbf{A}}_1^k \widehat{\mathbf{B}}^{-1}, k = 0, 1, \ldots \\
\widehat{\mathbf{D}} &= \widehat{\mathbf{B}}\widehat{\mathbf{\Omega}}\widehat{\mathbf{B}}'.
\end{aligned}$$

1.5.2 Recovering the SMA representation using the Choleski Factorization of Ω.

The SVAR representation based on a recursive causal ordering may be computed using the Choleski factorization of the reduced form covariance matrix $\mathbf{\Omega}$. Recall, the *Choleski factorization* of the positive semi-de&nite matrix Ω is given by

$$\mathbf{\Omega} = \mathbf{PP}'$$

where

$$\mathbf{P} = \begin{bmatrix} p_{11} & 0 \\ p_{21} & p_{22} \end{bmatrix}$$

is a lower triangular matrix with $p_{ii} \geq 0, i = 1, 2$. A closely related factorization obtained from the Choleski factorization is the *triangular factorization*

$$\mathbf{\Omega} = \mathbf{T\Lambda T}' \tag{11}$$

where $\mathbf{T}$ is a lower triangular matrix with $1's$ along the diagonal and $\mathbf{\Lambda}$ is a diagonal matrix with non-negative elements:

$$\mathbf{T} = \begin{bmatrix} 1 & 0 \\ t_{21} & 1 \end{bmatrix}, \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \lambda_i \geq 0, i = 1, 2.$$

Consider the reduced form VAR

$$\begin{aligned} \mathbf{y}_t &= \mathbf{a}_0 + \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{u}_t, \\ \mathbf{\Omega} &= E[\mathbf{u}_t \mathbf{u}_t'], \end{aligned} \tag{12}$$

and perform the triangular factorization (11) on the covariance matrix $\mathbf{\Omega}$. Now construct a pseudo SVAR model by premultiplying (12) by $\mathbf{T}^{-1}$:

$$\mathbf{T}^{-1}\mathbf{y}_t = \mathbf{T}^{-1}\mathbf{a}_0 + \mathbf{T}^{-1}\mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{T}^{-1}\mathbf{u}_t$$

or

$$\mathbf{B}\mathbf{y}_t = \boldsymbol{\gamma}_0 + \mathbf{\Gamma}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{13}$$

where

$$\mathbf{B} = \mathbf{T}^{-1}, \boldsymbol{\gamma}_0 = \mathbf{T}^{-1}\mathbf{a}_0, \mathbf{\Gamma}_1 = \mathbf{T}^{-1}\mathbf{A}_1, \boldsymbol{\varepsilon}_t = \mathbf{T}^{-1}\mathbf{u}_t.$$

Notice that the pseudo structural errors $\boldsymbol{\varepsilon}_t$ have a diagonal covariance matrix $\mathbf{\Lambda}$ since

$$\begin{aligned} E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] &= \mathbf{T}^{-1} E[\mathbf{u}_t \mathbf{u}_t'] \mathbf{T}^{-1\prime} \\ &= \mathbf{T}^{-1} \mathbf{\Omega} \mathbf{T}^{-1\prime} \\ &= \mathbf{T}^{-1} \mathbf{T\Lambda T}' \mathbf{T}^{-1\prime} \\ &= \mathbf{\Lambda}. \end{aligned}$$

In the pseudo SVAR (13),

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ b_{21} & 0 \end{bmatrix} = \mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 \\ -t_{21} & 1 \end{bmatrix}$$

so that $b_{12} = 0$ and $b_{21} = -t_{21}$[1].

The identi&cation of the SVAR using the triangular factorization depends on the ordering of the variables in $\mathbf{y}_t$. In the above analysis, it is assumed that $\mathbf{y}_t = (y_{1t}, y_{2t})'$ so that y_{1t} comes &rst in the ordering of the variables. When the triangular factorization is conducted and the pseudo SVAR (13) is computed the structural **B** matrix has the form

$$\mathbf{B} = \mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix}$$

where $b_{12} = 0$. If the ordering of the variables is reversed, $\mathbf{y}_t = (y_{2t}, y_{1t})'$, then the recursive causal ordering of the SVAR is reversed and the structural **B** matrix becomes

$$\mathbf{B} = \mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 \\ b_{12} & 1 \end{bmatrix}$$

where $b_{21} = 0$.

