Backus, Chernov, & Zin

Notes on Identification of Taylor Rules*

(Started: July 23, 2010; Revised: December 20, 2011)

Using yields

Suppose under Q, the state follows:

$$x_{t+1} = \Phi^Q x_t + \Sigma\epsilon_{t+1}, \tag{1}$$

where Φ^Q is a diagonal matrix and the diagonal of the covariance matrix $\Sigma\Sigma'$ is full of ones. These are restrictions that ensure identification of the Q parameters based on yields. Then under P, we have

$$x_{t+1} = \mu + \Phi x_t + \Sigma\epsilon_{t+1}. \tag{2}$$

There are no restrictions on μ and Φ. These two equations define risk premia $\Lambda_t = \Lambda_0 + \Lambda_x' x_t$ via:

$$\begin{aligned} \Phi^Q &= \Phi - \Sigma\Lambda_x' & (3)\\ 0 &= \mu - \Sigma\Lambda_0. & (4) \end{aligned}$$

Normally, one specifies the nominal interest rate as a linear function of these state variables, $i_t = \delta' x_t$. Yields data allow one to estimate the following set of parameters $(\Phi^Q, \Sigma, \mu, \Phi, \delta)$.

Assume that inflation is a linear function of these state variables, $\pi_t = \alpha' x_t$. Because we imposed a restriction on the diagonal of the covariance matrix, α can be free. Denote the real interest rate by r_t. Then the nominal and real log pricing kernels are related via:

$$m^{\$}_{t+1} = m_{t+1} - \pi_{t+1} = -r_t - 1/2 \cdot \Lambda_t'\Lambda_t - \Lambda_t'\epsilon_{t+1} - \alpha' x_{t+1}. \tag{5}$$

Therefore, the nominal interest rate rate is:

$$\begin{aligned} i_t &= y^{\$}_t(1) = -\log P^{\$}_t(1) = -E_t(m_{t+1} - \pi_{t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1} - \pi_{t+1}) \\ &= r_t + \underbrace{\alpha'(\mu + \Phi x_t)}_{E_t\pi_{t+1}} - \underbrace{1/2 \cdot \alpha'\Sigma\Sigma'\alpha}_{\text{convexity}} - \underbrace{\alpha'\Sigma\Lambda_t}_{\text{inflation premium}}. \qquad (6) \end{aligned}$$

This is a modern version of the Fisher equation that incorporates risk premia. Equations (3), (4) imply

$$i_t = r_t + \alpha'\Phi^Q x_t - 1/2 \cdot \alpha'\Sigma\Sigma'\alpha. \tag{7}$$

This expression for the nominal interest rate must imply the same as the original $i_t = \delta' x_t$. If one specifies the real rate r_t as a linear function of x, then one can explicitly compute δ if

*Working notes, no guarantee of accuracy or sense.

the set of loadings for r_t and α are known, or vice-versa, one can infer α if δ and the set of loadings for r_t are known. Here we do not want to take a stand on the real rate and treat δ as a nuisance parameter. Instead, we infer α from the monetary policy. We postulate monetary policy (MP) via the Taylor rule:

$$TR_t = \gamma_0 + r_t + \gamma_\pi \pi_t + \gamma' x_t. \tag{8}$$

In equilibrium, $i_t = TR_t$, and we can solve for π_t by solving for α :

$$\alpha' \Phi^Q x_t - 1/2 \cdot \alpha' \Sigma\Sigma' \alpha = \gamma_0 + \gamma_\pi \alpha' x_t + \gamma' x_t. \tag{9}$$

We obtain a system of two equations:

$$\begin{aligned} \alpha' \Phi^Q x_t &= \gamma_\pi \alpha' x_t + \gamma' x_t & (10) \\ -1/2 \cdot \alpha' \Sigma\Sigma' \alpha &= \gamma_0. & (11) \end{aligned}$$

The first equation implies:

$$\alpha' = \gamma' \left(\Phi^Q - \gamma_\pi I \right)^{-1}. \tag{12}$$

Because Φ^Q is diagonal, $\alpha_i = \gamma_i / (\phi^Q_{ii} - \gamma_\pi)$.

