

Structural vector autoregression models

by Kevin Kotzé

Sims (1980) introduced structural vector autoregressive (SVAR) models as an alternative to the large-scale macroeconometric models used in academic and policy work at the time. This was after he questioned the idea of developing sophisticated econometric models that were identified via what he called *incredible* (non-justified) exclusionary restrictions, which were neither innocuous nor essential to the construction of a model that was then used for policy analysis and forecasting purposes.

Since that point in time, this methodology has gained widespread use in applied time series research. These models are used today to address a number of important questions that include: What factors influence business cycle fluctuations? Are demand and supply disturbances equally important economic forces that influence cyclical behaviour over time and across countries? What is the effect of a monetary policy shock? Do oil prices shocks contribute to recessions? How does the central bank respond to shocks in macroeconomic variables?, etc.

This chapter explores the relationship between the reduced-form VAR, which was introduced previously, and the structural form of the model. In particular, we explain how one can identify the structural shocks from the reduced form VAR so that they match their theoretical counterparts. To obtain such behavioural relationships, we will investigate different identification methods that rely on short-run and long-run restrictions. One of the important features of SVAR models is that contemporaneous variables may be treated as explanatory variables, which is particularly important when the frequency of the data is relatively long (i.e. quarterly). In addition, these models also allow for one to impose several highly specific restrictions on the parameters in the coefficient and residual covariance matrices. This would allow for one to evaluate the effect of an independent shock, when the off-diagonal elements in the residual covariance matrix are set to zero.

The strategy that has been employed to construct a SVAR model is as follows: The first stage requires that the analyst use *a priori* knowledge to decide upon what variables should be included in the reduced-form model. Thereafter, the lag length of the autoregression, choice of deterministic components and appropriate treatment of the nonstationary components should be decided upon. This would give rise to an appropriate dynamic specification, which allows for endogenous interactions among the variables. Economic hypotheses could then be formulated and tested, while the historical dynamics of the data could also be examined, after sufficient structure has been imposed on the model. For example, it may then be possible to consider the in-sample effects of an independent shock on the rest of the system. This procedure may involve the computation of impulse responses and variance decompositions.

SVAR models have the advantage over traditional large-scale macroeconometric models in that the results are not hidden by a large and complicated structure (a *black box*), but are easy to obtain and interpret. Sims (1980) argued that SVAR models provide a more systematic approach to imposing restrictions, which could enable the researcher to capture empirical regularities, which remain hidden when using the techniques that were previously applied. In contrast, the results from policy exercises that use large-scale macroeconometric models are hard to compare and recreate, and can easily be amended by their users with judgemental ex-post decisions. In addition, the lack of consensus about the appropriate specification for a simultaneous equation model (that were developed under the Cowles commission) led to the relative popularity of SVAR models.

In what seems like an ever-growing area of research, few contributions have been as influential as that of Sims's description of the SVAR approach.¹ In what follows, we consider this approach in more detail, along with a number of practical issues that arise with the estimation of SVAR models.

1 Incorporating contemporaneous variables

To show how we could incorporate contemporaneous variables in a multivariate VAR, we start off by assuming that each variable is symmetrical. Thereafter, where we assume that we have a bi-variate model for $y_{1,t}$ and $y_{2,t}$, we may wish to allow for $y_{1,t}$ to be affected by current and past realizations of $y_{2,t}$, while $y_{2,t}$ is affected by current and past realizations of $y_{1,t}$. Where we also want to incorporate autoregressive lags of the left-hand-side variables, we could write

$$\begin{aligned} y_{1,t} &= b_{10} - b_{12}y_{2,t} + \gamma_{11}y_{1,t-1} + \gamma_{12}y_{2,t-1} + \varepsilon_{1,t} \\ y_{2,t} &= b_{20} - b_{21}y_{1,t} + \gamma_{21}y_{1,t-1} + \gamma_{22}y_{2,t-1} + \varepsilon_{2,t} \end{aligned} \quad (1.1)$$

In this case we assume that both $y_{1,t}$ and $y_{2,t}$ are stationary, where $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are white noise with constant variance terms, σ_1 and σ_2 . It is also assumed that $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are uncorrelated, which would allow for us to identify the effect of each independent shock. Hence, the covariance elements in Σ_ε are set to zero, such that the variance-covariance matrix of the residuals is

$$\Sigma_\varepsilon \sim \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

Given these conditions, the coefficient b_{12} would describe the contemporaneous effect of a change in $y_{2,t}$ on $y_{1,t}$. Similarly, b_{21} would describe the contemporaneous effect of a change in $y_{1,t}$ on $y_{2,t}$. Note that this would imply that there will be an indirect contemporaneous effect of $\varepsilon_{1,t}$ on $y_{2,t}$ if $b_{21} \neq 0$, while $\varepsilon_{2,t}$ affects $y_{1,t}$ if $b_{12} \neq 0$. This allows for a more elaborate characterisation of the dynamics by the model, however, it does also present a number of interesting challenges when we turn our attention to parameter estimation.

