You know, I don't know about that. 16 March 2020

- I wanna pause OSI for a sec b/c I can't seem to get it to work and I'm confused about how to get it time-varying anyway. So let's turn to estimation.

Estimation of the anchoring function

The issue is that we wanna estimate the anchoring fct together w/ the model. On the top of my head I can think of 3 ways of doing that:

1) IR-matching
2) Likelihood-based (either MLE or Bayesian)
3) VAR-representation ⇒ exist? est. that!
   ↳ it would be a time-varying one.

↳ I'm leaning toward #2 b/c 1) it's sexier
2) it's more general than conditional on shocks
3) a TV-VAR sounds challenging

The thing is: 17 March 2020

need to derive the log-likelihood of my model

Take the midsimple model w/ the TR.

In materials 21, this is eq. (5)-(10) + TR

$$\pi_t = \kappa x_t - (1-\alpha)\beta f_a(t) + \left[-\kappa\alpha\beta(I_3-\alpha\beta h_x)^{-1} - e_3(I_3-\alpha\beta h_x)^{-1}\right]s_t$$

$$x_t = -\sigma i_t - \sigma f_b(t) + \left[-(1-\beta)b_2(I_3-\beta h_x)^{-1} + \sigma\beta b_3(I_3-\beta h_x)^{-1} - \sigma e_1(I_3-\beta h_x)^{-1}\right]s_t$$

$$f_a(t) = \frac{1}{1-\alpha\beta}\bar{\pi}_{t-1} - b_1(I_3-\alpha\beta h_x)^{-1}s_t$$

$$f_b(t) = \frac{1}{1-\beta}\bar{\pi}_{t-1} - b_1(I_3-\beta h_x)^{-1}s_t$$

$$\bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1}\left(\pi_t - (\bar{\pi}_{t-1} + b_1 s_{t-1})\right)$$

$$k_t^{-1} = k_{t-1}^{-1} + d\left(\pi_t - (\bar{\pi}_{t-1} - b_1 s_{t-1})\right) + c$$

$$i_t = \psi_\pi \pi_t + \psi_x x_t + \bar{i}_t$$

This is a state-space model (believe it or not)

and I'm gonna eliminate some variables

$$\pi_t = \kappa x_t - \frac{(1-\alpha)\beta}{1-\alpha\beta}\bar{\pi}_{t-1} - (1-\alpha)\beta b_1(I_3-\alpha\beta h_x)^{-1}s_t$$
$$+ \left[-\kappa\alpha\beta(I_3-\alpha\beta h_x)^{-1} - e_3(I_3-\alpha\beta h_x)^{-1}\right]s_t$$

$$x_t = -\sigma\psi_\pi\pi_t - \sigma\psi_x x_t - \sigma\bar{i}_t - \frac{\sigma}{1-\beta}\bar{\pi}_{t-1} - \sigma b_1(I_3-\beta h_x)^{-1}s_t$$
$$+ \left[-(1-\beta)b_2(I_3-\beta h_x)^{-1} + \sigma\beta b_3(I_3-\beta h_x)^{-1} - \sigma e_1(I_3-\beta h_x)^{-1}\right]s_t$$

$$\pi_t = \kappa x_t - \frac{(1-\alpha)\beta}{1-\alpha\beta}\bar{\pi}_{t-1}$$

$$+ \left[-(1-\alpha)\beta b_1 (I_3 - \alpha\beta h x)^{-1} s_t - \kappa\alpha\beta (I_3 - \alpha\beta h x)^{-1} - e_3 (I_3 - \alpha\beta h x)^{-1}\right] s_t$$

$$x_t = -\sigma\psi_\pi \pi_t - \sigma\psi_x x_t - \sigma \bar{i}_t - \frac{\sigma}{1-\beta}\bar{\pi}_{t-1} - \sigma b_1 (I_3 - \beta h x)^{-1} s_t$$

$$+ \left[-(1-\beta) b_2 (I_3 - \beta h x)^{-1} + \sigma\beta b_3 (I_3 - \beta h x)^{-1} - \sigma e_1 (I_3 - \beta h x)^{-1}\right] s_t$$

$$(1+\sigma\psi_x) x_t = -\sigma\psi_\pi \pi_t \qquad - \frac{\sigma}{1-\beta}\bar{\pi}_{t-1}$$

$$\left[-\sigma e_2 - \sigma b_1 (I_3 - \beta h x)^{-1} - (1-\beta) b_2 (I_3 - \beta h x)^{-1} + \sigma\beta b_3 (I_3 - \beta h x)^{-1} - \sigma e_1 (I_3 - \beta h x)^{-1}\right] s_t$$

$$\Rightarrow \quad x_t = -\frac{\sigma\psi_\pi}{1+\sigma\psi_x}\pi_t - \frac{1}{1+\sigma\psi_x}\frac{\sigma}{1-\beta}\bar{\pi}_{t-1}$$

$$+ \frac{1}{1+\sigma\psi_x}\left[-\sigma e_2 - \sigma b_1 (I_3 - \beta h x)^{-1} - (1-\beta) b_2 (I_3 - \beta h x)^{-1} + \sigma\beta b_3 (I_3 - \beta h x)^{-1} - \sigma e_1 (I_3 - \beta h x)^{-1}\right] s_t$$

Can even sub x_t out!

$$\pi_t = -\frac{\kappa\sigma\psi_\pi}{1+\sigma\psi_x}\pi_t - \frac{\kappa}{1+\sigma\psi_x}\frac{\sigma}{(1-\beta)}\bar{\pi}_{t-1}$$

$$+ \frac{\kappa}{1+\sigma\psi_x}\left[-\sigma e_2 - \sigma b_1 (I_3 - \beta h x)^{-1} - (1-\beta) b_2 (I_3 - \beta h x)^{-1} + \sigma\beta b_3 (I_3 - \beta h x)^{-1} - \sigma e_1 (I_3 - \beta h x)^{-1}\right] s_t$$

$$- \frac{(1-\alpha)\beta}{1-\alpha\beta}\bar{\pi}_{t-1}$$

$$+ \left[-(1-\alpha)\beta b_1 (I_3 - \alpha\beta h x)^{-1} s_t - \kappa\alpha\beta (I_3 - \alpha\beta h x)^{-1} - e_3 (I_3 - \alpha\beta h x)^{-1}\right] s_t$$

