

1. Shocks

This version: February 15, 2006

Nr. 1

1.3. Factor models

- What if there are more shocks than variables in the VAR?
- What if there are only a few underlying shocks, explaining most of fluctuations?

In fact, the same question... Can think of most macroeconomic variables as being driven by a few common shocks, plus partly idiosyncratic shocks/measurement error:

Think of n variables depending on $q < n$ common shocks, and n idiosyncratic shocks. Then, we have both:

- More shocks ($q + n$) than variables, so standard small VAR wrong.
- Less “important” shocks ($q < n$) than variables.

1. Standard dynamic factor model

- Using Stock-Watson 2005 notation.
- n variables X_{it} , $i = 1, \dots, n$. (n may be large)
- q underlying factors (common shocks) f_{1t}, \dots, f_{qt} , $f_t \equiv [f_{1t} \dots f_{qt}]'$
- Each variable depends on a distributed lag of each factor, and an idiosyncratic component:

$$X_{it} = \tilde{\lambda}_i(L) f_t + u_{it}$$

where $\tilde{\lambda}_i(L) = [\tilde{\lambda}_{i1}(L) \dots \tilde{\lambda}_{iq}(L)]$ is a vector of lag polynomials.

- Idiosyncratic components are serially correlated, but mutually uncorrelated:

$$u_{it} = \delta_i(L)u_{it-1} + v_{it}$$

where innovations are mutually and serially uncorrelated, $E v_{it} v_{j\tau} = 0$, for all $t, \tau, i, j, i \neq j$

Last assumption unpalatable (X_{it} : consumption of non-durables, X_{jt} : consumption of services). Can be somewhat relaxed (and relaxation ok if n is indeed very large).

- Factors follow a multivariate autoregressive process:

$$f_t = \Gamma(L)f_{t-1} + \eta_t$$

where innovations can be mutually correlated, but are serially uncorrelated and uncorrelated with idiosyncratic components at all leads and lags; $E \eta_t u_{i\tau} = 0$ for all i, t, τ

Example 1. The simplest common factor model

$$X_{it} = \tilde{\lambda}_i f_t + v_{it}$$

where f_t is a single factor; no lags.

Can see how, if n is large, easy to recover to recover the single factor (Choice of variables, of units: not neutral)

Example 2. A simple dynamic single common factor model

$$X_{it} = \tilde{\lambda}_{i1} f_t + \tilde{\lambda}_{i2} f_{t-1} + v_{it}$$

For example, different components of consumption or investment respond differently to technological shocks.

Example 3. A simple structural factor model

$$X_{it} = \tilde{\lambda}_{i1}(L)f_{1t} + \tilde{\lambda}_{i2}(L)f_{2t} + v_{it}$$

where, for example, the first factor is a technological shock, and the second factor is a “demand ” shock.

Assume that the first variable in the VAR is the rate of change of GDP and that the “demand” shock has no effect on output in the long run. This assumption implies

$$\tilde{\lambda}_{12}(1) = 0$$

Estimating the dynamic factor model

Rewrite the equation for X_{it} so as to have white noise errors. From:

$$X_{it} = \tilde{\lambda}_i(L) f_t + u_{it}$$

and

$$u_{it} = \delta_i(L) u_{it-1} + v_{it}$$

Get:

$$X_{it} = \lambda_i(L) f_t + \delta_i(L) X_{it-1} + v_{it}$$

where $\lambda_i(L) \equiv (1 - \delta_i(L)L) \tilde{\lambda}_i(L)$.

Rewrite

$$X_{it} = \lambda_i(L) f_t + \delta_i(L) X_{it-1} + v_{it}$$

in matrix form:

$$X_t = \lambda(L) f_t + D(L) X_{t-1} + v_t$$

where $X_t \equiv [X_{1t} \dots X_{nt}]'$, $\lambda(L) \equiv [\lambda_1(L)' \dots \lambda_n(L)']$, $D(L)$ is a diagonal matrix, and

$$f_t = \Gamma(L) f_{t-1} + \eta_t$$

Note that $\lambda(L)$ and f_t are not separately identified:

$$\lambda(L) f_t = (\lambda(L) H^{-1})(H f_t)$$

so need some normalization.