1.5.3 Sensitivity Analysis

Since the ordering of the variables in y_t determines the recursive causal structure of the SVAR, and since this identi&cation assumption is not testable a sensitivity analysis is often performed to determine how the structural analysis based on the IRFs and FEVDs are in! uenced by the assumed causal ordering. This sensitivity analysis is based on estimating the SVAR for different orderings of the variables. If the IRFs and FEVDs change considerably for different orderings of the variables in y_t then it is clear that the assumed recursive causal structure heavily in! uences the structural inference.

Another way to determine if the assumed causal ordering in! uences the structural inferences is to look at the residual covariance matrix $\hat{\mathbf{\Omega}}$ from the estimated reduced form VAR (4). If this covariance matrix is close to being diagonal then the estimated value of **B** will be close to diagonal and so the ordering of the variables will not in! uence the structural inference. Of course, eye-balling the elements of $\hat{\mathbf{\Omega}}$ is not a rigorous test. A formal test of the null hypothesis that $\mathbf{\Omega}$ is diagonal can be easily computed using the LM statistic (see Greene (2000), pg. 601)

$$LM = T \cdot \widehat{\rho}^2,$$

where $\widehat{\rho}^2 = \hat{\omega}_{12}^2/\hat{\omega}_1^2\hat{\omega}_2^2$ is the square of the estimated residual correlation between u_{1t} and u_{2t}. Under the null that $\mathbf{\Omega}$ is diagonal, LM has an asymptotic chi-square distribution with 1 degree of freedom.

[1] It is straightforward to show that $-t_{21} = -\omega_{12}/\omega_{11}$.

2 Structural VAR Modeling for I(1) Data that is Not Cointegrated

Let $\mathbf{y}_t = (y_{1t}, y_{2t})'$ be $I(1)$ and not cointegrated. That is, y_{1t} and y_{2t} are both $I(1)$ and there is no linear combination of y_{1t} and y_{2t} that is $I(0)$. In this case, $\Delta\mathbf{y}_t = (\Delta y_{1t}, \Delta y_{2t})'$ is $I(0)$ and is assumed to have the SVAR representation

$$\mathbf{B}\Delta\mathbf{y}_t = \boldsymbol{\gamma}_0 + \boldsymbol{\Gamma}_1\Delta\mathbf{y}_t + \boldsymbol{\varepsilon}_t \tag{14}$$

or

$$\mathbf{B}(L)\Delta\mathbf{y}_t = \boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}_t$$

where $\boldsymbol{\varepsilon}_t \sim i.i.d.\ (0, \mathbf{D})$, $\mathbf{D}$ is diagonal, and $\mathbf{B}(L) = \mathbf{B} - \boldsymbol{\Gamma}_1 L$. Notice that the SVAR model for $\Delta\mathbf{y}_t$ is of the same form as the SVAR for $\mathbf{y}_t$ when we assumed $\mathbf{y}_t$ is $I(0)$.

The reduced form VAR for $\Delta\mathbf{y}_t$ is

$$\Delta\mathbf{y}_t = \mathbf{a}_0 + \mathbf{A}_1\Delta\mathbf{y}_{t-1} + \mathbf{u}_t \tag{15}$$

or

$$\mathbf{A}(L)\Delta\mathbf{y}_t = \mathbf{a}_0 + \mathbf{u}_t$$

where $\boldsymbol{\alpha}_0 = \mathbf{B}^{-1}\boldsymbol{\gamma}_0$, $\mathbf{A}_1 = \mathbf{B}^{-1}\boldsymbol{\Gamma}_1$, $\mathbf{u}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t$, $E[\mathbf{u}_t\mathbf{u}_t'] = \mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1\prime}$ and $\mathbf{A}(L) = \mathbf{I}_2 - \mathbf{A}_1 L$. In (15), it is assumed that the roots of $\det(I_2 - A_1 z) = 0$ lie outside the complex unit circle.