$$\Leftrightarrow \left(1 + \frac{\kappa\sigma\psi_\pi}{1+\sigma\psi_x}\right)\pi_t = -\left(\frac{\kappa\sigma}{1+\sigma\psi_x(1-\beta)} + \frac{(1-\alpha)\beta}{(1-\alpha\beta)}\right)\bar{\pi}_{t-1} + \text{stuff} \cdot s_t$$

$$\left(1+\frac{\kappa\sigma\psi_\pi}{1+\sigma\psi_x}\right)\pi_t = -\left(\frac{\kappa\sigma}{1+\sigma\psi_x(1-\beta)} + \frac{(1-\alpha)\beta}{(1-\alpha\beta)}\right)\bar{\pi}_{t-1} \quad +$$

$$\left\{\frac{\kappa}{1+\sigma\psi_x}\left[-\sigma e_2 - \sigma b_1(f_3-\beta hx)^{-1} \;-(1-\beta)b_2(f_3-\beta hx)^{-1} + \sigma\beta b_3(f_3-\beta hx)^{-1} - \sigma e_1(f_3-\beta hx)^{-1}\right]\right.$$

$$\left. + \left[-(1-\alpha)\beta b_1(f_3-\alpha\beta hx)^{-1}s_1 - \kappa\alpha\beta(f_3-\alpha\beta hx)^{-1} - e_3(f_3-\alpha\beta hx)^{-1}\right]\right\} s_t$$

$$\frac{1+\sigma\psi_x+\kappa\sigma\psi_\pi}{1+\sigma\psi_x}\,\pi_t = \text{same}$$

$$\Rightarrow \pi_t = -\frac{\cancel{1+\sigma\psi_x}}{1+\sigma\psi_x+\kappa\sigma\psi_\pi}\left(\frac{\kappa\sigma(1-\alpha\beta) + (1-\alpha)\beta(1-\beta)(1+\sigma\psi_x)}{\cancel{(1+\sigma\psi_x)}(1-\beta)(1-\alpha\beta)}\right)\bar{\pi}_{t-1}$$

$$+\left\{\frac{\kappa}{1+\sigma\psi_x+\kappa\sigma\psi_\pi}\left[-\sigma e_2 - \sigma b_1(f_3-\beta hx)^{-1} \;-(1-\beta)b_2(f_3-\beta hx)^{-1} + \sigma\beta b_3(f_3-\beta hx)^{-1} - \sigma e_1(f_3-\beta hx)^{-1}\right]\right.$$

$$\left. + \frac{1+\sigma\psi_x}{1+\sigma\psi_x+\kappa\sigma\psi_\pi}\left[-(1-\alpha)\beta b_1(f_3-\alpha\beta hx)^{-1}s_1 - \kappa\alpha\beta(f_3-\alpha\beta hx)^{-1} - e_3(f_3-\alpha\beta hx)^{-1}\right]\right\} \cdot s_t$$

$$\Rightarrow \pi_t = -\frac{\kappa\sigma(1-\alpha\beta) + \beta(1-\alpha)(1-\beta)(1+\sigma\psi_x)}{(1+\sigma\psi_x+\kappa\sigma\psi_\pi)(1-\beta)(1-\alpha\beta)}\bar{\pi}_{t-1} \quad +$$

$$\left\{\frac{\kappa}{1+\sigma\psi_x+\kappa\sigma\psi_\pi}\left[-\sigma e_2 - \sigma b_1(f_3-\beta hx)^{-1} \;-(1-\beta)b_2(f_3-\beta hx)^{-1} + \sigma\beta b_3(f_3-\beta hx)^{-1} - \sigma e_1(f_3-\beta hx)^{-1}\right]\right.$$

$$\left. + \frac{1+\sigma\psi_x}{1+\sigma\psi_x+\kappa\sigma\psi_\pi}\left[-(1-\alpha)\beta b_1(f_3-\alpha\beta hx)^{-1}s_1 - \kappa\alpha\beta(f_3-\alpha\beta hx)^{-1} - e_3(f_3-\alpha\beta hx)^{-1}\right]\right\} s_t$$

Damn dam! So we have $\equiv A$

$$\pi_t = -\frac{\kappa\sigma(1-\alpha\beta)+\beta(1-\alpha)(1-\beta)(1+\sigma\psi_x)}{(1+\sigma\psi_x+\kappa\sigma\psi_\pi)(1-\beta)(1-\alpha\beta)}\bar{\pi}_{t-1} +$$

$$\Bigg\{\frac{k}{1+\sigma\psi_x+\kappa\sigma\psi_\pi}\left[-\sigma e_2-\sigma b_1(f_3-\beta hx)^{-1}-(1-\beta)b_2(f_3-\beta hx)^{-1}+\sigma\beta b_3(f_3-\beta hx)^{-1}-\sigma e_1(f_3-\beta hx)^{-1}\right]$$

$$+\frac{1+\sigma\psi_x}{1+\sigma\psi_x+\kappa\sigma\psi_\pi}\left[-(1-\alpha)\beta b_1(f_3-\alpha\beta hx)^{-1}s_1-\kappa\alpha\beta(f_3-\alpha\beta hx)^{-1}-e_3(f_3-\alpha\beta hx)^{-1}\right]\Bigg\}s_t$$

$\equiv B$

$$\bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1}\left(\pi_t - (\bar{\pi}_{t-1} + b_1 s_{t-1})\right)$$

$$k_t^{-1} = k_{t-1}^{-1} + d\left(\bar{\pi}_t - (\bar{\pi}_{t-1} - b_1 s_{t-1})\right) + c$$