1.1 Identification of parameters

To express the above *structural-form* of the model as a *reduced-form* expression, we could take all the contemporaneous variables to the left-hand-side, before inserting them in the vector \mathbf{y}_t . This would allow us to write the model in (1.1) as,

$$B\mathbf{y}_t = \Gamma_0 + \Gamma_1\mathbf{y}_{t-1} + \varepsilon_t$$

where

$$\begin{aligned} B &= \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, \quad \mathbf{y}_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix}, \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} \\ \Gamma_1 &= \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \quad \text{and} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \end{aligned}$$

Premultiplication by B^{-1} gives us the VAR in *reduced-form*:

$$\mathbf{y}_t = A_0 + A_1\mathbf{y}_{t-1} + \mathbf{u}_t$$

where $A_0 = B^{-1}\Gamma_0$, $A_1 = B^{-1}\Gamma_1$ and $\mathbf{u}_t = B^{-1}\varepsilon_t$. This expression could be expressed in a similar way to the original model after using matrix multiplication, where $a_{i,0}$ is the i element in A_0 and $a_{i,j}$ is row i column j of matrix A_1 . Similarly, \mathbf{u}_t would contain the elements $u_{1,t}$ and $u_{2,t}$, such that

$$\begin{aligned} y_{1,t} &= a_{10} + a_{11}y_{1,t-1} + a_{12}y_{2,t-1} + u_{1,t} \\ y_{2,t} &= a_{20} + a_{21}y_{1,t-1} + a_{22}y_{2,t-1} + u_{2,t} \end{aligned} \quad (1.2)$$

Using the relationship $\mathbf{u}_t = B^{-1}\varepsilon_t$, or:

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

We can show that,

$$\begin{aligned} u_{1,t} &= (\varepsilon_{1,t} - b_{12}\varepsilon_{2,t}) / (1 - b_{12}b_{21}) \\ u_{2,t} &= (\varepsilon_{2,t} - b_{21}\varepsilon_{1,t}) / (1 - b_{12}b_{21}) \end{aligned}$$

Since $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are white noise processes, the residuals $u_{1,t}$ and $u_{2,t}$ have zero means, constant variances, and little autocorrelation. However, as \mathbf{u}_t is dependent upon both $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, there may be some evidence of covariation. The covariance of the two terms could be represented as,

$$\begin{aligned} \text{cov}[u_{1,t}, u_{2,t}] &= \mathbb{E}[(\varepsilon_{1,t} - b_{12}\varepsilon_{2,t})(\varepsilon_{2,t} - b_{21}\varepsilon_{1,t})] / (1 - b_{12}b_{21})^2 \\ &= -[(b_{21}\sigma_1^2 + b_{12}\sigma_2^2)] / (1 - b_{12}b_{21})^2 \end{aligned}$$

Since they are all time invariant, the variance/covariance matrix will be,

$$\Sigma_{\mathbf{u}} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

where $\text{var}[u_{i,t}] = \sigma_{ii}$ and $\sigma_{12} = \sigma_{21} = \text{cov}[u_{1,t}, u_{2,t}]$.

1.2 Estimation of reduced-form parameters

Unfortunately, the equations in the *structural-form* of the model can't be estimated directly, due to feedback effects from contemporaneous variables. However, the *reduced-form* as expressed in (1.2) contains only predetermined variables. In addition, the error terms are serially uncorrelated with constant variance. Hence we can use OLS to estimate the parameters in this form of the model, as it would be consistent and asymptotically efficient.

This would allow for us to generate values for the residuals, $u_{1,t}$ and $u_{2,t}$. In addition, we would also obtain coefficient values for the A_0 and A_1 matrices, which could possibly be used to recover the parameters *structural-form*, given the relationships between these expressions. However, as the *structural-form* contains ten parameters and the *reduced-form* contains nine parameters, this is not going to be possible, as there is no mapping that enables us to obtain all the *structural-form* parameters from the *reduced-form* parameters.²

However, it may be possible to show that if one variable in the *structural-form* is restricted to a calibrated value, then the structural system could be exactly identified. To accomplish this objective we assume that $b_{21} = 0$ in the *structural-form*, before making use of the method of recursive estimation (Sims 1980). This would imply that the *structural-form* of the model would be expressed as,

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

$$\text{such that } B^{-1} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix}$$

Premultiplying by B^{-1} yields

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

Given this expression, it is worth taking note of the left-hand-side variables and the error terms, where it would suggest that by setting $b_{21} = 0$, the shocks from $\varepsilon_{1,t}$ do not effect contemporaneous values of $y_{2,t}$. Furthermore, by returning to the relationship $\mathbf{u}_t = B^{-1}\varepsilon_t$, we note that

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

In this case we have $\varepsilon_{2,t} = u_{1,t}$, and after using $b_{12} = -\text{cov}[u_{1,t}, u_{2,t}]/\sigma_2^2$, we are able to get values for $\varepsilon_{1,t} = b_{12}\varepsilon_{2,t} + u_{1,t}$. In addition, after using the *reduced-form*, where all the coefficient matrices are premultiplied by B^{-1} , we are able to show that

$$\begin{aligned} a_{10} &= b_{10} - b_{12}b_{20} \\ a_{12} &= \gamma_{12} - b_{12}\gamma_{22} \\ a_{21} &= \gamma_{21} \\ a_{11} &= \gamma_{11} - b_{12}\gamma_{21} \\ a_{20} &= b_{20} \\ a_{22} &= \gamma_{22} \\ \text{var}[u_1] &= \sigma_1 + b_{12}\sigma_2 \\ \text{var}[u_2] &= \sigma_2 \\ \text{cov}[u_1, u_2] &= -b_{12}\sigma_2 \end{aligned}$$

In this example, we were able to recover the errors, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, using the relationship, $u_{1,t} = \varepsilon_{1,t} - b_{12}\varepsilon_{2,t}$ and $u_{2,t} = \varepsilon_{2,t}$. Hence, when $b_{21} = 0$, then $y_{1,t}$ does not have a contemporaneous effect on $y_{2,t}$ and $\varepsilon_{1,t}$ does not affect $y_{2,t}$. In addition, the observed values of $u_{2,t}$ would be attributed to the shocks in the $y_{2,t}$ equation.

