Bayesian Methods for Macroeconometrics

Frank Schorfheide
Department of Economics, University of Pennsylvania

(Lack of) Identification

- So far, we considered reduced form VARs, say,

$$y_t = \Phi_1 y_{t-1} + u_t, \quad \mathbb{E}[u_t u_t'] = \Sigma \tag{1}$$

- Error terms u_t have the interpretation of one-step ahead forecast errors.

- If the eigenvalues of Φ_1 are inside the unit-circle then y_t has the following moving-average (MA) representation in terms of u_t:

$$y_t = (I - \Phi_1 L)^{-1} u_t = \sum_{j=0}^{\infty} \Phi_1^j u_{t-j} = \sum_{j=0}^{\infty} C_j u_{t-j} \tag{2}$$

- DSGE models suggest that the one-step ahead forecast errors are functions of some fundamental shocks, such as technology shocks, preference shocks, or monetary policy shocks.

(Lack of) Identification

- Let ϵ_t a vector of such fundamental shocks and assume that $I\!E[\epsilon_t \epsilon_t'] = \mathcal{I}$. Moreover, assume that

$$u_t = \Phi_\epsilon \epsilon_t. \tag{3}$$

- Then we can express the VAR in structural form as follows

$$\begin{aligned} y_t &= \Phi_1 y_{t-1} + \Phi_\epsilon \epsilon_t \qquad (4) \\ \Phi_\epsilon^{-1} y_t &= \Phi_\epsilon^{-1} \Phi_1 y_{t-1} + \epsilon_t \end{aligned}$$

- The moving-average representation of y_t in terms of the structural shocks is given by

$$y_t = \sum_{j=0}^{\infty} \Phi_1^j \Phi_\epsilon \epsilon_{t-j} = \sum_{j=0}^{\infty} C_j \Phi_\epsilon \epsilon_{t-j}. \tag{5}$$

(Lack of) Identification

- For (1) and (4) the matrix Φ_ϵ has to satisfy the restriction

$$\Phi_\epsilon \Phi_\epsilon' = \Sigma \tag{6}$$

  Notice that the matrix Φ_ϵ has n^2 elements.

- The covariance relationship, unfortunately, generates only $n(n+1)/2$ restrictions and does not uniquely determine Φ_ϵ.

- This creates an identification problem since all we can estimate from the data is Φ_1 and Σ.

(Lack of) Identification

- In order to make statements about the propagation of structural shocks ϵ_t we have to make further assumptions. The papers (see course outline) by Cochrane (1994), Christiano and Eichenbaum (1999), and Stock and Watson (2001) survey such identifying assumptions. A cynical view of this literature is the following:

  1. Propose an identification scheme, that determines all elements of Φ_ϵ.
  2. Compute impulse response functions.
  3. If impulse response functions are plausible, then stop; else, declare a "puzzle" and return to 1.

(Lack of) Identification

- Here are some famous "puzzles:"

  1. "Liquidity Puzzle:" When identifying monetary policy shocks as surprise changes in the stock of money one often finds that interest rates fall when the money stock is lowered.

  2. "Price Puzzle:" When identifying monetary policy shocks as surprise changes in the Federal Funds Rate, one often finds that prices fall after a drop in interest rates.

- These "puzzles" are typically resolved by considering more elaborate identification schemes.

Identification Schemes

- We begin by decomposing the covariance matrix into the product of lower triangular matrices (Cholesky Decomposition):

$$\Sigma_u = AA', \tag{7}$$

where A is lower triangular. If Σ_u is non-singular the decomposition is unique.

- Let Ω be an orthonormal matrix, meaning that $\Omega\Omega' = \Omega'\Omega = \mathcal{I}$.

- We can characterize the relationship between the reduced form and the structural shocks as follows

$$u_t = A\Omega\epsilon_t \tag{8}$$

- Notice that

$$I\!\!E[u_t u_t'] = I\!\!E[A\Omega\epsilon_t\epsilon_t'\Omega' A'] = A\Omega I\!\!E[\epsilon_t\epsilon_t']\Omega' A' = A\Omega\Omega' A' = AA' = \Sigma_u. \tag{9}$$

Identification Schemes

- In general, it is quite tedious to characterize the space of orthonormal matrices. Let's try for $n = 2$:

$$\Omega(\varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \tag{10}$$

  where $\varphi \in (-\pi, \pi]$.

- Notice that, for instance,

$$\Omega(\pi/2) = -\Omega(-\pi/2) \tag{11}$$

  which means that only the signs of the impulse responses change but not the shape.

- Identification schemes impose restrictions on φ.