Define $F_t \equiv [f'_t \ f'_{t-1} \ \dots \ f'_{t-p+1}]$, so we can rewrite the DFM as:

$$X_t = \Lambda F_t + D(L)X_{t-1} + v_t \quad (1)$$

$$F_t = \Phi(L)F_{t-1} + G\eta_t \quad (2)$$

The dimension of F_t is (at most) $r = qp$. The r elements of F_t are called “static factors”. The underlying q elements of f_t are called “dynamic factors”.

We ultimately want to recover the dynamic factors, and the effects of their innovations, the η 's on the X_i .

Can be estimated using a [Kalman filter](#) (Hamilton, Chapter 13):

First equation is the “observation equation”, second equation is “state equation”. But hard to do for large n .

So other approach: [Principal components](#).

An approach to estimation (for more, see Stock-Watson 2-4).

- First step: Assume some $D_0(L)$ in equation (1) and construct $\tilde{X}_t \equiv X_t - D_0(L)X_{t-1}$. Derive F_t as the first r principal components of \tilde{X}_t . (How to choose r ?)
- Regress X_{it} on estimated F_t and lagged X_{it} . Get $D_1(L)$. Iterate until convergence. This gives us estimates for the r **static** factors F_t .
- Second step: Regress (estimated) F_t on F_{t-1} and choose the first q principal components of the residuals. This gives us the innovations to the dynamic factors. (How to choose q ?)

(If we had F_t exactly, the covariance matrix should be singular.)

Back to the second example.

$$X_{it} = \lambda_{i1}f_t + \lambda_{i2}f_{t-1} + v_{it}$$

Assume that f_t is white noise:

$$f_t = \eta_t, E\eta_t = 0, V(\eta_t) = 1$$

So:

$$X_{it} = \lambda_{i1}\eta_t + \lambda_{i2}\eta_{t-1} + v_{it}$$

and defining $F_t \equiv [\eta_t \ \eta_{t-1}]'$

$$X_{it} = [\lambda_{i1} \ \lambda_{i2}] F_t + v_{it}$$

This is a one dynamic factor, two static factors, DFM. Can we recover the two static factors, and the one dynamic factor?

- Step 1: Start with $D_0(L) = 0$, do principal components, and iterate.

Get $D(L) = 0$ and two static factors, g_{1t} , g_{2t} . Both are white noise, mutually uncorrelated. Relation to the two static factors, f_{1t} and f_{2t} (η_t and η_{t-1})?

Let $G_t \equiv [g_{1t} \ g_{2t}]'$. Then,

$$G_t = H F_t$$

for H is an orthogonal matrix, so $V(F) = V(G) = I$. In this 2x2 case, H is determined up to a scalar θ . The relation between the true shocks and the estimated factors can be expressed as:

$$\eta_t = h_{11} g_{1t} + h_{21} g_{2t}$$

$$\eta_{t-1} = h_{12} g_{1t} + h_{22} g_{2t}$$

- Step 2. Regress g_{1t} , g_{2t} on g_{1t-1} , g_{2t-1} .

Recall that there is a linear combination of g_{1t} and g_{2t} which is equal to η_{t-1} , and a linear combination of g_{1t-1} and g_{2t-1} which is also equal to η_{t-1} . So the rank of the matrix of residuals will be one, the number of dynamic factors.

Doing principal components on the error term (the term in brackets) recovers η_t , the single dynamic factor.

In this case, identification is easy. But in general, need further assumptions to go back from dynamic factors to underlying shocks. (Same as from VAR to structural VAR)

Steps behind the previous slide:

$$\begin{bmatrix} f_{1t} \\ f_{2t} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f_{1t-1} \\ f_{2t-1} \end{bmatrix} + \begin{bmatrix} \eta \\ 0 \end{bmatrix}$$

Using $F_t = H' G_t$:

$$H' \begin{bmatrix} g_{1t} \\ g_{2t} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} H' \begin{bmatrix} g_{1t-1} \\ g_{2t-1} \end{bmatrix} + \begin{bmatrix} \eta \\ 0 \end{bmatrix}$$

Premultiplying by H and using $HH' = I$:

$$\begin{bmatrix} g_{1t} \\ g_{2t} \end{bmatrix} = H \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} H' \begin{bmatrix} g_{1t-1} \\ g_{2t-1} \end{bmatrix} + H \begin{bmatrix} \eta \\ 0 \end{bmatrix}$$

Note that the residuals in the last equation are both linear in η_t , so the matrix of residuals has rank one.