Compare this to the usual implementation of Taylor rules in affine term structure models. Authors specify the nominal interest rate in the form $i_t = \psi' x_t + \psi_\pi \pi_t$ and interpret ψ_π at the degree of the policy response. However, $\pi_t = \alpha' x_t$ implies that $x_{jt} = \alpha_j^{-1} \pi_t - \alpha'_{-j} \alpha_j^{-1} x_{-jt}$. Therefore, $i_t = (\psi_\pi + \psi_j \alpha_j^{-1}) \pi_t + f(x_{-jt})$ and the policy response is not identified.

Back to identification, the vector $y_t = (\pi_t, i_t)'$ is related to the shocks x_t (ignoring constants and r_t) via

$$y_t = C x_t, \tag{13}$$

where, assuming that $x = (x_1, x_2, \ldots, x_n)'$,

$$C = \begin{pmatrix} \gamma_1/(\phi^Q_{11} - \gamma_\pi) & \gamma_2/(\phi^Q_{22} - \gamma_\pi) & \ldots & \gamma_n/(\phi^Q_{nn} - \gamma_\pi) \\ \gamma_1 \phi^Q_{11}/(\phi^Q_{11} - \gamma_\pi) & \gamma_2 \phi^Q_{22}/(\phi^Q_{22} - \gamma_\pi) & \ldots & \gamma_n \phi^Q_{nn}/(\phi^Q_{nn} - \gamma_\pi) \end{pmatrix}. \tag{14}$$

The y_t follows the following VAR :

$$y_t = C \Phi C^{-1} y_{t-1} + \text{shock}. \tag{15}$$

We can estimate n^2 parameters of $\Phi^* = C \Phi C^{-1}$, while we have $n+1$ unknowns (n γ's and γ_π, Φ and Φ^Q are known from yields). So if we have more than one factor in a model, all the parameters, including γ_π, should be identified. Suppose Φ is not known either (why would this be? – the model is estimated by NLS). Then $n+1$ parameters are not identified. How could we achieve identification with additional restrictions?

1. One strategy is to say that Λ_t is too flexible and restrict risk premia, make Φ as close to Φ^Q as possible. This idea is connected to the literature stemming from the CP factor and constrained Sharpe ratios. For example, we can set $n+1$ elements of Λ_1 to zero. Note that Cochrane's example corresponds to $n = 2$ with all elements Λ_1 equal to zero. This is an overly restricted model.

2. Another strategy is to control how many x's act as shocks to the Taylor rule. Note that at most, we can impose n restrictions as there n γ's. Thus, this strategy on its own is not sufficient.

Using survey-based inflation expectations

Another way to write the equilibrium $i = TR$ is

$$E_t(\pi_{t+1}) - \alpha'\Sigma\Lambda_t = \gamma_\pi \pi_t + \gamma' x_t. \tag{16}$$

Forward-looking Taylor rule

[xxx Incomplete xxx]

We postulate monetary policy (MP) via the forward-looking Taylor rule:

$$TR_t = \gamma_0 + r_t + \gamma_\pi \frac{1}{\tau} E_t \left(\sum_{j=1}^{\tau} \pi_{t+j} \right) + \gamma' x_t. \tag{17}$$

The expected inflation can be computed

$$E_t \left(\sum_{j=1}^{\tau} \pi_{t+j} \right) = \alpha' \left(\Psi^\tau \mu + \Phi^\tau x_t \right), \tag{18}$$

where

$$\Psi^\tau \equiv \sum_{k=0}^{\tau-1} \Phi^k = (I - \Phi)^{-1} (I - \Phi^\tau). \tag{19}$$

In equilibrium, $i_t = TR_t$, and we can solve for π_t by solving for α :

$$\alpha'\Phi^Q x_t - 1/2 \cdot \alpha'\Sigma\Sigma'\alpha = \gamma_0 + \tilde{\gamma}_\pi \alpha' \left(\Psi^\tau \mu + \Phi^\tau x_t \right) + \gamma' x_t, \tag{20}$$

where $\tilde{\gamma}_\pi = \gamma_\pi / \tau$. We obtain a system of two equations:

$$\begin{aligned} \alpha'\Phi^Q x_t &= \gamma_\pi \alpha' \Phi^\tau x_t + \gamma' x_t & (21) \\ -1/2 \cdot \alpha'\Sigma\Sigma'\alpha &= \gamma_0 + \tilde{\gamma}_\pi \alpha' \Psi^\tau \mu. & (22) \end{aligned}$$