This procedure of setting the lower triangle of the B coefficient matrix equal to zero is an application of the Cholesky decomposition. It turns out that the number of restrictions that we need to impose is equivalent to the number of terms in the lower (or upper) triangle of the B matrix, which is $[(K^2 - K)/2]$, where K represents the number of endogenous variables in the model. The alternative ordering of the Cholesky decomposition would involve setting $b_{12} = 0$, which would be equivalent to setting the values in the upper triangle equal to zero.

2 Impulse Response Functions

To show how each of the shocks affect the respective variables in an autoregressive structure, we would usually choose to express the model in its moving average form. In the univariate case, we were able to show that a stable AR(p) process has a corresponding MA(q) representation, which may be used to derive impact multipliers and impulse response functions. For example, the stationary AR(1) model, $y_t = \phi y_{t-1} + \varepsilon_t$ could be represented by the MA(∞) expression

$$y_t = \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}$$

Just as every stable autoregressive, AR(p), model has a moving average, MA(q), representation; every stable vector autoregressive model, VAR(p), has a vector moving average, VMA(q), representation.

Therefore, it would be possible to show that the stable *reduced-form* bivariate VAR(1) model,

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

Has a vector moving average representation,

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \cdot \begin{bmatrix} u_{1,t-i} \\ u_{2,t-i} \end{bmatrix} \quad (2.1)$$

where μ_1 and μ_2 are mean values for $y_{1,t}$ and $y_{2,t}$. To extend this representation to a SVAR model, we would need to make use of the explicit relationships that exist between the respective *structural* and *reduced-form* of the model. For example, since $\mathbf{u}_t = B^{-1}\varepsilon_t$, where we are able to expand upon B^{-1} to establish,

$$B^{-1} = \frac{1}{\det} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} = \frac{1}{1 - b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

we have,

$$\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} = \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}^i \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \quad (2.2)$$

Such that after combining (2.2) and (2.1), the moving average representation of the SVAR model could be written as,

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \cdot \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-i} \\ \varepsilon_{2,t-i} \end{bmatrix}$$

This expression could then be used to describe the effect of a shock in ε_t on the endogenous variables. As the notation has become somewhat cumbersome, we may summarise the impact multipliers, which describe the effect of shocks on the endogenous variables, with the aid of matrix Θ_i , such that

$$\Theta_i = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} \\ \theta_{2,1} & \theta_{2,2} \end{bmatrix}_i = \frac{A_1^i}{1 - b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

where $\mu = [\mu_1 \ \mu_2]'$. This would imply that we could express $\mathbf{y}_t = [y_{1,t} \ y_{2,t}]'$ as a VMA(∞), using the notation

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\Theta}_i \boldsymbol{\varepsilon}_{t-i}$$

This is a particularly useful expression, as the $\boldsymbol{\Theta}_i$ matrix describes the effects of the shocks, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ on the entire paths of $y_{1,t}$ and $y_{2,t}$. For example, where the numbers in brackets refer to the lags of $\theta_{jk}(i)$:

- $\theta_{12}(0)$ is the instant impact of 1 unit change in $\varepsilon_{2,t}$ on $y_{1,t}$
- $\theta_{11}(1)$ is the instant impact of 1 unit change in $\varepsilon_{1,t-1}$ on $y_{1,t}$
- $\theta_{12}(1)$ is the instant impact of 1 unit change in $\varepsilon_{2,t-1}$ on $y_{1,t}$

The impact multipliers $\theta_{11}(i)$, $\theta_{12}(i)$, $\theta_{21}(i)$ and $\theta_{22}(i)$ may then be used to generate the impulse response functions for different values of i . This may be displayed with the aid of a visual representation that describes the behaviour of $y_{1,t}$ and $y_{2,t}$ in response to the various shocks, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$. To avoid the problem of an underidentified system we use the Cholesky decomposition;

$$\begin{aligned} u_{1,t} &= \varepsilon_{1,t} - b_{12}\varepsilon_{2,t} \\ u_{2,t} &= \varepsilon_{2,t} \end{aligned}$$

Using this representation, we note once again that all the errors from $u_{2,t}$ are attributed to $\varepsilon_{2,t}$, where we can derive $\varepsilon_{1,t}$ using b_{12} , $u_{1,t}$ and $\varepsilon_{2,t}$. It is also important to note that while the Cholesky decomposition constrains the system, such that $\varepsilon_{1,t}$ has no direct effect on $y_{2,t}$, it would nevertheless be possible to show that lagged values of $y_{1,t}$ would affect the contemporaneous value of $y_{2,t}$.

At this point it should also be noted that the ordering of the Cholesky decomposition (i.e. whether to set b_{12} or b_{21} to zero) could affect the results of the impulse response function. The degree to which the ordering affects the results depends upon the magnitude of the correlation between $u_{1,t}$ and $u_{2,t}$, which may be summarised as, $\rho_{12} = \sigma_{12} / (\sqrt{\sigma_{11}}\sqrt{\sigma_{22}})$.