Short-run Restrictions

- Suppose that

$$y_t = \begin{bmatrix} \text{Fed Funds Rate} \\ \text{Output Growth} \end{bmatrix}, \quad \epsilon_t = \begin{bmatrix} \epsilon_{R,t} \\ \epsilon_{z,t} \end{bmatrix} = \begin{bmatrix} \text{Monetary Policy Shock} \\ \text{Technology Shock} \end{bmatrix}.$$

Moreover, we assume that the central bank does not react contemporaneously to technology shocks because data on aggregate output only become available with a one-quarter lag.

This assumption can be formalized through $\varphi = 0$. Then

$$u_t = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \epsilon_{R,t} \\ \epsilon_{z,t} \end{bmatrix}. \tag{12}$$

Reference: Sims (1980).

Long-run Restrictions

- Now suppose that

$$y_t = \begin{bmatrix} \text{Inflation} \\ \text{Output Growth} \end{bmatrix}, \quad \epsilon_t = \begin{bmatrix} \epsilon_{R,t} \\ \epsilon_{z,t} \end{bmatrix} = \begin{bmatrix} \text{Monetary Policy Shock} \\ \text{Technology Shock} \end{bmatrix}$$

Moreover,

$$y_t = (\sum_{j=0}^{\infty} C_j L^j) u_t = C(L) u_t. \tag{13}$$

Consider the following assumption: monetary policy shocks do not raise output in the long-run.

Long-run Restrictions

- Let's examine the moving average representation of y_t in terms of the structural shocks

$$
\begin{aligned}
y_t &= \begin{bmatrix} c_{11}(L) & c_{12}(L) \\ c_{21}(L) & c_{22}(L) \end{bmatrix} \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} \epsilon_{R,t} \\ \epsilon_{z,t} \end{bmatrix} \\
&= \begin{bmatrix} \cdot & \cdot \\ a_{11}\cos\varphi c_{21}(L) + (a_{21}\cos\varphi + a_{22}\sin\varphi)c_{22}(L) & \cdot \end{bmatrix} \begin{bmatrix} \epsilon_{R,t} \\ \epsilon_{z,t} \end{bmatrix} \\
&= \begin{bmatrix} d_{11}(L) & d_{12}(L) \\ d_{21}(L) & d_{22}(L) \end{bmatrix} \begin{bmatrix} \epsilon_{R,t} \\ \epsilon_{z,t} \end{bmatrix}
\end{aligned}
$$

- Suppose that in period $t = 0$ log output and log prices are equal to zero. Then the log-level of output and prices in period $t = T > 0$ is given by

$$
y_T^c = \sum_{t=1}^{T} y_t = \sum_{t=1}^{T} \sum_{j=0}^{\infty} D_j \epsilon_{t-j} \tag{14}
$$

Long-run Restrictions

- Now consider the derivative

$$\frac{\partial y_T^c}{\partial \epsilon_1'} = \sum_{j=0}^{T-1} D_j \tag{15}$$

- Letting $T \longrightarrow \infty$ gives us the long-run response of the level of prices and output to the shock ϵ_1:

$$\frac{\partial y_\infty^c}{\partial \epsilon_1'} = \sum_{j=0}^{\infty} D_j = D(1) \tag{16}$$

- Here, we want to restrict the long-run effect of monetary policy shocks on output:

$$d_{21}(1) = 0 \tag{17}$$

Long-run Restrictions

- This leads us to the equation

$$[a_{11}c_{21}(1) + a_{21}c_{22}(1)] \cos \varphi + a_{22}c_{22}(1) \sin \varphi = 0. \tag{18}$$

- Notice that the equation has two solutions for $\varphi \in (-\pi, \pi]$. Under one solution a positive monetary policy shock is contractionary, under the other solution it is expansionary. The shape of the responses is, of course, the same.

- Reference: Blanchard and Quah (1989).

Sign Restrictions

- Again consider

$$y_t = \begin{bmatrix} \text{Inflation} \\ \text{Output Growth} \end{bmatrix}, \quad \epsilon_t = \begin{bmatrix} \epsilon_{R,t} \\ \epsilon_{z,t} \end{bmatrix} = \begin{bmatrix} \text{Monetary Policy Shock} \\ \text{Technology Shock} \end{bmatrix}$$

- Our identification assumption is: upon impact, a monetary policy shock raises both prices and output. It can be verified that

$$\frac{\partial y_t}{\partial \epsilon_{R,t}} = \begin{bmatrix} a_{11} \cos \varphi c_{11,1} + (a_{21} \cos \varphi + a_{22} \sin \varphi) c_{12,1} \\ a_{11} \cos \varphi c_{21,1} + (a_{21} \cos \varphi + a_{22} \sin \varphi) c_{22,1} \end{bmatrix}. \tag{19}$$