Identification

- Write the factor model in infinite MA form. From above,

$$X_t = (I - D(L)L)^{-1} \lambda(L)(1 - \Gamma(L)L)^{-1} \eta_t + (I - D(L)L)^{-1} v_t$$

Each X_i depends on both innovations to the dynamic factors, and current and past own idiosyncratic shocks.

Defining matrix lag polynomials C and J appropriately:

$$X_t = C_0 \eta_t + C(L) \eta_{t-1} + J_0 v_t + J(L) v_{t-1}$$

- Using the same notation as in topic 1-2, assume the true model is given by:

$$X_t = A_0 e_t + A(L) e_{t-1} + K_0 w_t + K(L) w_{t-1}$$

where the e_t and w_t are the underlying structural common and idiosyncratic shocks. Without loss of generality, assume $\Sigma_e = I$.

- Comparison of the factor and the true model imply:

$$C_0 \eta_t = A_0 e_t$$

Let Σ_η be the covariance matrix of the innovations to the common factors. Then A_0 must satisfy:

$$V(A_0 e_t) = A_0 A_0' = V(C_0 \eta_t) = C_0 \Sigma_\eta C_0'$$

General treatment of identification left to Stock Watson, sections 3-2 (short-run restrictions) and 3-3 (long-run restrictions).

Example 1. Two common structural shocks, with innovations e_{1t} and e_{2t} .

Define $x_{it}^c \equiv c_{0i1} \eta_{1t} + c_{0i2} \eta_{2t}$ be the innovation in X_{it} associated with common factors, so the relation of the x^c to the two common shocks is given by:

$$x_{1t}^c = a_{011} e_{1t} + a_{012} e_{2t}$$

$$x_{2t}^c = a_{021} e_{1t} + a_{022} e_{2t}$$

...

$$x_{nt}^c = a_{0n1} e_{1t} + a_{0n2} e_{2t}$$

- Clearly not identified (consider orthogonal transformations of e , $z = He$).
- Zero restriction: Assume that the innovation to the second structural shock does not affect x_1^c contemporaneously, so $a_{012} = 0$. Then, identified: $a_{011} = \sigma_{x_1^c}$, and $e_{1t} = x_{1t}^c / \sigma_{x_1^c}$.
- Estimate all the other equations by OLS.

Example 2. q common structural shocks

Assume that the relation between the first q x^c s and the innovations to the structural shocks are given by:

$$\begin{bmatrix} x_{1t}^c \\ x_{2t}^c \\ \dots \\ x_{qt}^c \end{bmatrix} = \begin{bmatrix} a_{011} & 0 & 0 & \dots & 0 \\ a_{021} & a_{022} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ a_{0q1} & a_{0q2} & a_{0q3} & \dots & a_{0qq} \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \\ \dots \\ e_{qt} \end{bmatrix}$$

and there are no restrictions on the relation of the remaining x_{it} , $i = q+1, \dots, n$ to the structural shocks.

Then $A_0 = S$, where S is the lower triangular Choleski matrix associated with the covariance matrix of the first q x^c s.

Logic carries to more general zero restrictions, and long-run restrictions.

Turning to results. Three applications.

Application 1. Stock-Watson 2005. Non structural factor model.

- X_{it} : 132 macro time-series, log, and log differenced if needed. Monthly, 1959:1 to 2003:12.
- $D(L)$: 4/6 lags. Find 7-9 static factors. Find 7 dynamic factors.

Major conclusions? Not sure... (Problem: from innovations and reduced form shocks to a structural interpretation)

- A first image: Variance decomposition at different horizons. Table 2 summary.

Importance of idiosyncratic shocks.

Importance of the first shock, but also of others.

Partly misleading: On to Table 2b: variable by variable

- **First factor** dominates for quantities.