In those cases where the correlation coefficient, ρ_{12} , is close to zero then the ordering is immaterial. In contrast, when the correlation is close to unity then it would be inappropriate to attribute the shock to a single source. In addition, if the correlation is between zero and one then you usually need to consider both ordering - if the results are different then you need to investigate further.

To avoid these problems it is usually a good idea (wherever possible) to relate the ordering to a theoretical consideration. For example, when modelling exchange rates, we could conceive that a shock to the U.S. exchange rate may affect the South African exchange rate contemporaneously, while it is unlikely that a shock to the South African exchange rate would have a contemporaneous affect the U.S. exchange rate.

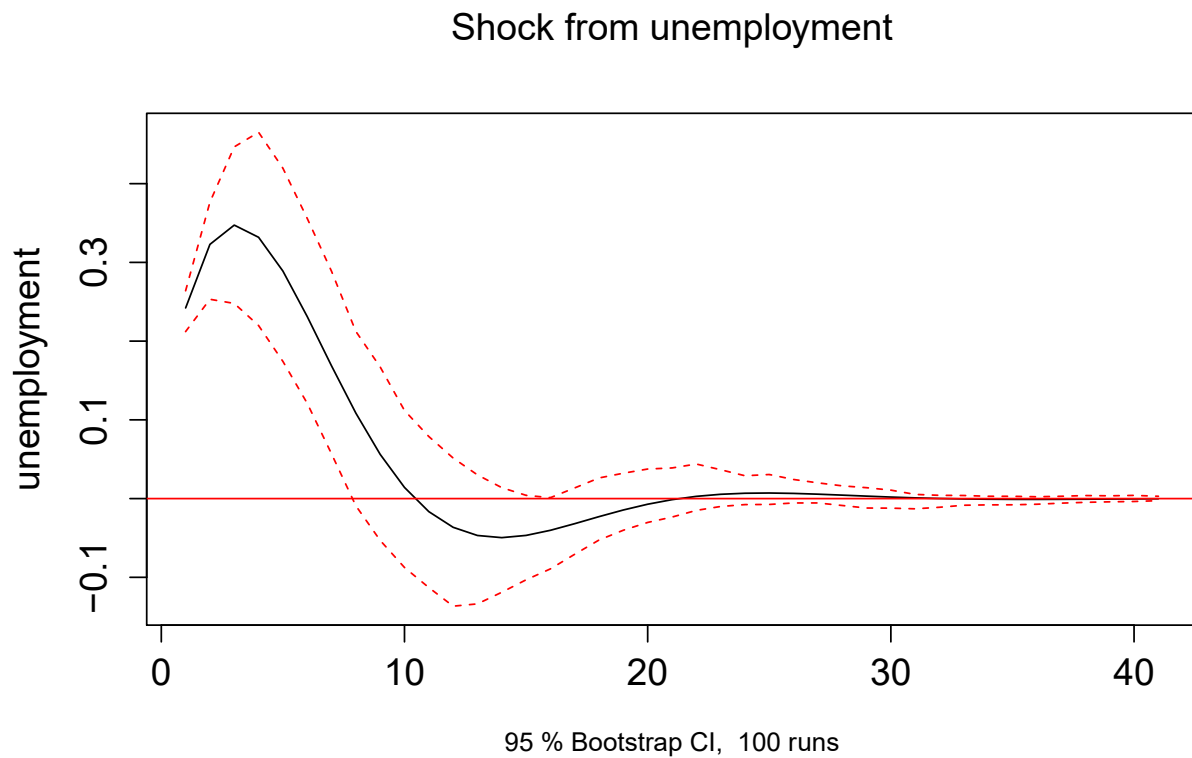
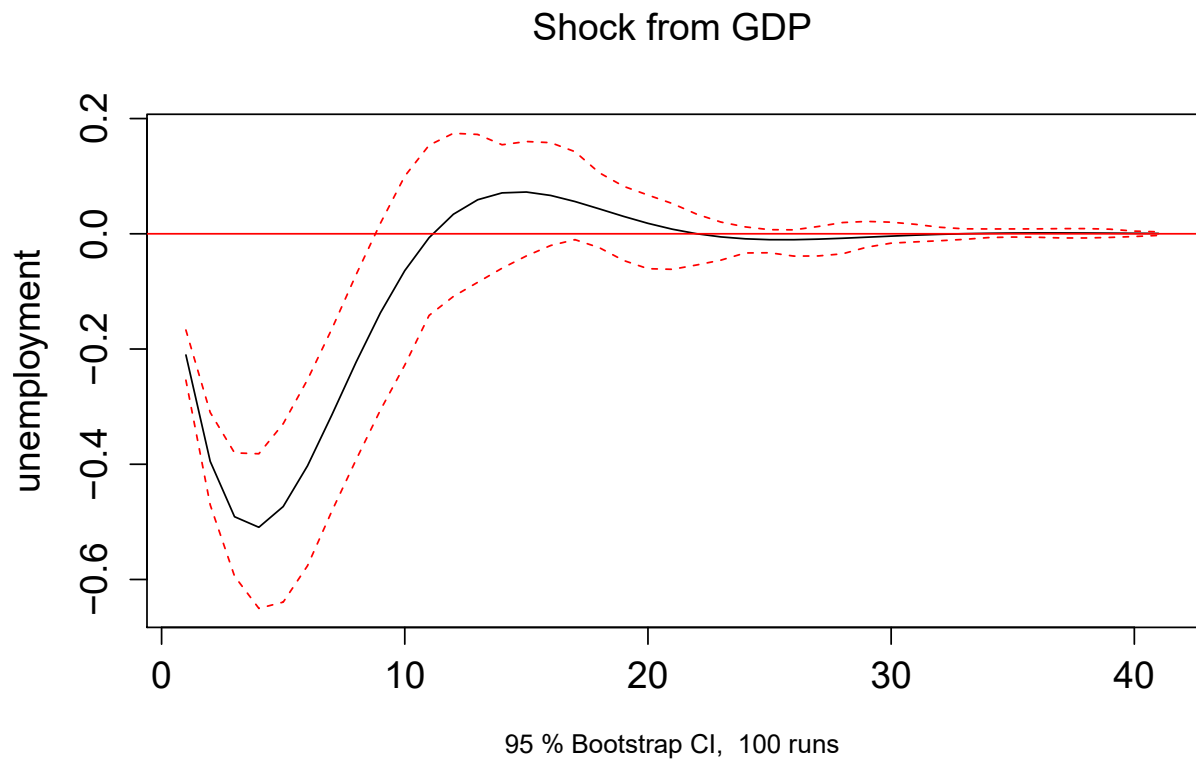