$\hookrightarrow$ 1 jump (π_t), 3 exog states ($s_t = \begin{bmatrix} r_t^n \\ \bar{i}_t \\ u_t \end{bmatrix}$) and 2 endog states $\xi_t = \begin{bmatrix} \bar{\pi}_t \\ k_t^{-1} \end{bmatrix}$ (or $\begin{bmatrix} \bar{\pi}_{t-1} \\ k_t^{-1} \end{bmatrix}$) so $X_t = \begin{bmatrix} \xi_t \\ s_t \end{bmatrix}$ (states)

$$\pi_t = A\,\bar{\pi}_{t-1} + B\,s_t = \begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} \xi_t \\ s_t \end{bmatrix}$$

$$\bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1}\left(\pi_t - (\bar{\pi}_{t-1} + b_1 s_{t-1})\right)$$

$$k_t^{-1} = k_{t-1}^{-1} + d\Big(\underbrace{\bar{\pi}_t - (\bar{\pi}_{t-1} - b_1 s_{t-1})}_{\text{call this the state (?) } fe_{t-1}}\Big) + c$$

$$\pi_t = A\bar{\pi}_{t-1} + B s_t = \begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} \xi_t \\ s_t \end{bmatrix} \qquad Y_t = gx \cdot X_t$$

$$\bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1} fe_{t-1} \qquad X_{t+1} = hx\, X_t + \eta\, \varepsilon_t$$

$$k_t^{-1} = k_{t-1}^{-1} + d \cdot fe_{t-1} + c$$

$$fe_{t-1} = \pi_t - \bar{\pi}_{t-1} - b_1 s_{t-1}$$

Several issues y'all:

1) fe state or jump? depends on π_t (a jump)

→ there's gotta be some trick (like π_t in LEMP) to make it a pure state

2.) $\bar{\pi}_t$ nonlinear LOM!

↙

ok - that's troubling. But let's pause it for a sec & let's read what Lütkepohl has to say about MLE & log-likelihoods.

Supp. our VAR(p) looks like

$$y_t - \mu = A_1(y_{t-1} - \mu) + \ldots + A_p(y_{t-p} - \mu) + u_t \quad (3.3.1)$$

p. 75

Then if the VAR(p) is Gaussian, that is

$$u \equiv \mathrm{vec}(U) = \begin{bmatrix} u_1 \\ \vdots \\ u_T \end{bmatrix} \sim N(0, I_T \otimes \Sigma_u), \quad \text{which}$$

equivalently means that the prob. density of u is

$$f_u(u) = \frac{1}{(2\pi)^{kT/2}} \left| I_T \otimes \Sigma_u \right|^{-1/2} \exp\left[-\frac{1}{2} u' \left(I_T \otimes \Sigma_u^{-1} \right) u \right]$$

then we can use the fact that $u = y - \mu^* - (X' \otimes I_K)\alpha$

where $\underset{k^2p \times 1}{\alpha} := \mathrm{vec}(A)$, $A := (A_1, \ldots, A_p)$ $\quad k \times kp$

$Y^0 := (y_1 - \mu, \ldots, y_T - \mu)$ $\quad k \times T$

$Y_t^0 := \begin{bmatrix} y_t - \mu \\ \vdots \\ y_{t-p+1} - \mu \end{bmatrix}$ $\quad (kp \times 1)$

$X := (Y_0^0, \ldots, Y_{T-1}^0)$,

to write

$$f_y(y) = \left| \frac{\partial u}{\partial y} \right| f_u(u)$$

$$= \frac{1}{(2\pi)^{kT/2}} \left| I_T \otimes \Sigma_u \right|^{-1/2} \exp\left[-\frac{1}{2} \left(y - \mu^* - (X' \otimes I_K)\alpha \right)' \left(I_T \otimes \Sigma_u^{-1} \right) \cdot \left(y - \mu^* - (X' \otimes I_K)\alpha \right) \right] \quad (3.4.4)$$

and

$$\ln l(\mu, \alpha, \Sigma_u) = -\frac{kT}{2} \ln(2\pi) - \frac{T}{2} \ln|\Sigma_u| - \frac{1}{2} \mathrm{tr}(Y^0 - AX)' \Sigma_u^{-1} (Y^0 - AX)$$

is the log-likelihood. $\quad (3.4.5)$

Given that the updating eq. w/ endog. gain 18 March 2020 introduces non-linearities, I'm afraid that even a "simple" & quick estimation has to involve some form of particle filter. But let's see whether Lütkepohl has anything interesting to say about 1) state-space models 2) non-linearities

Lütkepohl, State-space models (ch. 13, p. 415 ff.)

$$\underset{N\times 1}{z_{t+1}} = \underset{N\times N}{B_t} z_t + \underset{N\times m}{F_t} x_t + \underset{N\times 1}{w_t} \leftarrow \text{noise} \qquad t = 0, 1, 2, \ldots \qquad (13.2.1a)$$

$$\underset{k\times 1}{y_t} = \underset{k\times N}{H_t} z_t + \underset{k\times M}{G_t} \underset{M\times 1}{x_t} + \underset{k\times 1}{v_t} \qquad t = 1, 2, \ldots \qquad (13.2.2)$$

measurement error

transition matrix

measurement matrix

input matrices

inputs / instruments / policy vars

and $\begin{bmatrix} w_t \\ v_t \end{bmatrix} \sim WN\left[0, VC_t\right]$ $\quad VC_t = \begin{bmatrix} \Sigma_{w_t} & \Sigma_{w_t v_t} \\ \Sigma_{v_t w_t} & \Sigma_{v_t} \end{bmatrix}$

Nonlinear state-space models

$$z_{t+1} = b_t(z_t, x_t, w_t, \delta_1)$$

$$y_t = h_t(z_t, x_t, v_t, \delta_2)$$

vectors of params

Example of nonlin state-space is the "bilinear" model:

$y_t = \alpha y_{t-1} + u_t + \beta y_{t-1} u_{t-1}$ p. 427 w/ Refs.

$\hookrightarrow \quad z_{t+1} = B z_t + w_t + C\, \text{vec}(z_t z_t')$ (13.2.33)

$y_t = [I_K \; 0 \ldots 0]\, z_t$ (13.2.34)

$\Rightarrow$ Bunch of refs on bilinear systems, univariate & multivariate on p. 427 bottom.

MLE of state-space models p. 434

Gather the time-invariant params from $B, F, H_t, G, \Sigma_w, \Sigma_v$ and Σ_0 & μ_0 in δ as

$$\delta = \begin{bmatrix} \text{vec}[\nu, A_1, \ldots, A_p] \\ \text{vech}(\Sigma_u) \end{bmatrix}$$

where vech = "half-vectorization", $\text{vech}\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{bmatrix} a \\ b \\ d \end{bmatrix}$.