The Wold MA representation of (15) is

$$\Delta\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Psi}(L)\mathbf{u}_t, \tag{16}$$

where $\mu = A(1)^{-1}a_0$ and $\boldsymbol{\Psi}(L) = \mathbf{A}(L)^{-1}$, and the SMA representation is

$$\Delta\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Theta}(L)\boldsymbol{\varepsilon}_t, \tag{17}$$

where $\boldsymbol{\Theta}(L) = \boldsymbol{\Psi}(L)\mathbf{B}^{-1}$.

2.1 Impulse Response Functions

Consider the SMA representation (17) at time $t+s$

$$\begin{bmatrix} \Delta y_{1t+s} \\ \Delta y_{2t+s} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+s} \\ \varepsilon_{2t+s} \end{bmatrix} + \cdots + \begin{bmatrix} \theta_{11}^{(s)} & \theta_{12}^{(s)} \\ \theta_{21}^{(s)} & \theta_{22}^{(s)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \cdots.$$

The structural dynamic multipliers are

$$\begin{aligned} \frac{\partial \Delta y_{1t+s}}{\partial \varepsilon_{1t}} &= \theta_{11}^{(s)}, \frac{\partial \Delta y_{1t+s}}{\partial \varepsilon_{2t}} = \theta_{12}^{(s)} \\ \frac{\partial \Delta y_{2t+s}}{\partial \varepsilon_{1t}} &= \theta_{21}^{(s)}, \frac{\partial \Delta y_{2t+s}}{\partial \varepsilon_{2t}} = \theta_{22}^{(s)} \end{aligned} \tag{18}$$

which give the impact of the structural shocks on the *first difference* of y at horizon $t+s$. Often we are more interested in the impact of the structural shocks on the level of y. Using the fact that

$$y_{it+s} = y_{it-1} + \Delta y_{it} + \Delta y_{it+1} + \cdots + \Delta y_{it+s},\ i = 1, 2$$

we have

$$\begin{aligned}
\frac{\partial y_{it+s}}{\partial \varepsilon_{jt}} &= \frac{\partial \Delta y_{it}}{\partial \varepsilon_{jt}} + \frac{\partial \Delta y_{it+1}}{\partial \varepsilon_{jt}} + \cdots + \frac{\partial \Delta y_{it+s}}{\partial \varepsilon_{jt}} \\
&= \theta_{ij}^{(0)} + \theta_{ij}^{(1)} + \cdots + \theta_{ij}^{(s)} \\
&= \sum_{k=0}^{s} \theta_{ij}^{(k)},\ i, j = 1, 2.
\end{aligned}$$

Hence, the impact of ε_{jt} on y_{it+s} is equal to the cumulative impact of ε_{jt} on Δy_i through horizon s. The long-run impact of a shock to ε_{it} on the level of y_j is then

$$\lim_{s\to\infty} \frac{\partial y_{it+s}}{\partial \varepsilon_{jt}} = \theta_{ij}(1),\ i, j = 1, 2. \tag{19}$$

For stationary y this long-run impact is always zero but for nonstationary y this impact may or may not be zero for some combination of i and j.