Figure 1: **Impulse response functions**

Figure 1 contains two examples of impulse response functions. The top panel shows the extent of a decrease in unemployment that follows a positive output shock of one standard deviation, while the bottom panel shows the extent of the persistence in unemployment that follows a positive unemployment shock.

3 Forecast error variance decompositions

The forecast error variance decomposition describes the proportion of the expected variance in a variable that is due to each of the structural shocks in the model, at different horizons. To derive the forecast error variance decomposition, we may choose to derive a number of forecasts at different horizons. Therefore, after estimating the coefficient matrices A_0 and A_1 , we would be able to derive an h -step ahead forecast for $\mathbb{E}[\mathbf{y}_{t+h}]$ conditional on \mathbf{y}_t . For example, the conditional expectation of $\mathbb{E}[\mathbf{y}_{t+1}]$ is,

$$\mathbb{E}_t[\mathbf{y}_{t+1}] = A_0 + A_1\mathbf{y}_t$$

while the conditional expectation of $\mathbb{E}[\mathbf{y}_{t+2}]$ is

$$\mathbb{E}_t[\mathbf{y}_{t+2}] = [I + A_1]A_0 + A_1^2\mathbf{y}_t \quad (3.1)$$

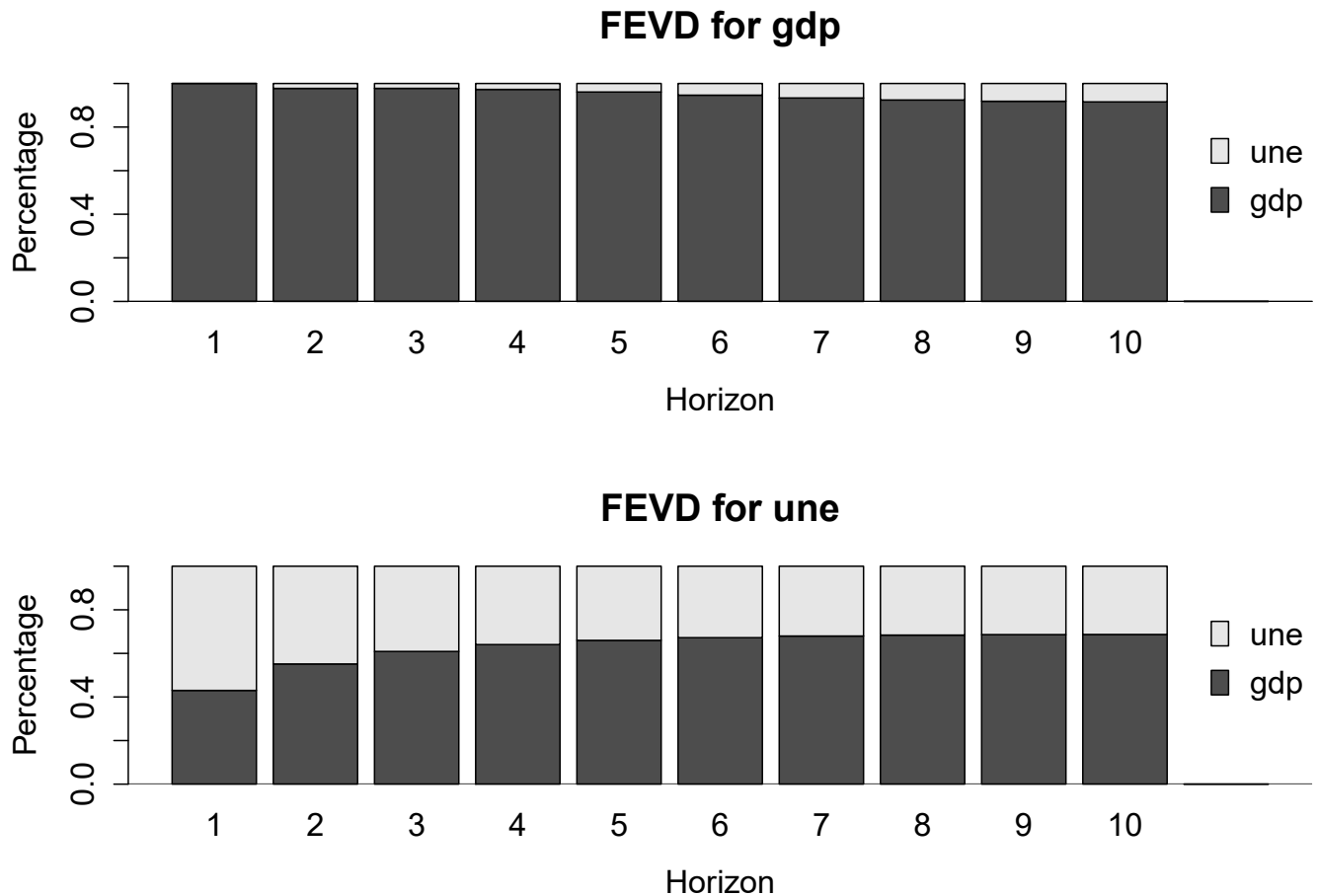
such that the conditional expectation of $\mathbb{E}[\mathbf{y}_{t+H}]$ is