IP: 93% of two-year ahead forecast error. First factor very close to IP

High for nearly all components (70 to 90%, except for non-durables, 44%), and for unemployment (85%).

Does not explain prices at all (mostly explained by idiosyncratic components, and the third factor)

Explains some of interest rates, not asset prices.

- **Second factor** important for interest rates and asset prices.

Does not explain quantities; does explain some of the movement in prices.

Image removed due to copyright restrictions.

Table 2. Forecast Error Variance Decomposition with Respect to Factor Innovations. pp. 54-57.

Stock, J. and M. Watson, "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper No. 11467, July 2005. pp. 1-65.
Please see: <http://www.nber.org/papers/W11467>

Image removed due to copyright restrictions.

Table 2. Forecast Error Variance Decomposition with Respect to Factor Innovations. pp. 54-57.

Stock, J. and M. Watson, "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper No. 11467, July 2005. pp. 1-65.

Please see: <http://www.nber.org/papers/W11467>

Image removed due to copyright restrictions.

Table 2. Forecast Error Variance Decomposition with Respect to Factor Innovations. pp. 54-57.

Stock, J. and M. Watson, "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper No. 11467, July 2005. pp. 1-65.
Please see: <http://www.nber.org/papers/W11467>

Image removed due to copyright restrictions.

Table 2. Forecast Error Variance Decomposition with Respect to Factor Innovations. pp. 54-57.

Stock, J. and M. Watson, "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper No. 11467, July 2005. pp. 1-65.

Please see: <http://www.nber.org/papers/W11467>

Image removed due to copyright restrictions.

Table 2. Forecast Error Variance Decomposition with Respect to Factor Innovations. pp. 54-57.

Stock, J. and M. Watson, "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper No. 11467, July 2005. pp. 1-65.
Please see: <http://www.nber.org/papers/W11467>

Image removed due to copyright restrictions.

Table 2. Forecast Error Variance Decomposition with Respect to Factor Innovations. pp. 54-57.

Stock, J. and M. Watson, "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper No. 11467, July 2005. pp. 1-65.

Please see: <http://www.nber.org/papers/W11467>

Structural implications? Not obvious for the usual reasons: reduced form, not structural. Some leads

- First factor reflects the (important) fact that most quantities (investment, consumption), move together. Proof that there is a common factor? Not necessarily. Strong interactions between components. For example, simple *IS* model, with two shocks:

$$y = c + x, \quad c = .4y + \epsilon_c, \quad x = .4y + \epsilon_x, \quad \sigma_c = \sigma_x = 1$$

Then first factor will be output: $y = 5(\epsilon_c + \epsilon_x)$ And a regression of c on y (with OLS bias) gives:

$$c = .5y + \eta, \quad \eta = (\epsilon_c - \epsilon_x)$$

Then, proportion of variance of consumption explained by output ($V(.5y)/V(c) = 75\%$) Need some metric to assess SW results.

- Fact that first factor does not explain prices?

Supply/demand

$$y = -ap + \epsilon_d, \quad y = bp + \epsilon_s$$

So:

$$p = (a + b)^{-1}(\epsilon_d - \epsilon_s), \quad y = (a + b)^{-1}(b\epsilon_d + a\epsilon_s)$$

Suppose first factor explains quantities. Correlation with second factor?

$$E(py) = (a + b)^{-2}(-a\sigma_s^2 + b\sigma_d^2)$$

So: Similar variances and parameters? or large variance of supply shocks, and small a , or ?

- First, second factors and asset prices? The need for more information: leads and lags. And more a-priori structure.

Application 2. Bernanke-Boivin-Eliasz 2005. Monetary policy shocks.

Slightly different in structure and intent: Identify the effects of one shock, to monetary policy, when Fed looks at a lot more than just output and inflation.

(One) motivation: The “price puzzle”. CPI appears to increase in response to an “exogenous” increase in the FF rate.

Potential explanation. Fed sees an increase in the price of materials, and reacts in advance. Just looking at output and inflation, we do not see what the Fed reacted to.