2.2 Beveridge-Nelson Decomposition

Using the Wold MA representation for $\Delta \mathbf{y}_t$, the multivariate BN decomposition of y_t is

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{y}_0 + \boldsymbol{\mu}\mathbf{t} + \boldsymbol{\Psi}(1) \sum_{k=1}^{t} \mathbf{u}_k + \tilde{\mathbf{u}}_t - \tilde{\mathbf{u}}_0, \\
\tilde{\mathbf{u}}_t &= \tilde{\boldsymbol{\Psi}}(L)\mathbf{u}_t,
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{\Psi}(1) &= \sum_{k=0}^{\infty} \boldsymbol{\Psi}_k = (\mathbf{I}_2 - \mathbf{A}_1)^{-1} \\
\tilde{\boldsymbol{\Psi}}(L) &= \sum_{k=0}^{\infty} \tilde{\boldsymbol{\Psi}}_k L^k, \tilde{\boldsymbol{\Psi}}_k = -\sum_{j=k+1}^{\infty} \boldsymbol{\Psi}_j
\end{aligned}$$

The BN decomposition gives the multivariate stochastic trends in $\mathbf{y}_t$ in terms of the reduced form error terms $\mathbf{u}_t$

$$\mathbf{TS}_t = \boldsymbol{\Psi}(1) \sum_{k=1}^{t} \mathbf{u}_k.$$

Using $u_t = B^{-1}\varepsilon_t$ and $\boldsymbol{\Theta}(1) = \boldsymbol{\Psi}(1)\mathbf{B}^{-1}$ the multivariate stochastic trends in $\mathbf{y}_t$ may also be represented in terms of the structural errors $\boldsymbol{\varepsilon}_t$

$$\mathbf{TS}_t = \boldsymbol{\Theta}(1) \sum_{k=1}^{t} \boldsymbol{\varepsilon}_k.$$

2.3 Testing Long-run Neutrality[2]

King and Watson (1997), hereafter KW, survey the use of bivariate SVAR models to test some simple long-run neutrality propositions in macroeconomics. The key feature of long-run neutrality propositions is that changes in nominal variables have no effect on real economic variables in the long-run. Some examples are long-run neutrality propositions are: (1) A permanent change in the nominal money stock has no long-run effect on the level of real output; (2) A permanent change in the rate of in! ation has no long-run effect on unemployment (a vertical Phillips curve); (3) A permanent change in the rate of in! ation has no long-run effect on real interest rates (the long-run Fisher relationship). For the analysis in this section, we focus on testing the proposition that money is neutral in the long-run.

Let $\mathbf{y}_t = (y_{1t}, y_{2t})'$ where y_{1t} denotes the natural logarithm of real output and y_{2t} denotes the logarithm of nominal money. KW show that testing long-run neutrality within a SVAR framework requires the data to be $I(1)$. They characterize long-run neutrality of money using the SMA representation for $\Delta\mathbf{y}_t$ written as

$$\begin{aligned}\Delta y_{1t} &= \mu_1 + \theta_{11}(L)\varepsilon_{1t} + \theta_{12}(L)\varepsilon_{2t} \\ \Delta y_{2t} &= \mu_2 + \theta_{21}(L)\varepsilon_{1t} + \theta_{22}(L)\varepsilon_{2t}\end{aligned}$$

where ε_{1t} represents exogenous shocks to output that are uncorrelated with exogenous shocks to nominal money, ε_{2t}, and $\theta_{ij}(L) = \sum_{k=0}^{\infty} \theta_{ij}^{(k)} L^k$ for $i, j = 1, 2$.

Long-run neutrality of money involves the answer to the question: does an unexpected and exogenous permanent change in the level of money (y_2) lead to a permanent change in the level of output (y_1)? If the answer is no, then money is long-run neutral towards output. In terms of the SMA representation, ε_{2t} represents exogenous unexpected changes in money. The permanent effect of ε_{2t} on future values of the level of money is, by (19), $\theta_{22}(1)\varepsilon_{2t}$. Similarly, the permanent effect of ε_{2t} on future values of the level of output is $\theta_{12}(1)\varepsilon_{2t}$. Since the data are in logs, the long-run elasticity of output with respect to permanent changes in money is

$$\eta_{12} = \frac{\theta_{12}(1)}{\theta_{22}(1)}.$$

Hence, money is neutral in the long-run when $\theta_{12}(1) = 0$, or equivalently, when $\eta_{12} = 0$. That is, money is neutral in the long-run when the exogenous shocks that permanently alter money, ε_{2t}, have no permanent effect on output.