$$\mathbb{E}_t[\mathbf{y}_{t+H}] = [I + A_1 + A_1^2 + \dots + A_1^{H-1}]A_0 + A_1^H\mathbf{y}_t$$

This would allow for the calculation of the one-step ahead forecast error, which would be $(\mathbf{y}_{t+1} - \mathbb{E}_t[\mathbf{y}_{t+1}])$. Note that as $\mathbb{E}_t[\mathbf{y}_{t+1}] = A_0 + A_1\mathbf{y}_t$ and $\mathbf{y}_{t+1} = A_0 + A_1\mathbf{y}_t + \mathbf{u}_{t+1}$, the one-step ahead forecast error is equal to \mathbf{u}_{t+1} . Using the expression in (3.1), we could then express the two-step ahead forecast error as $(\mathbf{u}_{t+2} + A_1\mathbf{u}_{t+1})$. In this way we are able to calculate the H -step ahead forecast error as $(\mathbf{u}_{t+H} + A_1\mathbf{u}_{t+H-1} + A_1^2\mathbf{u}_{t+H-2} + \dots + A_1^{H-1}\mathbf{u}_{t+1})$.

After calculating the forecast errors of the *reduced-form* model, we could then derive appropriate values for the forecast errors in terms of the *structural-form* errors, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$. This would then allow for us to calculate the proportion of the expected variance in a variable that is due to each of the structural shocks in the model. For example, if $\varepsilon_{2,t}$ explains none of the forecast error variance of $y_{1,t}$; then $y_{1,t}$ is exogenous as it evolves independent of $\varepsilon_{2,t}$ and $y_{2,t}$. Similarly, if $\varepsilon_{2,t}$ explains all the forecast error variance of $y_{1,t}$; then $y_{1,t}$ is entirely endogenous.

As we would need to make use of identification restrictions to recover the structural shocks, we should be aware of the impact of these restrictions on the results. In this case the use of the Cholesky decomposition necessitates that all one period forecast errors of $y_{2,t}$ is due to $\varepsilon_{2,t}$. However lagged values of $\varepsilon_{1,t}$ could affect values of $y_{2,t}$. When using the alternate ordering of the Cholesky decomposition this would provide different results. For this reason it is often useful to examine the variance decompositions at different horizons, where as H increases the decompositions should converge.

Figure 2: **Forecast error variance decompositions**

An example of the results of a forecast error variance decomposition is provided in Figure 2. In this case the model includes data for unemployment and output, where we note that the variance in output is largely attributable to output shocks, while the variance in unemployment is attributable to both output and unemployment shocks.

4 Alternative restrictions for coefficient matrix

Sims (1986) and Bernanke (1986) provide examples of SVAR models that impose identification restrictions that may differ from the Cholesky decomposition (or alternate ordering of the Cholesky decomposition). The motivation for such a decomposition is that the upper or lower triangle may not allow for restrictions that are consistent with a particular theory that is to be modelled.

For example, in a three variable model, where $C = B^{-1}$ the Cholesky decomposition would impose restrictions such that the structural errors may be identified from the *reduced-form* errors, after calculating

$$\begin{aligned}
 u_{1,t} &= \varepsilon_{1,t} \\
 u_{2,t} &= c_{21}\varepsilon_{1,t} + \varepsilon_{2,t} \\
 u_{3,t} &= c_{31}\varepsilon_{1,t} + c_{32}\varepsilon_{2,t} + \varepsilon_{3,t}
 \end{aligned}$$

As an alternative, we could conceive that each of the structural shocks could be identified from a combination of two *reduced-form* errors. For example, consider the decomposition,

$$u_{1t} = \varepsilon_{1t} + c_{13}\varepsilon_{3t}$$

$$u_{2t} = c_{21}\varepsilon_{1t} + \varepsilon_{2t}$$

$$u_{3t} = c_{31}\varepsilon_{2t} + \varepsilon_{3t}$$

This structural decomposition clearly differs to that of the Cholesky decomposition, as we have lost the triangular structure of the identification restrictions, where each structural shock is affected by its own structural innovation and the structural innovation in one other variable. In this case the condition for $(K^2 - K)/2$ restrictions are satisfied, so the conditions for exact identification of the structural parameters in the model is maintained.