The log-likelihood for the Gaussian state-space model is:

$$\ln l(\delta | y_1, \ldots, y_T) = -\frac{KT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \ln |\Sigma_y(t|t-1)|$$

$$- \frac{1}{2} \sum_{t=1}^{T} (y_t - y_{t|t-1})' \Sigma_y(t|t-1)^{-1} (y_t - y_{t|t-1}) \quad (13.4.1)$$

Denoting the fcst error $e_t(\delta) := y_t - y_{t|t-1}$

and $\Sigma_t(\delta) := \Sigma_y(t|t-1)$, we can rewrite this as

$$\ln \mathcal{L}(\delta) = -\frac{kT}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left[\ln|\Sigma_t(\delta)| + e_t'(\delta)\Sigma_t(\delta)^{-1}e_t(\delta)\right] \tag{13.4.3}$$

which makes explicit

1) the dependence of $\ln \mathcal{L}$ on δ,
2) that all quantities in the $\ln \mathcal{L}$ all functions of δ and can (must) be computed using the Kalman filter.

locally identified: when in a subspace of the param space, δ is uniquely determined.

vs. globally identified: when δ is uniquely determined in the entire param space.

↳ identification: we need a min of $-\ln \mathcal{L}$, so we need some sort of Hessian = pos. def. ⟹ the information matrix, $= E[\text{Hessian}] = E\left[\frac{\partial^2(-\ln \mathcal{L})}{\partial\delta\partial\delta}\Big|_{\delta_0}\right]$

A quick note on DSEMs (dynamic simultaneous equations models a.k.a. "linear systems") p. 323

essentially, these are linear VARMAX(p,s,q) models

$$A_0 y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + B_0 x_t + B_1 x_{t-1} + \dots + B_s x_{t-s} + w_t \quad (10.1.1)$$

- VARMAX(p,s,q) is w_t is MA(q)
- VARX(p,s) if $w_t \sim WN$.

VAR(p) models w/ time-varying coefficients p. 391 ff.

periodic VARs → eg. w/ seasonal dummies

intervention models → DGP_1 is replaced by DGP_2 at time τ.

$$y_t = \nu_t + A_{1t} y_{t-1} + \dots + A_{pt} y_{t-p} + u_t \quad (12.2.1)$$

↳ $WN(0, \Sigma_t)$

also time-varying

(not identically distrib!)

Rewrite the VAR(p) as a VAR(1)

$$Y_t = v_t + A_t Y_{t-1} + U_t \qquad (12.2.2)$$

$$Y_t := \begin{bmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{bmatrix}, \quad v_t := \begin{bmatrix} v_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad A_t := \begin{bmatrix} A_{1,t} & \cdots & A_{p-1,t} & A_{p,t} \\ I_k & \ddots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & I_k & 0 \end{bmatrix}, \quad U_t := \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$kp \times 1$ $\quad$ $kp \times 1$ $\quad$ $kp \times kp$ $\quad$ $kp \times 1$

And by recursive sub we get

$$Y_t = \left[\prod_{j=0}^{h-1} A_{t-j}\right] Y_{t-h} + \sum_{i=0}^{h-1} \left[\prod_{j=0}^{i-1} A_{t-j}\right] v_{t-i} + \sum_{i=0}^{h-1} \left[\prod_{j=0}^{i-1} A_{t-j}\right] U_{t-i} \qquad (12.2.3)$$

Defining $J := [I_k \ 0]$ such that $y_t = JY_t$, we can premultiply (12.2.3) by J, define

$\Phi_{it} := J\left[\prod_{j=0}^{i-1} A_{t-j}\right] J'$ to get

$$y_t = \mu_t + \sum_{i=0}^{\infty} \Phi_{it} u_{t-i} \qquad (12.2.6)$$

where $\mu_t = E[y_t]$

$\Rightarrow$ Then the MSE (or FEV of the FE $y_{t+h} - y_t(h)$) is

$$\Sigma_t(h) := \sum_{i=0}^{h-1} \Phi_{i,t+h} \Sigma_{t+h-i} \Phi'_{i,t+h} \qquad (12.2.10)$$

MLE of TV-VAR   p. 394

Write (12.2.1) as $y_t = B_t Z_{t-1} + u_t$ (12.2.11)

where $B_t := [\nu_t, A_{1t}, \dots, A_{pt}]$, $Z_{t-1} := (1, Y_{t-1}')'$

B_t depend on the vector $\underset{N\times 1}{\gamma}$ of time-invariant params.

Σ_t depend on $\underset{M\times 1}{\sigma}$ of fixed params.

If $u_t \sim N(0, \Sigma_t)$, the log-likelihood is

$$\ln l(\gamma, \sigma) = -\frac{KT}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \ln |\Sigma_t| - \frac{1}{2} \sum_{t=1}^{T} u_t' \Sigma_t^{-1} u_t$$

(12.2.12)

where initial conditions have been ignored.

You can also derive an info-matrix.

So it seems like you just min $-\ln l(\gamma, \sigma)$!

It also seems like for certain special cases you can even derive the estimators in closed-form!

Ryan meeting 18 March 2020

Not yet convinced that the procedure is sensible
(above "procedure": param)

→ the only procedure that can work is

$f_{z,t}$ and more RHS-variables: k^{-1}, $\bar{\pi}$

$$z_t = \bar{z}_t + f_{z,t} z_{t-1} + f_{u,t} u_t$$

things in $f_{z,t}$ will be 0 or simply described by the LOMs

It may be too restrictive.

A computationally intensive exercise is:

- Conj. path for the int. rate (exog.)
- solve the model for that int. rate
- check the target criterion, compute the resid
- fmincon to min that resid

⇒ find simulated optimal plan.

· Simulate the Ramsey model
· Simulate the model w/ Taylor-rule
→ see how close you can get
It might then be that most results won't be pencil & paper.

Comments:

· Concerning w/ linearness of (10)
→ "negative surprises cause me to be unanchored" shouldn't be "big mistakes" → so take the square
↳ like smooth one better than jumpy

· Estim: filter the data → App paper w/ Robert:
match the params to moments cors of data
⇒ would give you results

they est an NK model: HP-filter both data and model, compute moments and try to match those.

Work after

It's the primal economy paper w/ Robert that Ryan meant. → chahrourWorthit.pdf in "literature."

⇒ They [1] take data on real per-capita output, inflation, nom interest rates & per-capita hours, [2] (BK) high-pass filter them, [3] Use GMM to estimate the process $\hat{\tau}_t$ that best-matches the autocov. structure of the data.