The above characterization of long-run neutrality clearly shows why the data need to be $I(1)$ in order to be able to test long-run neutrality. If the data are $I(0)$ then the long-run impacts of shocks to the levels of the series are always zero (see (9) above).

The restriction that money is long-run neutral for output imposes the restriction that the long-run impact matrix $\mathbf{\Theta}(1)$ is lower triangular. The lower triangularity of

[2] This section is based on King and Watson (1997).

$\mathbf{\Theta}(1)$ implies that the multivariate stochastic trend for y_t has the form

$$\begin{bmatrix} TS_{1t} \\ TS_{2t} \end{bmatrix} = \begin{bmatrix} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix} \begin{bmatrix} \sum_{k=1}^{t} \varepsilon_{1k} \\ \sum_{k=1}^{t} \varepsilon_{2k} \end{bmatrix}.$$

Hence, the stochastic trend in y_{1t}, TS_{1t}, only involves shocks to ε_1.

To test the long-run neutrality proposition, the SVAR model for $\Delta \mathbf{y}_t$ must be identi&ed and estimated and then the long-run impact coefficients $\theta_{12}(1)$ and $\theta_{22}(1)$ can be estimated from the derived SMA model. To illustrate, assume that $\Delta \mathbf{y}_t$ has the SVAR representation (14). From the previous discussion of identi&cation, at least one restriction on the parameters of (14) is need for identi&cation. KW consider the following identifying assumptions:

- the impact elasticity of y_1 (output) with respect to y_2 (money), b_{12}, is known,
- the impact elasticity of y_2 (money) with respect to y_1 (output) , b_{21}, is known,
- the long-run elasticity of y_1 (output) with respect to y_2 (money), η_{12}, is known,
- the long-run elasticity of y_2 (money) with respect to y_1 (output), η_{21}, is known.

Instead of reporting results based on a single indentifying restriction, KW summarize results for a wide range of observationally equivalent estimated models based on the (just) identifying assumptions listed above[3]. This approach allows the reader to gauge the robustness of conclusions about long-run neutrality to speci&c assumptions about the values of $b_{12}, b_{21}, \eta_{12}$ and η_{22}.

2.3.1 Estimating the SVAR assuming b_{12} or b_{21} is known

Consider the estimating the SVAR (14) under the restriction that b_{12} is known. Given that b_{12} is known the SVAR (14) may be rewritten as

$$\begin{aligned} \Delta y_{1t} + b_{12}\Delta y_{2t} &= \gamma_{10} + \gamma_{11}\Delta y_{1t-1} + \gamma_{12}\Delta y_{2t-1} + \varepsilon_{1t} \\ \Delta y_{2t} &= \beta_{20} - b_{21}\Delta y_{1t} + \gamma_{21}\Delta y_{1t-1} + \gamma_{22}\Delta y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

The &rst equation () may be estimated by OLS since only lagged values of Δy_1 and Δy_2 are on the right-hand-side. However, the second equation () cannot be estimated by OLS because Δy_{1t} will be correlated with ε_{2t} unless $b_{12} = 0$. If $b_{12} \neq 0$, the second equation may be estimated by instrumental variables (IV) using the residual from the estimated &rst equation, $\hat{\varepsilon}_{1t}$, together with Δy_{1t-1} and Δy_{2t-1} as instruments The residual $\hat{\varepsilon}_{1t}$ is a valid instrument because

$$p \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \hat{\varepsilon}_{1t} \Delta y_{1t} \neq 0$$

[3] Since each of the identifying assumptions just identi&es the SVAR, each SVAR has the same reduced form VAR and hence each SVAR model is observationally equivalent.

since $E[\varepsilon_{1t}\Delta y_{1t}] \neq 0$ and

$$p\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{1t}\varepsilon_{2t} = 0$$

since $E[\varepsilon_{1t}\varepsilon_{2t}] = 0$. Hausman, Newey and Taylor (1987) show that this procedure produces the maximum likelihood estimates of the parameters of the SVAR.