Sign Restrictions

- Thus, we obtain the sign restrictions

$$
\begin{aligned}
0 &< a_{11} \cos\varphi c_{11,1} + (a_{21}\cos\varphi + a_{22}\sin\varphi)c_{12,1} \\
0 &< a_{11} \cos\varphi c_{21,1} + (a_{21}\cos\varphi + a_{22}\sin\varphi)c_{22,1}
\end{aligned}
$$

  which restrict φ to be in a certain subset of $(-\pi, \pi]$ and will generate a range of responses.

- References: Canova and De Nicolo (2002), Faust (1998), Uhlig (2005).

Impulse Responses and Variance Decompositions

- Impulse responses are defined as

$$\frac{\partial y_{t+h}}{\partial \epsilon_t'} = C_h \Phi_\epsilon \tag{20}$$

and correspond to the MA coefficient matrices in the moving average representation of y_t in terms of structural shocks.

Impulse Responses and Variance Decompositions

- The covariance matrix of y_t is given by

$$\Gamma_{yy,0} = \sum_{j=0}^{\infty} C_j \Phi_\epsilon \mathcal{I} \Phi_\epsilon' C_j' \tag{21}$$

Let $\mathcal{I}^i$ be matrix for which element i, i is equal to one and all other elements are equal to zero. Then we can define the contribution of the i'th structural shock to the variance of y_t as

$$\Gamma_{yy,0}^{(i)} = \sum_{j=0}^{\infty} C_j \Phi_\epsilon \mathcal{I}^{(i)} \Phi_\epsilon' C_j' \tag{22}$$

Thus the fraction of the variance of $y_{l,t}$ explained by shock i is

$$\frac{[\Gamma_{yy,0}^{(i)}]_{ll}}{[\Gamma_{yy,0}]_{ll}}. \tag{23}$$

Implementation

- Consider a simple VAR of the form $y_t = \Phi_1 y_{t-1} + u_t$, $u_t = A\Omega(\varphi)\epsilon_t$, $\Phi = \Phi_1'$. For $s = 1, \dots, n_{sim}$:

  1. Generate a draw from the posterior distribution of (Φ, Σ), e.g., using sampling techniques for the $\mathcal{IW} - \mathcal{N}$ distribution. Let $A = chol(\Sigma)$.
  2. Compute moving average representation $y_t = \sum_{j=0} C_j(\Phi) u_t$.
  3. Short-run and long-run identification schemes: determine φ as function of A and the $C_j(\Phi)$'s.

     Sign Restrictions: conditional on Φ and A assign a prior distribution to the set of φ's for which the sign restrictions are satisfied. Generate a draw φ from this prior. Note: the sample has no information about φ given Φ, A. Hence prior equals posterior.

4. Once φ is determined, compute impulse responses and variance decompositions.

- This algorithm leaves you with n_{sim} draws from the posterior of the impulse responses and variance decompositions. You can now compute summary statistics for this posterior, such as means, medians, standard deviations, and (pointwise) confidence sets.

Alternative Setups

- Sims and Zha (1998) and subsequent work start out from the specification

$$A_0 y_t = A_1 y_{t-1} + \dots A_p y_{t-p} + \epsilon_t$$

  where ϵ_t's are structural shocks.

- Analyze the model directly in terms of A matrices.

- Impose identification restrictions on A_0.

References

- Blanchard, Oliver Jean, and Danny Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, **79**, 655-673.
- Canova, F. and G. De Nicoló, "Monetary Disturbances Matter for Business Cycle Fluctuations in the G-7." *Journal of Monetary Economics* **49**(6) (2002), 1131-1159.
- Cochrane, John H. (1994): "Shocks," *Carnegie-Rochester Conference Series on Public Policy*, **41**, 295-364.
- Doan, Thomas, Robert Litterman, and Christopher Sims (1984): "Forecasting and Conditional Projections Using Realistic Prior Distributions," *Econometric Reviews*, **3**, 1-100.
- Faust, Jon (1998): "The Robustness of Identified VAR Conclusions about Money,"

Carnegie Rochester Conference Series, **49**, 207-244.

- Sims, Christopher, and Tao Zha (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, **39**, 949-968.

- Stock, James J. and Mark W. Watson (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, **15**, 101-115.

- Uhlig, Harald (2005) "What are the Effects of Monetary Policy? Evidence from an Agnostic Identification Procedure," *Journal of Monetary Economics*.