The first equation implies:

$$\alpha' = \gamma' \left(\Phi^Q - \tilde{\gamma}_\pi \Phi^\tau \right)^{-1}. \tag{23}$$

Using nominal yields, start with the real ones

Suppose under Q, the state follows:

$$x_{t+1} = \Phi^Q x_t + \Sigma\epsilon_{t+1}, \tag{24}$$

where Φ^Q is a diagonal matrix and the diagonal of the covariance matrix $\Sigma\Sigma'$ is full of ones. These are restrictions that ensure identification of the Q parameters based on yields.

Assume that the the real interest rate is

$$r_t = \rho_0 + \rho_x' x_t \tag{25}$$

and the real pricing kernel is

$$\log m_{t+1} = -r_t - \Lambda_t'\Lambda_t - \Lambda_t\epsilon_{t+1}, \tag{26}$$

where

$$\Lambda_t = \Lambda_0 + \Lambda_x' x_t. \tag{27}$$

Now, suppose we have a nominal shock z_t that is independent from x. Under Q z is:

$$z_{t+1} = \phi_z^Q z_t + u_{t+1}. \tag{28}$$

Assume inflation is a linear function of all the factors:

$$\pi_t = \alpha_0 + \alpha_x' x_t + \alpha_z z_t. \tag{29}$$

Then the nominal and real log pricing kernels are related via:

$$\log m_{t+1}^{\$} = \log m_{t+1} - \pi_{t+1} = -r_t - 1/2 \cdot \Lambda_t'\Lambda_t - \Lambda_t'\epsilon_{t+1} - \alpha_0 - \alpha_x' x_{t+1} - \alpha_z z_{t+1}. \tag{30}$$

Therefore, the nominal interest rate rate is:

$$\begin{aligned} i_t &= y_t^{\$}(1) = -\log P_t^{\$}(1) = -E_t(\log m_{t+1} - \pi_{t+1}) - \frac{1}{2}\mathrm{Var}_t(\log m_{t+1} - \pi_{t+1}) \\ &= r_t + E_t\pi_{t+1} - \underbrace{[\alpha_x'\Sigma\Sigma'\alpha_x + \alpha_z^2]/2}_{\text{convexity}} - \underbrace{\alpha_x'\Sigma\Lambda_t}_{\text{inflation premium}} \\ &= \underbrace{\rho_0 + \rho_x' x_t}_{r_t} + \underbrace{\alpha_0 + \alpha_x'\Phi^Q x_t + \alpha_z\phi_z^Q z_t}_{E_t^Q\pi_{t+1}} - \underbrace{[\alpha_x'\Sigma\Sigma'\alpha_x + \alpha_z^2]/2}_{\text{convexity}}. \end{aligned} \tag{31}$$

We postulate monetary policy (MP) via the Taylor rule:

$$TR_t = \gamma_0 + \gamma_\pi\pi_t + \gamma_z z_t. \tag{32}$$

In equilibrium, $i_t = TR_t$, and we can solve for π_t by solving for α :

$$\rho_0 + \rho_x' x_t + \alpha_0 + \alpha_x'\Phi^Q x_t + \alpha_z\phi_z^Q z_t - [\alpha_x'\Sigma\Sigma'\alpha_x + \alpha_z^2]/2 = \gamma_0 + \gamma_\pi\pi_t + \gamma_z z_t. \tag{33}$$

We obtain a system of two equations:

$$\rho_x' x_t + \alpha_x' \Phi^Q x_t + \alpha_z \phi_z^Q z_t = \gamma_\pi \alpha_x' x_t + \gamma_\pi \alpha_z z_t + \gamma_z z_t \quad (34)$$

$$\rho_0 + \alpha_0 - [\alpha_x' \Sigma \Sigma' \alpha_x + \alpha_z^2]/2 = \gamma_0 + \gamma_\pi \alpha_0. \quad (35)$$