3 Structural VARs with Combinations of I(1) and I(0) Data

Consider two observed series y_{1t} and y_{2t} such that y_{1t} is $I(1)$ and y_{2t} is $I(0)$. For example, in the analysis in Blanchard and Quah (1989), hereafter BQ, y_1 is the log of real GDP and y_2 is the unemployment rate. De&ne $\mathbf{y}_t = (\Delta y_{1t}, y_{2t})'$ so that $\mathbf{y}_t$ is $I(0)$. Suppose $\mathbf{y}_t$ has the structural representations

$$\begin{aligned}\mathbf{B}\mathbf{y}_t &= \boldsymbol{\gamma}_0 + \boldsymbol{\Gamma}_1\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \\ y_t &= \boldsymbol{\mu} + \boldsymbol{\Theta}(L)\boldsymbol{\varepsilon}_t,\end{aligned}$$

and reduced form representations

$$\begin{aligned}\mathbf{y}_t &= \mathbf{a}_0 + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{u}_t, \\ &= \boldsymbol{\mu} + \boldsymbol{\Psi}(L)\mathbf{u}_t,\end{aligned}$$

where $E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] = \mathbf{D}, \mathbf{D}$ is diagonal, $E[\mathbf{u}_t\mathbf{u}_t'] = \boldsymbol{\Omega} = \mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1'}, \boldsymbol{\Psi}(L) = (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1}$ and $\boldsymbol{\Theta}(L) =.\boldsymbol{\Psi}(L)\mathbf{B}^{-1}$. Regarding the structural innovations, BQ loosely interpret ε_{1t} as a (permanent) supply shock since it is the innovation to the $I(1)$ real output series y_{1t} and interpret ε_{2t} as a (transitory) demand shock since it is the innovation to the $I(0)$ unemployment series.

The IRFs are given by

$$\begin{aligned}\frac{\partial \Delta y_{1t+s}}{\partial \varepsilon_{1t}} &= \theta_{11}^{(s)}, \frac{\partial \Delta y_{1t+s}}{\partial \varepsilon_{2t}} = \theta_{12}^{(s)} \\ \frac{\partial y_{2t+s}}{\partial \varepsilon_{1t}} &= \theta_{21}^{(s)}, \frac{\partial y_{2t+s}}{\partial \varepsilon_{2t}} = \theta_{22}^{(s)}\end{aligned} \quad (20)$$

Since y_{1t} is $I(1)$, the long-run impacts on the *level* of y_1 of shocks to ε_1 and ε_2 are

$$\begin{aligned}\lim_{s\to\infty}\frac{\partial y_{1t+s}}{\partial \varepsilon_{1t}} &= \theta_{11}(1) = \sum_{s=0}^{\infty}\theta_{11}^{(s)}, \\ \lim_{s\to\infty}\frac{\partial y_{1t+s}}{\partial \varepsilon_{2t}} &= \theta_{12}(1) = \sum_{s=0}^{\infty}\theta_{12}^{(s)}.\end{aligned} \quad (21)$$

Since y_{2t} is $I(0)$, the long-run impacts on the *level* of y_2 of shocks to ε_1 and ε_2 are zero:

$$\lim_{s\to\infty} \frac{\partial y_{2t+s}}{\partial \varepsilon_{jt}} = \lim_{s\to\infty} \theta_{ij}^{(s)} = 0. \tag{22}$$

For y_2, $\theta_{21}(1) = \sum_{s=0}^{\infty} \theta_{21}^{(s)}$ and $\theta_{22}(1) = \sum_{s=0}^{\infty} \theta_{22}^{(s)}$ represent the *cumulative impact* of shocks to ε_1 and ε_2 on the level of y_2.