App. D p.46 (Mac).

Wedges are an MA(14):

$$\tau_t = \Phi_\varepsilon(L)\,\varepsilon_t + \Phi_u(L)\,u_t$$

ε_t: $WN(0, \Sigma_\varepsilon)$ $\quad$ u_t: $2\times1\ WN(0, I_2)$

⇒ params to be estimated are

$$\gamma_{ma} \equiv (\Phi_\varepsilon, \Phi_u, \Sigma_\varepsilon)$$

I think that the following is the target: $k=8$

$$\tilde{\Omega}_T = \text{vech}\left\{\text{Var}\left[\tilde{q}_t^{data}, \ldots, \tilde{q}_{t-k}^{data}\right]\right\} \quad q: y_t, \pi_t, i_t, n_t$$

Filtering: Baxter-King filter w/ truncation horizon 32, lag-length 12

→ $\tilde{q}_t \equiv BK_{32}(q_t)$, so $\tilde{q}_t$ is filtered data.

Then, they do a trick in converting γ_{ma} into γ_{ar} to finally estimate γ as:

$$\hat{\gamma}_{ar} = \underset{\gamma_{ar}}{\operatorname{argmin}} \left(\tilde{\Omega}_T - \tilde{\Omega}(\gamma_{ar})\right)' W^{-1} \left(\tilde{\Omega}_T - \tilde{\Omega}(\gamma_{ar})\right)$$

- W is a diagonal matrix w/ the bootstrapped variances of $\tilde{\Omega}_T$ along the main diagonal.
- The model analogue $\tilde{\Omega}(\gamma_{ar})$ is computed after the model data has been similarly filtered as the data.

↳ Hold it there, now I know what I need to do.

↳ That and

- the numerical implementation of the target criterion ⇒ both are things to do once I have the big screen. So now give a last try to do the time-varying one.

The one - a first shot

Ryan wrote:

$$z_t = \bar{z}_t + f_{z,t} z_{t-1} + f_{u,t} u_t$$

I think he meant $z_t = \bar{z}_t + h_{z,t} z_{t-1} + f_{z,t} u_t$

Let me ignore r_t^n b/c it just blows things up and I just wanna see if it works. Ignore $\bar{z}_t$ too.

$$h_{\pi,t}\pi_{t-1} + f_{\pi,t}u_t - \kappa(h_{x,t}x_{t-1} + f_{x,t}u_t)$$

$$-(1-\alpha)\beta\left(h_{fa,t}\,fa(t-1) + f_{fa,t}u_t\right) + exog_1 \cdot u_t = 0 \qquad (9)$$

$$h_{x,t}x_{t-1} + f_{x,t}u_t + \sigma(h_{i,t}i_{t-1} + f_{i,t}u_t)$$

$$-\sigma\left(h_{fb,t}\,fb(t-1) + f_{fb,t}u_t\right) + exog_2 \cdot u_t = 0 \qquad (10)$$

$\Leftrightarrow$

$$h_{\pi,t}\pi_{t-1} - \kappa h_{x,t}x_{t-1} - (1-\alpha)\beta h_{fa,t}\,fa(t-1)$$

$$+ \underbrace{\left(f_{\pi,t} - \kappa f_{x,t} - (1-\alpha)\beta f_{fa,t} + exog_1\right)}_{\text{this is exactly what I have in materials 21}} u_t = 0 \qquad (10)$$

$$h_{x,t}x_{t-1} + \sigma h_{i,t}i_{t-1} - \sigma h_{fb,t}\,fb(t-1)$$

$$+\left(f_{x,t} + \sigma f_{i,t} - \sigma f_{fb,t} + exog_2\right)u_t = 0$$

$$z_t = \bar{z}_t + h_{z,t}\, z_{t-1} + f_{z,t}\, u_t$$

$$h_{fa,t}\, f_a(t-1) + f_{fa,t}\, u_t - \frac{1}{1-\alpha\beta}\left(h_{\bar{\pi},t-1}\, \bar{\pi}_{t-2} + f_{\bar{\pi},t-1}\, u_{t-1}\right) + exog_3\, u_t$$

$$h_{fa,t}\, f_a(t-1) - \frac{1}{1-\alpha\beta} h_{\bar{\pi},t-1}\, \bar{\pi}_{t-2} + \underbrace{(f_{fa,t} + exog_3)\, u_t - \frac{1}{1-\alpha\beta} f_{\bar{\pi},t-1}\, u_{t-1} = 0}_{\text{same as in materials 21}} \quad (13)$$

$$h_{fb,t}\, f_b(t-1) - \frac{1}{1-\beta} h_{\bar{\pi},t-1}\, \bar{\pi}_{t-2} + (f_{fb,t} + exog_3)\, u_t - \frac{1}{1-\beta} f_{\bar{\pi},t-1}\, u_{t-1} = 0 \quad (14)$$

Honestly, I don't even think I will complete this b/c we can see that if, in say (14), $h_{fb,t} = 0$ and $h_{\bar{\pi},t-1} = 0$ then we're back to exactly the system I had in mat 21.

So suppose $h_{fb,t} \neq 0$. But then in each equation we have an earlier iteration of itself. Now I might argue $h_\pi = h_x = h_i = 0$, but if I ass that for $h_{fa,t} = h_{fb,t} = 0$ then again the link between (9)-(10) and (15)-(16) is broken. Then (14) would give an additional constraint

$$h_{fb,t}\, f_b(t-1) = \frac{1}{1-\beta} h_{\bar{\pi},t-1}\, \bar{\pi}_{t-2}$$

In two unknowns. Mm... You'd get a proliferation of unknowns...! I'll stop here, this wasn't for endog states, merde!