The first equation implies:

$$\rho_x' + \alpha_x' \Phi^Q = \gamma_\pi \alpha_x' \quad (36)$$

$$\alpha_z \phi_z^Q = \gamma_\pi \alpha_z + \gamma_z \quad (37)$$

and

$$\alpha_x = -\rho_x' (\Phi^Q - \gamma_\pi I)^{-1} \quad (38)$$

$$\alpha_z = \gamma_z / (\phi_z^Q - \gamma_\pi) \quad (39)$$

Back to identification, the vector $y_t = (\pi_t, i_t)'$ is related to the shocks $\tilde{x}_t = (x_t', z_t)'$ via (ignoring constants)

$$y_t = C \tilde{x}_t, \quad (40)$$

where, assuming that $x = (x_1, x_2, \ldots, x_n)'$,

$$C = \begin{pmatrix} -\rho_{x,1}/(\phi_{11}^Q - \gamma_\pi) & \ldots & -\rho_{x,n}/(\phi_{nn}^Q - \gamma_\pi) & \gamma_z/(\phi_z^Q - \gamma_\pi) \\ \rho_{x,1}(1 - \phi_{11}^Q/(\phi_{11}^Q - \gamma_\pi)) & \ldots & \rho_{x,n}(1 - \phi_{nn}^Q/(\phi_{nn}^Q - \gamma_\pi)) & \gamma_z \phi_z^Q/(\phi_z^Q - \gamma_\pi) \end{pmatrix}. \quad (41)$$

The y_t follows the following VAR :

$$y_t = C \Phi C^{-1} y_{t-1} + \text{shock}. \quad (42)$$

We can estimate $(n+1)^2$ parameters of $\Phi^* = C \Phi C^{-1}$, while we have 2 unknowns (γ_π and γ_z, ρ_x, Φ and Φ^Q are known from yields). So if we have more than one factor in a model, all the parameters, including γ_π, should be identified.

Square-root case

Why does it work?

I will revisit the same basic example using state-space notation, rather than the MA one.

$$i_t = r + E_t p_{t+1} + x_{1t}$$

$$i_t = r + \tau p_t + x_{2t}.$$

Suppose both x's are AR(1) with a joint (diagonal) persistence matrix Φ and look for p_t in the form $\alpha_1 x_{1t} + \alpha_2 x_{2t}$.

Recall, yet again, that the identification problem is: the vector $y_t = (\pi_t, i_t)'$ is related to the shocks x_t (ignoring constants) via

$$y_t = Cx_t. \tag{43}$$

Therefore, the y_t follows the following VAR :

$$y_t = C\Phi C^{-1} y_{t-1} + \text{shock}. \tag{44}$$

Can we recover τ on the basis of the estimated $A = C\Phi C^{-1}$?

When we have only one shock, this VAR collapses to one equation. In the Cochrane example we have:

$$i_t = r + \phi_2 \alpha_2 x_{2t} = \phi_2 p_t.$$

Such a "VAR" tells us nothing about τ. In the opposite case we have:

$$i_t = r + \tau p_t.$$

This "VAR" tells us all we need to know about τ.

One more try

$$\begin{aligned} y_t &= Cx_t \\ x_t &= \Phi x_{t-1} + Gw_t \end{aligned}$$

Then, the VAR is

$$y_t = C^{-1}\Phi C y_{t-1} + Gw_t$$

What are the general conditions for recovering C from the VAR?

In our case,

$$\begin{aligned} i_t &= E_t p_{t+1} + x_{1t} \\ i_t &= \tau p_t + x_{2t}. \end{aligned}$$

Guess,

$$p_t = \alpha' x_t.$$

Then,

$$\alpha' = (e_2' - e_1')(\Phi - \tau I)^{-1},$$

where e_i is a vector of zeros with a one in position i, and I is the identity matrix. Then

$$C = \begin{pmatrix} \alpha' + e_1' \\ \alpha' \end{pmatrix} \equiv \begin{pmatrix} \tau\alpha' + e_2' \\ \alpha' \end{pmatrix}.$$

Regardless of which shock we shut down, x_1 or x_2, C becomes singular. It easy to see now why we are still recovering τ when $x_2 = 0$.