3.1 Identifying the SVAR Using Long-Run Restrictions

BQ achieve identi&cation of the SVAR/SMA by assuming that demand shocks (shocks to ε_2) have no long-run impact on the level of output or unemployment. They allow supply shocks (shocks to ε_1) to have a long-run impact on the level of output but not on the level of unemployment. In terms of the long-run impacts discussed in the previous section, BQ long-run restriction may be represented as follows. Using (21), the restriction that shocks to ε_2 have no long-run impact on the level out y_1 implies that

$$\theta_{12}(1) = \sum_{s=0}^{\infty} \theta_{12}^{(s)} = 0. \tag{23}$$

The restriction that shocks to ε_1 and ε_2 have no long-run effect on the level of y_2 is just a restatement of the result in (22) which holds because y_2 is $I(0)$.

The long-run restriction (23) makes the long-run impact matrix $\mathbf{\Theta}(1)$ lower triangular

$$\mathbf{\Theta}(1) = \begin{bmatrix} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix}$$

To see how the lower triangularity of $\mathbf{\Theta}(1)$ can be used to identify $\mathbf{B}$ in the SVAR, consider the *long-run covariance matrix*[4] of $\mathbf{y}_t$ de&ned from the Wold MA representation

$$\begin{aligned} \mathbf{\Lambda} &= \mathbf{\Psi}(1)\mathbf{\Omega}(1)' \\ &= (\mathbf{I}_2 - \mathbf{A}_1)^{-1}\mathbf{\Omega}(\mathbf{I}_2 - \mathbf{A}_1)^{-1\prime}. \end{aligned} \tag{24}$$

Since $\mathbf{\Omega} = \mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1\prime}$ and $\mathbf{\Theta}(1) = \mathbf{\Psi}(1)\mathbf{B}^{-1} = (\mathbf{I}_2 - \mathbf{A}_1)^{-1}\mathbf{B}^{-1}$ the long-run covariance matrix $\mathbf{\Lambda}$ may be re-expressed as

$$\begin{aligned} \mathbf{\Lambda} &= (\mathbf{I}_2 - \mathbf{A}_1)^{-1}\mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1\prime}(\mathbf{I}_2 - \mathbf{A}_1)^{-1\prime} \\ &= \mathbf{\Theta}(1)\mathbf{D}\mathbf{\Theta}(1)'. \end{aligned} \tag{25}$$

In order to identify $\mathbf{B}$, BQ make the additional assumption

$$\mathbf{D} = \mathbf{I}_2 \tag{26}$$

[4]Recall, the long-run covariance of $\mathbf{y}_t$ is the asymptotic covariance of $\sqrt{T}(\bar{\mathbf{y}} - \boldsymbol{\mu})$.

so that the structural shocks ε_{1t} and ε_{2t} have unit variances. Inserting (26) into (25), the long-run variance becomes

$$\mathbf{\Lambda} = \mathbf{\Theta}(1)\mathbf{\Theta}(1)'. \tag{27}$$

Notice that since $\mathbf{\Theta}(1)$ is lower triangular, the factorization in (27) can be obtained using the Choleski factorization; that is, $\mathbf{\Theta}(1)$ can be computed as the lower triangular Choleski factor of $\mathbf{\Lambda}$. Given that $\mathbf{\Theta}(1)$ can be computed, $\mathbf{B}$ can then be computed using $\mathbf{\Theta}(1) = \mathbf{\Psi}(1)\mathbf{B}^{-1} = (\mathbf{I}_2 - \mathbf{A}_1)^{-1}\mathbf{B}^{-1}$ so that

$$\mathbf{B} = \left[(\mathbf{I}_2 - \mathbf{A}_1)\mathbf{\Theta}(1)\right]^{-1}.$$

3.2 Estimating the SVAR in the Presence of Long-Run Restrictions

The estimation of $\mathbf{B}$ and $\mathbf{\Theta}(L)$ using the BQ identi&cation scheme can be accomplished in two steps.