① GMM of midsimple

19 March 2020

② Numerical implementation of target criterion

① GMM of midsimple

sim_learnLH.m w/ PLM = constant_only

→ Need to:

1) add smooth function as a third criterion

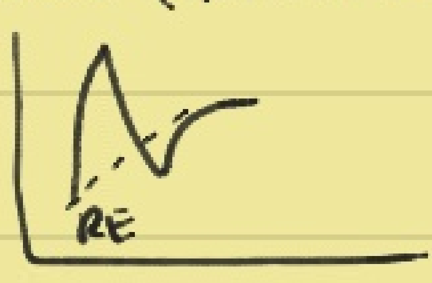
2) To have "midsimple", we need that x (and i) aren't learned, so input $a = \begin{bmatrix} a_1 \\ 0 \\ 0 \end{bmatrix}$, $b = b^{RE}$

↳ ok, I've learned one thing (comparing w/ IRFs of materials 9)

⇒ when only π is learned, IRFs are less oscillatory in the sense that they overshoot only once:

vs.

when $(\pi, x, i)'$ learned



→ did I ever conclude this? I thought so, but now it doesn't seem like it. In materials12 I just state that vector vs. scalar learning is pretty much the same

...And it almost is, just not quite.

⇒ so mental note!

Learning only LOM(π) dampens the oscillations b/c $E(\cdot)$ in NKIS & NKPC are moving less, and therefore $E(i_{t+k})$ is moving less. I think more importantly since the bulk of action is in x, the fact that $FE(x)$ is now cut out gets rid of the FE that was most oscillatory.

Ok I just wasted an hour trying to customize keys in Matlab...

Back to 1) the anchoring function:

$$k_t^{-1} - k_{t-1}^{-1} = c + d(FE)$$

has the interpretation that the gain decreases when $FE < 0$. Ryan is right: the easiest thing that makes Δk_t^{-1} big when FE is big in absolute value is FE^2

$$\rightarrow \quad k_t^{-1} - k_{t-1}^{-1} = c + d\, FE_t^2$$

↑ does this even make sense?

> Not really: it's saying that the gain always changes, even if the $FE^2 = 0$.

↳ So then gather data and implement

- HP
- BK
- Hamilton filters

But there is a problem: for this current form

$$k_t^{-1} - k_{t-1}^{-1} = c + d \cdot FE^2,$$

y always explodes, even more so if $c=0$.
And it's b/c $k \to 0$, so I guess k^{-1} explodes.

↳ Need to be smart about this!

Normally, a dgain would be <u>20 March 2020</u>

$k_{t+1} = k_t + 1$

So I could just do $k_t = k_{t-1} + \frac{1}{FE^2}$

so that if in the limit $FE^2 \to \infty$, $k_t = k_{t-1}$ (cgain)

Then one could have

$$k_t = k_{t-1} + d \frac{1}{FE^2}$$ in this case I guess d small $d<1$

or $$k_t = k_{t-1} + \frac{1}{d\,FE^2}$$ in this case I guess d big $d>1$

↳ both of them work and exactly the way I thought

or even $k_t = k_{t-1} + \frac{1}{(dFE)^2} = k_{t-1} + (dFE)^{-2}$

works.

HP-filter

A time series y_t (in logs) is

$$y_t = \underbrace{g_t}_{\text{growth component}} + \underbrace{c_t}_{\text{cyclical component}}$$

Obtain g_t as

$$\min_{\{g_t\}_{t=-1}^{T}} \left\{ \sum_{t=1}^{T} c_t^2 + \lambda \sum_{t=1}^{T} \left[(g_t - g_{t-1}) - (g_{t-1} - g_{t-2}) \right]^2 \right\}$$

where $\lambda = 1600$.

Hamilton provides a closed-form solution to this problem as $g^* = (H'H + \lambda Q'Q)^{-1} H' y$ (2)

where $\tilde{T} := T + 2$

$\underset{T\times 1}{y} := (y_T, y_{T-1}, \ldots, y_1)'$ $\underset{\tilde{T}\times 1}{g} = (g_T, g_{T-1}, \ldots, g_0, g_{-1})'$

and $\underset{T\times\tilde{T}}{H} = \begin{bmatrix} \underset{T\times T}{I_T} & \underset{T\times 2}{0} \end{bmatrix}$ and $\underset{T\times\tilde{T}}{Q} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0\ 1 & -2 & 1 \end{bmatrix}$

Hamilton filter

21 March 2020

Idea: $c = y_{t+h} - E[y_{t+h} | y_t]$

choose $h=8$ for quarterly data.

How to calc $E[y_{t+h} | y_t]$?

just regress y_{t+h} on $y_t, y_{t-1}, y_{t-2}, y_{t-3}$

and take the fitted value from that reg.

In fact, Hamilton's eq (21) gives us c directly:

$$c = \hat{v}_{t+h} = y_{t+h} - \hat{\beta}_0 - \hat{\beta}_1 y_t - \hat{\beta}_2 y_{t-1} - \hat{\beta}_3 y_{t-2} - \hat{\beta}_4 y_{t-3}$$

So: $Y = y_{t+h}$, $X = \overset{4\times 1}{\begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \end{bmatrix}}$ $\Rightarrow \underset{1\times 4}{\beta} = (\underset{1\times 4}{X'}\underset{4\times 1}{X})^{-1}\underset{1\times 4}{X'}\underset{1\times 1}{Y}$

1x1

and $\hat{v}_{t+h} = y_{t+h} - \beta X$

Baxter & King filter (BK filter, bandpass filter)

Def: business cycle = cyclical components of a ts that last between 6 and 32 quarters in duration

Notation: $BP_K(p, q)$ = filter that passes cycles between $q = 32$ and $p = 6$ cycle length, and is truncated at leads/lags K ($\Rightarrow$ recommend $K = 12$)

$\hookrightarrow$ Such a filter is a MA(K).

So we're interested in constructing the filtered series

$$y_t^* = \sum_{k=-K}^{K} a_k \, y_{t-k} \qquad (1)$$

and the question is how to obtain a_k?