To consider how this would impact on the results of a model, consider a hypothetical example for a two variable model that has a sample size of five. This model would provide five residuals for $u_{1,t}$ and $u_{2,t}$, which may take on the values that are provided in Table 1.

	1	2	3	4	5
$u_{1,t}$	1.0	-0.5	0.0	-1.0	0.5
$u_{2,t}$	0.5	-1.0	0.0	-0.5	1.0

Table 1: **Reduced-form residuals**

Note that in this hypothetical example, both $u_{1,t}$ and $u_{2,t}$ sum to zero, while $\sigma_1 = 0.5$, $\sigma_{12} = \sigma_{21} = 0.4$, and $\sigma_2 = 0.5$. This provides the variance-covariance matrix,

$$\Sigma_u = \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 0.5 \end{bmatrix}$$

Since we had initially premultiplied ε_t by B^{-1} to derive u_t , we would be able to calculate values for Σ_ε from Σ_u with the use of

$$\Sigma_\varepsilon = B\Sigma_u B'$$

This expression could be expanded upon, where after inserting the above variance-covariance matrix,

$$\begin{bmatrix} \text{var}(\varepsilon_1) & 0 \\ 0 & \text{var}(\varepsilon_2) \end{bmatrix} = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 0.5 \end{bmatrix} \begin{bmatrix} 1 & b_{21} \\ b_{12} & 1 \end{bmatrix}$$

After completing the matrix multiplication we are left with the following four equations,

$$\text{var}(\varepsilon_1) = 0.5 + 0.8b_{12} + 0.5b_{12}^2 \quad (4.1)$$

$$0 = 0.5b_{21} + 0.4b_{21}b_{12} + 0.4 + 0.5b_{12} \quad (4.2)$$

$$0 = 0.5b_{21} + 0.4b_{21}b_{12} + 0.4 + 0.5b_{12} \quad (4.3)$$

$$\text{var}(\varepsilon_2) = 0.5b_{21}^2 + 0.8b_{21} + 0.5 \quad (4.4)$$

Since (4.2) and (4.3) are identical, we have three independent equations that need to solve for four unknowns.³ If we were to impose a Cholesky decomposition then we would need to set $b_{12} = 0$. This would ensure that we would now be left with three unknowns, which would allow us to solve (4.2) to (4.3) as follows,

$$\begin{aligned}
\text{var}(\varepsilon_1) &= 0.5 \\
0 &= 0.5b_{21} + 0.4 \quad \text{s.t.} \quad b_{21} = -0.8 \\
0 &= 0.5b_{21} + 0.4 \quad \text{s.t.} \quad b_{21} = -0.8 \\
\text{var}(\varepsilon_2) &= 0.5b_{21}^2 + 0.8b_{21} + 0.5 = 0.18
\end{aligned}$$

Since $\varepsilon_{1,t} = u_{1,t}$ and $\varepsilon_{2,t} = -0.8u_{1,t} + u_{2,t}$, we could then recover the structural shocks $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, which have been derived in Table 2.

	1	2	3	4	5
$\varepsilon_{1,t}$	1.0	-0.5	0.0	-1.0	0.5
$\varepsilon_{2,t}$	-0.3	-0.6	0.0	0.3	0.6

Table 2: **Structural-form residuals**

While in this example we made use of the identification restriction $b_{12} = 0$, different calibrated values could also be used. For example, if one of the shocks, $\varepsilon_{2,t}$ has a one-for-one affect on $y_{1,t}$ then we could set $b_{12} = 1$. This would allow us to solve for the structural shocks, such that,

$$\begin{aligned}
\text{var}(\varepsilon_1) &= 0.5 + 0.8b_{12} + 0.5b_{12}^2 = 1.8 \\
&\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots
\end{aligned}$$

where after solving for the unknowns, we would be able to recover values for the structural shocks, ε_t . Similarly, although there is little theory that informs us on the variance of shocks, we could elect to calibrate one of the values of the variance terms to $\text{var}(\varepsilon_1) = 1.8$, which would allow us to calculate values for b_{12} .

$$\begin{aligned}
\text{var}(\varepsilon_1) &= 1.8 = 0.5 + 0.8b_{12} + 0.5b_{12}^2 \\
&\vdots \qquad \vdots \qquad \vdots
\end{aligned}$$

which would also allow for the recovery of the structural shocks in ε_t . Alternatively, we could also impose the restriction $b_{12} = b_{21}$. Then replacing b_{21} with b_{12} in the following expression would allow for us to calculate

$$\begin{aligned}
0 &= 0.5b_{21} + 0.4b_{21}b_{12} + 0.4 + 0.5b_{12} \\
&\vdots \qquad \qquad \qquad \vdots
\end{aligned}$$

where after deriving values for b_{12} , we can then solve for the rest.

5 Long-run restrictions

Economic theory does not always provide us with enough meaningful restrictions that could be imposed on the contemporaneous coefficients. In these circumstances, we could impose restrictions on the long-run properties of shocks. For example, we could impose conditions that allow for the neutrality of the effects of certain shocks over time.

Such restrictions were imposed in Blanchard and Quah (1989), which makes use of a bivariate structure to consider the relationship between output growth and unemployment; where output growth is responsible for the demand-side shocks and unemployment is responsible for the supply-

side shocks. Given the fact that there are only two variables in this model, we would need to impose a single restriction, to allow for output growth and unemployment to be determined by orthogonal structural shocks.