- Estimate the reduced form VAR by OLS equation by equation:

$$\begin{aligned} \mathbf{y}_t &= \hat{\mathbf{a}}_0 + \widehat{\mathbf{A}}_1\mathbf{y}_{t-1} + \hat{\mathbf{u}}_t \\ \hat{\mathbf{\Omega}} &= \frac{1}{T}\sum_{t=1}^{T}\hat{\mathbf{u}}_t\hat{\mathbf{u}}_t' \end{aligned}$$

- Compute a parametric estimate of the long-run covariance matrix:

$$\hat{\mathbf{\Lambda}} = (\mathbf{I}_2 - \hat{\mathbf{A}}_1)^{-1}\hat{\mathbf{\Omega}}(\mathbf{I}_2 - \hat{\mathbf{A}}_1)^{-1\prime}$$

- Compute the Choleski factorization of $\hat{\mathbf{\Lambda}}$:

$$\hat{\mathbf{\Lambda}} = \hat{\mathbf{P}}\hat{\mathbf{P}}'$$

- De&ne the estimate of $\mathbf{\Theta}(1)$ as the lower triangular Choleski factor of $\hat{\mathbf{\Lambda}}$:

$$\hat{\mathbf{\Theta}}(1) = \hat{\mathbf{P}}$$

- Estimate $\mathbf{B}$ using

$$\hat{\mathbf{B}} = \left[(\mathbf{I}_2 - \hat{\mathbf{A}}_1)\hat{\mathbf{\Theta}}(1)\right]^{-1}$$

- Estimate $\mathbf{\Theta}_k$ using

$$\begin{aligned} \hat{\mathbf{\Theta}}_k &= \hat{\mathbf{\Psi}}_k\hat{\mathbf{B}}^{-1} \\ &= \hat{\mathbf{A}}_1^k\hat{\mathbf{B}}^{-1}. \end{aligned}$$

4 References

References

[1] Blanchard, O.J. and D.Quah (1989), “The Dynamic Effects of Aggregate Demand and Supply Disturbances,” *American Economic Review*, Vol. 79, 1146-64.

[2] Cochrane, J.H. (1994), “Permanent and Transitory Components of GNP and Stock Prices,” *Quarterly Journal of Economics*, Vol. 109, 241-265.

[3] Gali, J. (1992), “How Well Does the IS-LM Model Fit Postwar U.S. Data?” *Quarterly Journal of Economics*, Vol. 107, 709-38.

[4] Gonzalo, J. and S. Ng (1999), “A Systematic Framework for Analyzing the Dynamic Effects of Permanent and Transitory Shocks,” unpublished manuscript, Department of Economics, Carlos III de Madrid.

[5] Greene, W. (2000), *Econometric Analysis*, Fourth Edition. Prentice Hall: New Jersey.

[6] Hausman, J.A., W.K. Newey, and W.E. Taylor (1987), “Efficient Estimation and Identi&cation of Simultaneous Equation Models with Covariance Restrictions,” *Econometrica*, Vol. 55, 849-74.

[7] King, R.G. and M. W. Watson, (1997), “Testing Long-Run Neutrality,” Federal Reserve Bank of Richmond *Economic Quarterly*, Vol. 83(3), 69-101.

[8] Levtchenkova, S., A. Pagan, and J. Robertson, (1999): “Shocking Stories,” chapter 3 in M. McAleer and L. Oxley (eds.) *Practical Issues in Cointegration Analysis*s. Basil Blackwell: Oxford.

[9] Shapiro, M. and M.W. Watson, (1988), “Sources of Business Cycle Fluctuations,” *NBER Macroeconomics Annual*, Vo. 3, 111-56.

[10] Watson, M.W. (1994), “Vector Autoregressions and Cointegration,” in R. Engle and D. McFadden (eds.), *Handbook of Econometrics*, Vol IV. Elsevier: Amsterdam.

[11] Watson, M.W. (1998), “Lecture notes on VAR models”, mimeo, Department of Economics, Princeton University.