(where the MA is symmetric, i.e. $a_k = a_{-k}$)

Eq. (10) gives the answer:

$$a_h^{BP} = (\bar{b}_h - \underline{b}_h) + (\bar{\theta} - \underline{\theta})$$

where the bars denote things pertaining to low-pass filters that only allow stuff below the frequency $\bar{\omega}$ & $\underline{\omega}$, i.e periodicities $\bar{p}$ & $\underline{p}$,

where $p = \frac{2\pi}{\omega}$

θ is a normalizing constant (coming from truncation)

$$\theta = \left(1 - \sum_{h=-k}^{k} b_h\right) \frac{1}{2k+1}$$

and $b_0 = \frac{\omega}{\pi}$ and $b_h = \frac{\sin(h\underline{\omega})}{h\pi}$ for $h = 1, 2, \ldots$ (7)

Ok, so we have $\underline{p} = 32$, $\bar{p} = 6$ ($q = 32$ $p = 6$)

$$\Rightarrow \underline{\omega} = \frac{2\pi}{\underline{p}} = \frac{2 \cdot \pi}{32} = \frac{\pi}{16} , \quad \bar{\omega} = \frac{2\pi}{6} = \frac{\pi}{3}$$

Autocovariances of VARs (Lütkepohl, p. 21)

For a stationary VAR(1)

$$y_t = \nu + A_1 y_{t-1} + u_t$$
$$\hookrightarrow N(0, \Sigma_u)$$

define $\Gamma_y(h) \equiv E[(y_t - \mu)(y_{t-h} - \mu)']$

where $\mu := E[y_t]$. Then if you know $\Gamma_y(0)$,

you can compute

$$\Gamma_y(h) = A_1 \Gamma_y(h-1) \quad \text{recursively!} \qquad (2.1.31)$$

If you know A_1 and Σ_u, you can compute

$$\text{vec}\,\Gamma_y(0) = (I_{k^2} - A_1 \otimes A_1)^{-1}\,\text{vec}\,\Sigma_u \qquad (2.1.32)$$

But I don't want to est a VAR... I just want the empirical auto-covariance matrix of the data and of the simulated data.

So, the GMM of the anchoring fit works 23 March 2020

so far. Issues.

1.) I'm not sure if I should model the raw data w/ a time series. So e.g. if they were a VAR(1) then I could estimate a RF-VAR, I could bootstrap using the residuals and I could estimate the autocov structure as in (2.1.32) & (2.1.31) in Lütkepohl.

This issue shows up for $W = \begin{bmatrix} \hat{\sigma}^2_{acf_1} & & \\ & \ddots & \\ & & \hat{\sigma}^2_{acf_{15}} \end{bmatrix}$

b/c $\hat{\sigma}^2_{acf,i}$ are super small, thus W^{-1} is huge, and therefore the estimation does not move.

2) Filters may not be working?

→ I think now they are!

↳ check bootstrap! ✓ It's fine.

Write up materials22_peter. ✓

Read "Stochastic". On it!

I'm not sure what I'm learning from <u>24 March 2020</u>
the GMM so far, so until I talk to Ryan, I'll postpone it and turn to questions of implementation of the target criterion.

<u>Peter meeting</u> <u>24 March 2020</u>

- Ask Clough

- Intro: CBs always talk about anchoring, eq (2) formalizes that!

- Where to go #1:

- Suggestions

  (1): A common theme: does it make sense to have $h_t = k_t^{-1}$ and just track the evolution of that.

(2): Eq. (2)

If you work w/ nonlinear diff eqs, they are harder but not imposs to solve

→ mit-shock business is all about that w/ perfect foresight

- take FOCs
- keep 'em in nonlinear form
- ass. the econ is hit by 1-time shock
- trace out transition to new st. st.

A precise way of putting it:

any dynamic econ: all vars → X_t

$$F(X_t, X_{t-1}, \varepsilon_t) = 0$$

If F has a lin form, $AX_t + BX_{t+1} + C\varepsilon_t = 0$

then we can solve it numerically quite easily.

If F is nonlinear, but nonstochastic:

$$F(x_t, x_{t-1}) = 0$$

is harder, but doable.

↳ Sect 2. point (1)

"can use analytic methods to describe eqs in opt policy & key variables"

& simulate vars under those policies.

⇒ allows you to answer

1) How close does a TR come to opt. policy

2) Opt TR will try to do the same thing as opt policy and you'll see what they both are trying to do.

Suggestion overall:

- simulate numerically the model under some kind of optimal pol that's more flex than a TR.

2nd suggestion w/ nonlinearity:

· Analogy: stochastic Ramsey model (neoclassical growth model)

no closed-form sol once depreciation isn't $=1$ and not log u. So to solve this:

- loglin
- discretize the state-space
  - TFP follows not an AR1 but one of 3 values
  - then decision rules are also 3-form

<u>1option</u> continue working w/ full-blown nonlinear model but replace the stoch. disturbances

$$u = \begin{matrix} u^{high} \\ u^{low} \end{matrix} \qquad r^n = \begin{matrix} r^{high} \\ r^{low} \end{matrix}$$

$\Rightarrow$ 4 states of the world

$\hookrightarrow$ calculating $E(\cdot)$ of ∞ future terms in (2) may be easier

2option: Take (1)

$$f_t = f_{t-1} + h_t (\pi_t - f_{t-1})$$

and loglin

Another way to solve a DSGE model

loglin ass's that decision rules are lin functions of the states

-> well you can say that they are quadratic function of the states

↳ so instead of Taylor-approx, maybe the right thing is a cubic or spline

In eq (3) I try to approx something of unknown form w/ a class

Simulate the model under some optimal targeting rule & compare w/ behavior under an optimized TR economy.

Draft for end of semester
→ she'll read it then
and afterwards too.

↳ Talk to Ryan now:
What is the most promising avenue?

Tomorrow noon → send Peter email for macro seminar

Work after

Takeaways:

1) In my introduction, I talk about how CB-ers often talk about anchoring. The target criterion (eq. (2) in materials22 peter) formalizes this!

2) Small detail: replace $h_t \equiv k_t^{-1}$

3.) The main point: need to identify the set of results (and avenues) that can be achieved and then written up. For him it seems to be that this set of results is the target criterion and its numerical implementation so we can make statements about the optimal plan and how it compares to the optimal Taylor rule.

↳ he seemed adamant that I settle this w/ Ryan.

4) Taking the analogy of the neoclassical growth model, no closed-form sol exists for the nonlinear model unless log u and $\beta = 1$. So if we wanted a tighter analytical result, we could 1) either loglin, or 2) discretize the stochastic disturbance

$$\text{e.g. } u_t = \begin{cases} u^{high} \\ u^{low} \end{cases}$$

↳ this way we might more fully describe the target criterion.

5) As for the loglin, which is a Taylor-Approx, maybe a higher order approx (cubic or spline) would be sensible.