In this case, it is assumed that output growth is affected by both demand and supply shocks. However, in accordance with the natural rate hypothesis, it is assumed that demand-side shocks have no long-run effects on output growth, while supply-side productivity shocks have a lasting effect on output growth.

If the logarithm of output, $y_{1,t}$, is $I(1)$ then output growth, $\Delta y_{1,t}$, is $I(0)$. In addition, if we are to assume that the rate of unemployment, $y_{2,t}$, is $I(0)$, then we are able to write the bivariate moving average representation as,

$$\mathbf{y}_t = \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i}$$

where \mathbf{y}_t is a vector of both variables. This expression could be expanded as,

$$\begin{aligned} \begin{bmatrix} \Delta y_{1,t} \\ y_{2,t} \end{bmatrix} &= \begin{bmatrix} \theta_{11}(0) & \theta_{12}(0) \\ \theta_{21}(0) & \theta_{22}(0) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} + \dots \\ &\quad \begin{bmatrix} \theta_{11}(1) & \theta_{12}(1) \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix} + \dots \end{aligned} \quad (5.1)$$

where the lags have been included within the brackets (\cdot) . Therefore, the effect of $\varepsilon_{1,t-1}$ on $\Delta y_{1,t}$ is summarized by $\theta_{11}(1)$.

Now, if $\varepsilon_{1,t}$ has no long-run cumulative impact on $\Delta y_{1,t}$ we could impose the restriction

$$\sum_{i=0}^{\infty} \theta_{11}(i) = 0$$

which may be included in the coefficient matrix for the moving average representation,

$$\sum_{i=0}^{\infty} \Theta_i = \begin{bmatrix} 0 & \sum_{i=0}^{\infty} \theta_{12,i} \\ \sum_{i=0}^{\infty} \theta_{21,i} & \sum_{i=0}^{\infty} \theta_{22,i} \end{bmatrix} = \sum_{i=0}^{\infty} \begin{bmatrix} 0 & \theta_{12}(i) \\ \theta_{21}(i) & \theta_{22}(i) \end{bmatrix}$$

Hence, we can impose restrictions on either the short-run contemporaneous parameters, or the long-run moving average components to satisfy the identification restrictions in the model. Alternatively we could also use a combination of short-run and long-run restrictions. The only condition that needs to be enforced is that the number of restrictions must equal $[(K^2 - K)/2]$.

6 Conclusion

Sims (1980) introduced the structural vector autoregressive technique as an alternative to the large-scale macroeconomic models that were used during the 1970s. This methodology has gained widespread use in applied time series research, as it allows for the incorporation of contemporaneous variables and an investigation into the impact of individual shocks.

To identify the structural VAR model, we need to impose certain restrictions on the parameters in the model. Widely-used identification schemes could rely on either short-run or long-run restrictions. The short-run restrictions were originally suggested by Sims (1980), while Blanchard and Quah (1989) introduced a method for imposing long-run restrictions.

When seeking to model a system of K variables, we would need to impose $(K^2 - K)/2$ restrictions for exact identification. The use of the Cholesky decomposition would ensure that the identified shocks from the VAR model will be orthogonal (uncorrelated) and unique. However, the choice of this method for imposing restrictions could affect the results of the model. Notable alternative methods for imposing identification restrictions have been considered in Sims (1986) and Bernanke (1986).

We can investigate the effects of a shock on the VAR model, with the aid of an impulse response function, which describes how a given structural shock affects a variable over time. In addition, the forecast error variance decomposition attributes the forecast error variance to specific structural shocks at different horizons.

Before concluding, it is worth noting that a major limitation of the traditional VAR approach is that it is highly parameterised, which would imply that we could encounter degrees of freedom problems when including several variables or many lags. In addition, as all of the effects of omitted variables will be contained in the residuals, this may lead to major distortions in the impulse responses and large measurement errors (or a potential mis-specification). Imposing some structure on the model (by calibrating the values of certain parameters) may alleviate some of these potential problems.

7 References

Bernanke, Ben S. 1986. "Alternative Explanations of the Money-Income Correlation." *Carnegie-Rochester Conference Series on Public Policy* 25: 49–99.

Blanchard, Olivier J., and Danny Quah. 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review* 79 (4): 655–73.

Sims, Christopher A. 1980. "Comparison of Interwar and Postwar Business Cycles." *American Economic Review* 70 (2): 250–57.

———. 1986. "Are Forecasting Models Usable for Policy Analysis?" *Federal Reserve Bank of Minneapolis Quarterly Review* 10: 2–16.

-
1. For this reason Christopher A. Sims (together with Thomas J. Sargent) received the Sveriges Riksbank Price in Economic Sciences in Memory of Alfred Nobel (Nobel Prize in Economics) in 2011.↵
 2. The structural parameters would include, $b_{10}, b_{20}, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}, b_{12}, b_{21}, \sigma_1, \sigma_2$, while the reduced-form parameters would include, $a_{10}, a_{20}, a_{11}, a_{12}, a_{21}, a_{22}, \text{var}(u_{1,t}), \text{var}(u_{2,t}), \text{cov}(u_{1,t}, u_{2,t})$.↵
 3. The unknowns would include $\text{var}(\varepsilon_1), \text{var}(\varepsilon_2), b_{12}$ and b_{21} .↵