Way out? DFM

Consider following economy (backward looking NK)

$$AS \quad \pi_t = \delta\pi_{t-1} + k(y_{t-1} - y_{t-1}^n) + s_t$$

$$AD \quad y_t = \phi y_{t-1} - \psi(R_{t-1} - \pi_{t-1}) + d_t$$

$$y_t^n = \rho y_{t-1}^n + \eta_t$$

$$s_t = \alpha s_{t-1} + v_t$$

$$TR \quad R_t = \beta\pi_t + \gamma(y_t - y_t^n) + \epsilon_t$$

If all LHS variables observable, then standard VAR, with R_t ordered last, gives dynamic effects of ϵ_t .

What if y_t^n, s_t not observable to the econometrician and run a 3-variable VAR?

What if y_t and π_t observed by the cb and the econometrician with noise?

With this motivation, consider the model:

$$\begin{bmatrix} F_t \\ R_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ R_{t-1} \end{bmatrix} + \begin{bmatrix} v_{Ft} \\ v_{Rt} \end{bmatrix} \quad V(v) = Q \quad (3)$$

where F_t are k unobservable factors, and R_t is the policy variable, an observable factor. (In the example above, $F_t = [\pi_t \ y_t \ y_t^n \ s_t]'$).

Could instead have some of the variables (π_t, y_t) for example, be observable factors.

Also observe m (a large number of other) macroeconomic variables X_t , which satisfy:

$$X_t = \Lambda^F F_t + \Lambda^R R_t + e_t \quad (4)$$

Relation to DFM earlier: Some factors observable. Only static factors (no direct effect of lagged R on X except through effect through other factors.)

Estimation/Identification

- 2-step.

Do principal components on (4). Treating all factors as unobservable (?). Gives the right space, not the individual factors.

- Decompose space between F_t and R_t . Run equation (3) with identification restriction. Factors not affected contemporaneously by money shocks. Recursive ordering, with R last.
- Maximum likelihood of (4) and (3)? Hard if m is very large. Likelihood-based Gibbs sampling. (Victor will explain).

Results

- 120 monthly time series, transformed to be stationary. 1959-1 to 2001-8

- Figure 1. Impulse responses.

VAR ($Y = IP, CPI, FFR, k=0$)

1-factor FAVAR ($Y = IP, CPI, FFR, k=1$)

3-factor-FAVAR ($Y = FFR, k = 3.$)

- Reduces the “price puzzle”
- Issues.

Image removed due to copyright restrictions.

Figure 1. Estimated Impulse Response to an Identified Policy Shock for Alternative FAVAR Specifications, Based on the Two-Step Principal Component's Approach. p. 406.

Bernanke, B., J. Boivin, and P. Elias. "Measuring the Effects of Monetary Policy; A Factor Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics* 120, no. 1 (February 2005): 387-422.

Application 3. Forni, Gianone, Lippi, and Reichlin 2006. Technology shocks.

- Standard DFM.
- 89 quarterly series for X_t , transformed to be stationary. 1950:1 to 1988:4
- $r = 12 - 18$, $q = 3$.
- Identification assumption. Only one shock (technological shock?) has a long run effect on per capita output. So partial identification.
- Impulse response. Figure 3. Very similar to Gali 1999. Variance decomposition: Table 1. 78% for output after 2 years. (BQ: 18 to 80% after 2 years)

Image removed due to copyright restrictions.

Figure 3. The Impulse Response Function of the Long-run Shock on Output. p. 25.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin. "Opening the Black Box: Structural Factor Models with Large Cross-Sections." European Central Bank Working Paper Series No. 712, January 2007. pp. 1-40.

Image removed due to copyright restrictions.

Table 1. Fraction of the Forecast-Error Variance Due to the Long-run Shock. p. 28

Forni, M., D. Giannone, M. Lippi, and L. Reichlin. "Opening the Black Box: Structural Factor Models with Large Cross-Sections." European Central Bank Working Paper Series No. 712, January 2007. pp. 1-40.

Dynamic factor models. How many shocks? Tentative conclusions

- Clearly the right approach in many contexts.
People/the Fed look at many variables.
Less big shocks than variables, more shocks than any number of variables we can include.
- In terms of results. Verdict out yet. A few common shocks? Common shocks, quantities, prices, and asset prices.
- In general, the usual: Need for identification.
- Non-linear true models and factors. For example, one shock, quadratic model: Two factors: mean and variance.