Here's the deal: I don't think I wanna pursue further analytical work. Not w/ the discretization, not w/ a higher-order approx, although both might work. I think that building on the existing

work, I want to go on to simulating the optimal Ramsey policy and comparing it to the opt. TR.
⇒ I think that that is 1) feasible 2) makes sense in terms of storytelling & also is likely to provide additional inputs to that storytelling.

I will continue reading Stochastic Opt. in Continuous Time, not b/c I really hope for inputs for more analytical work, more for my intellectual completion – as diff. eqs contributed too.

50: numerical implementation of the target criterion <u>25 March 2020</u>

- I wonder if I couldn't do it via value function iteration → can I not solve the nonlinear system numerically and obtain the optimal paths like that?
→ which would allow me to back out i^{Ramsey} from NKIS?

g_π :

$y \quad k_t = k_{t-1} + \underbrace{(d \cdot fe_{t-1})^{-2}}_{=g}$

$$k_t = k_{t-1} + d^{-2} (\pi_t - \bar{\pi}_{t-1} - b_1 s_{t-1})^{-2}$$

Then $g_\pi(t) = d^{-2}(-2)(\pi_t - \bar{\pi}_{t-1} - b_1 s_{t-1})^{-3}$

$$= -2 (d\, fe_{t-1})^{-2} fe_{t-1}^{-1}$$

Now what I need to solve is the fact that when $\{i_t\}$ is exog, the model sol y = function (matrices, f_a, f_b) changes. In particular, if we go back to the model summary in materials5 in (1) – (3) we know that (3) doesn't hold (TR), and $\{i_t\}$ influences the econ via the NKIS curve (1).

$$x_t = -\sigma i_t + [\sigma, (1-\beta), -\sigma\beta] f_b + \sigma [1,0,0] (I_{nx} - \beta h x)^{-1} s_t$$

$$\pi_t = \kappa x_t + [(1-\alpha)\beta, \kappa\alpha\beta, 0] f_a + [0,0,1] (I_{nx} - \beta h x)^{-1} s_t$$

$$\begin{bmatrix} 0 & 1 \\ -\kappa & 1 \end{bmatrix} \begin{bmatrix} \pi_t \\ x_t \end{bmatrix} = \begin{bmatrix} -\sigma i_t + [\sigma, (1-\beta), -\sigma\beta] f_b + \sigma [1,0,0] (I_{nx} - \beta h x)^{-1} s_t \\ [(1-\alpha)\beta, \kappa\alpha\beta, 0] f_a + [0,0,1] (I_{nx} - \beta h x)^{-1} s_t \end{bmatrix}$$

So

$$\begin{bmatrix} \pi_t \\ x_t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\kappa & 1 \end{bmatrix}^{-1} \begin{bmatrix} -\sigma i_t + [\sigma, (1-\beta), -\sigma\beta] f_b + \sigma[1, 0, 0](I_{nx} - \beta h_x)^{-1} s_t \\ [(1-\alpha)\beta, \kappa\alpha\beta, 0] f_a + [0, 0, 1](I_{nx} - \beta h_x)^{-1} s_t \end{bmatrix}$$

where f_b, f_a unchanged.

<u>Ryan meeting</u> <u>25 March 2020</u>

$$k_t = k_{t-1} + \frac{1}{(d\,fe)^2} - c$$

so that when fe large, the gain grows

$E(fe)^{RE}$ would be the size of c.

- check if drawing w/ replacement
  - take raw data: run a VAR
  - bootstrap from residuals

- put sim data in filter

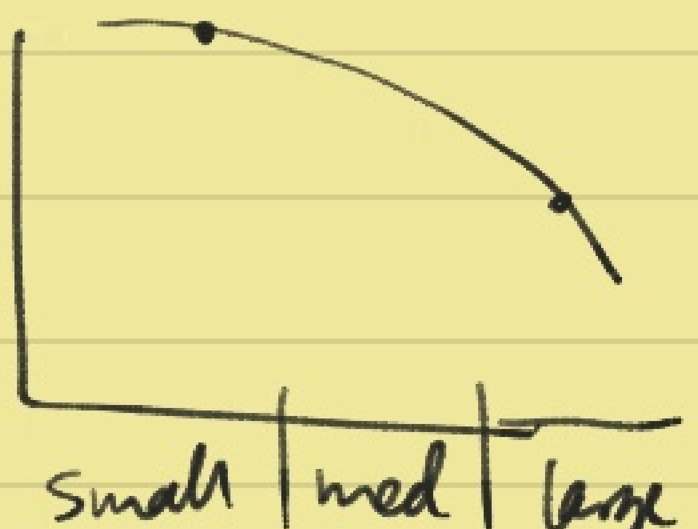
· g should
  - tend to go down
  - but when it goes up, it goes up a lot
    i.e. for very large FE, it ↑ a lot.

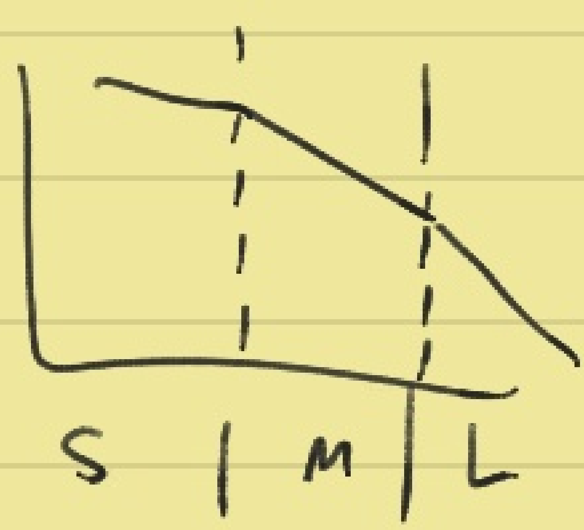
· check if you're close to the target moments

$k_t = k_{t-1} + \hat{g}(fe^c)$

$\hat{g}$ is a spline w/ x nodes

piecewise linear w/ x nodes





Each node comes w/ 2 params
→ so can est a 4-param family

↳ here you just est 3 slopes

→ eventually you could even est. what a small, large, FE is.

shape-preserving splines: are either always convex or always concave

Judd, the textbook, Empirical Methods

piece-wise linear interpolation
spline / quadratic spline
basis functions

Econ librarian: Sonia Ensins

Hamilton $\hat{d}$ lower b/c Var ↑

4 states, so VFI could be done.