1 Identification in the generic case

The starting point is the last slide of Dave's notes:

$$\begin{aligned} & y_t = \lambda E_t y_{t+1} + e_1' x_t + e_2' x_t \\ \Rightarrow \quad & (e_1 + e_2)'(I - \lambda A)^{-1} x_t = \beta' x_t \end{aligned}$$

For clarity sake, I (MC) prefer to think of x as observable (so A should be replaced by $\hat{A}$), but I will keep the notation the same as in Dave's slides.

We know e_1 because it comes from the pricing kernel (either estimate from yields or use structural restrictions from a GE model). Our idea is to impose identifying assumptions on e_2. Some of the elements of x_t affect the Taylor rule only (think MP shocks). We have to assume that at least one of x's does not effect the pricing equation. To be specific, assume that $e_{2,1} = 0$. Can we identify λ in this case?

We will study local identification. That is, a derivative of the matrix

$$(e_1 + e_2)'(I - \lambda A)^{-1} - \beta'$$

that is implied by the second equation should have a full rank. A derivative w.r.t. what? This is w.r.t. vector of the unknowns $(\lambda, e_{2,2}, e_{2,3}, \ldots, e_{2,n})$.

The derivative with respect to $e_{2,i}$, $i = 2, \ldots, n$ gives us the ith row of the matrix

$$(I - \lambda A)^{-1}.$$

The derivative with respect to λ is:

$$(e_1 + e_2)'(I - \lambda A)^{-1} A (I - \lambda A)^{-1}.$$

The resulting matrix has a full rank if A is "regular."

2 Isolating the "MP shock" state

Suppose we ended up with n observable states where the first element is i_t and the rest is a $n - 1 \times 1$ vector of inflation forecasts:

$$\begin{pmatrix} f_t \\ i_t \end{pmatrix} = \hat{A} \begin{pmatrix} f_{t-1} \\ i_{t-1} \end{pmatrix} + \hat{C} w_t.$$

We rotate this vector into a new one $(f_t', z_t)'$, where z_t will serve as the MP shock. The idea of the MP shock is that it may affect other states, but it is not affected by other states. The rotation is:

$$\begin{pmatrix} f_t \\ z_t \end{pmatrix} = B \begin{pmatrix} f_t \\ i_t \end{pmatrix},$$

where

$$B = \left(\begin{array}{cc} I_{n-1} & 0 \\ b_n' & b_{nn} \end{array} \right),$$

where I_{n-1} is a $(n-1) \times (n-1)$ identity matrix, and b_n is a vector of size $n-1$. Therefore,

$$\left(\begin{array}{c} f_t \\ z_t \end{array} \right) = B\hat{A}B^{-1} \left(\begin{array}{c} f_{t-1} \\ z_{t-1} \end{array} \right) + B\hat{C}w_t.$$

We can pick any matrix B to rotate the state without changing the implications for i_t or p_t. We choose a specific rotation that matches our view of z_t as an MP shock. First,

$$B^{-1} = \left(\begin{array}{cc} I_{n-1} & 0 \\ -b_{nn}^{-1}b_n' & b_{nn}^{-1} \end{array} \right).$$

Therefore, the nth row of $B\hat{A}B^{-1}$ consists of the following terms. The first $n-1$ elements are of the form:

$$\sum_{j=1}^{n} B_{nj} \left(\hat{A}_{ji} - B_{ni}/B_{nn} \cdot \hat{A}_{jn} \right), i = 1, \ldots, n-1.$$

The last element is:

$$1/B_{nn} \cdot \sum_{j=1}^{n} B_{nj}\hat{A}_{jn}.$$

We need the first $n-1$ elements to be equal to zero, and for simplicity we want the last element to be equal to 1. This results in n equations in n unknowns.

Thus, we obtain z_t up to a scale. This is OK because it enters the MP equation as $e'z_t$ where we treat e as unknown.