Work after

① proj. facility invoked

i Projection facility isn't invoked for constant-only (in fact, it's not implemented for that!)

ii For slope and constant learning, it always is both for CEMP & CUSUM criterion, and cgain and dgain. Darn!

iii ↳ the above was for proj. facility w/ eig(R). (the "mini projection facility")

If you comment it out, then it turns out that all is fine! dgain, cgain, CEMP, CUSUM learning slope & constant.

⇒ So go on living life w/o a projection facility for now, and conclude that eig(R) isn't the way to go.

↳ I'll leave the proj. facility issue for now and return to it when I have to.

② det(·)

Sims 2003 p. 671, eq (8) says:

Let $\underset{n\times 1}{X} \sim$ multivariate $N(\mu, \Omega)$. Then

$$\text{entropy}(X) = H(X) = \frac{n}{2}\log_2(2\pi e) + \frac{1}{2}\log_2|\Omega|$$

↳ the only specific thing to X is $|\Omega| = \det(VC_{\text{matrix}})$

=> the determinant is a good summary of the info a matrix carries.

crit-CUSUM with $\det[\omega^{-1} ff' - \theta_{t-1}]$

-> much more anchored. ↑ this doesn't make sense b/c θ scalar

also $\det[\omega^{-1} ff'] - \theta_{t-1}$

-> much more anchored

ALWAYS anchored! b/c $\det[\omega^{-1} ff']$ is tiny! 10^{-31}

(can't do det(·) for CEMP b/c $\phi - [F, G]$ isn't square!)

scalar CUSUM:

initially anchoring $\uparrow$ in ψ_π, but once $\psi_\pi \geq 2$, anchoring $\downarrow$ in $\psi_\pi \uparrow$.

$\sqrt{\text{cusum}}$ for scalar case: much more anchoring (needs much lower $\tilde{\theta} = 0.00005$)

↳ still has the same feature that as $\psi_\pi \uparrow$, anchoring $\uparrow$

but when ψ_π is very small ($\psi_\pi \approx 1.001$) anchoring > than when $\psi_\pi \approx 1.5$.

$\sqrt{\text{cusum}}$ for vector case: much more anchoring (also needs like $\tilde{\theta} = 0.00005$)

$\sqrt{\cdot}$ seems to make it less monotonous.

In general, as $\psi_\pi \uparrow$, anchoring $\uparrow$ but it's not monotonous. In particular, for small ψ_π (< 1.01) anchoring $\downarrow$ when $\psi_\pi \uparrow$

$\Rightarrow$ so the ^{2} seems to add monotonicity: w/o it,

$\Rightarrow$ I need to understand the Brown-Durbin thing better!

I wanna go back to the Woodford thing <u>30 Jan 2020</u>

the "oni" := optimal noninertial plan.

So we have $\mathcal{L}^{stab} = \mathcal{L}_\pi + \lambda \mathcal{L}_x$

and we want $(f_\pi^{oni}, f_x^{oni}) = \arg\min \mathcal{L}^{stab}$ s.t. (III) & (IV).

Let's focus on $\mathcal{L}_\pi$:

$$\mathcal{L}_\pi = \frac{1}{1-\beta} f_\pi (I_{nx} - h_x h_x')^{-1} I_{nx} f_\pi' - f_\pi (I_{nx} - h_x h_x')^{-1} (I_{nx} - \beta h_x h_x')^{-1} (h_x h_x') f_\pi'$$

This seems to consist of parts w/ the following structure:

$a' x a$ where $f_\pi = a'$, x = longer uglier stuff (matrix)

and from my metrics notes (Econometrics SUM Part 1, p. 2)

$$\frac{d(a' x a)}{da} = (x + x') a$$

so that suggests that

$$\frac{\partial \mathcal{L}_\pi}{\partial f_\pi} = \frac{1}{1-\beta} \left[\left((I_{nx} - h_x h_x')^{-1} I_{nx} \right) + \left((I_{nx} - h_x h_x')^{-1} I_{nx} \right)' \right] f_\pi'$$

$$- \left[\left((I_{nx} - h_x h_x')^{-1} (I_{nx} - \beta h_x h_x')^{-1} (h_x h_x') \right) + \left((I_{nx} - h_x h_x')^{-1} (I_{nx} - \beta h_x h_x')^{-1} (h_x h_x') \right)' \right] f_\pi'$$

$$= \underset{3\times 3}{A_p} \cdot \underset{3\times 1}{f_\pi'} \quad \text{where}$$

$$A_p = \frac{1}{1-\beta} \left[\left((I_{nx} - h_x h_x')^{-1} I_{nx} \right) + \left((I_{nx} - h_x h_x')^{-1} I_{nx} \right)' \right]$$

$$- \left[\left((I_{nx} - h_x h_x')^{-1} (I_{nx} - \beta h_x h_x')^{-1} (h_x h_x') \right) + \left((I_{nx} - h_x h_x')^{-1} (I_{nx} - \beta h_x h_x')^{-1} (h_x h_x') \right)' \right]$$

Since $\mathcal{L}_x$ is symmetric to $\mathcal{L}_\pi$,

$$\frac{\partial \mathcal{L}_x}{\partial f_x} = A_p \cdot f_x'$$

$$\Rightarrow \frac{\partial \mathcal{L}^{stat}}{\partial f_\pi} = A_p \cdot f_\pi' \quad \frac{\partial \mathcal{L}^{stat}}{\partial f_x} = \lambda \cdot A_p \cdot f_x' .$$

I'm confused though b/c we now have a weird system of equations:

$$\left.\begin{array}{l} A_p \cdot f_\pi' = 0 \\ \lambda A_p \cdot f_x' = 0 \end{array}\right\}$$ This can only be if $A_p = 0$ or $f_\pi' = f_x' = 0$ or unconstrained since $\lambda = 0$.

and (III) & (IV) $\rightarrow$ but then you only solve (III) & (IV)?

I'm also surprised b/c why do we have 4 eqs in 2 unknowns?

To clear this up, let's go back to Woodford's simple example on p. 511

$$(f_\pi^{opt}, f_x^{opt}) = \operatorname{argmin} (f_\pi^2 + \lambda f_x^2) \quad \text{s.t. } (1-\beta\rho) f_\pi = k f_x + 1$$

$$f_\pi = \frac{k}{1-\beta\rho} f_x + \frac{1}{1-\beta\rho}$$

$$\rightarrow f_x^{\text{oni}} = \operatorname{argmin} \left(\frac{k}{1-\beta\rho} f_x + \frac{1}{1-\beta\rho}\right)^2 + \lambda f_x^2$$

$$= \left(\frac{k}{1-\beta\rho}\right)^2 f_x^2 + \frac{2k}{(1-\beta\rho)^2} f_x + \left(\frac{1}{1-\beta\rho}\right)^2 + \lambda f_x^2$$

FOC $\left[\left(\frac{k}{1-\beta\rho}\right)^2 + \lambda\right] f_x + \frac{2k}{(1-\beta\rho)^2} = 0$

$$f_x = \frac{-\frac{2k}{(1-\beta\rho)^2}}{\left(\frac{k}{1-\beta\rho}\right)^2 + \lambda} = - \frac{\frac{2k}{\cancel{(1-\beta\rho)^2}}}{\frac{k^2 + \lambda(1-\beta\rho)^2}{\cancel{(1-\beta\rho)^2}}}$$

$$= - \frac{2k\lambda^{-1}}{k^2\lambda^{-1} + (1-\beta\rho)^2}$$ which is almost = (3.5).

Oh ok, so taking $\frac{\partial \mathcal{L}^{\text{stat}}}{\partial f_i}$ doesn't make sense of course: we either need the Lagrangian or we need to sub in.

What is still confusing though is that in Woodford's ex, there's only the NKPC and therefore you only have 1 constraint. I have two, however: (III) & (IV).

Ok, I see: Woodford writes an "$\mathcal{L}^{stab}$" conditional on each shock separately. And he doesn't consider shocks to the TR.

Let's work thru this example.

Shocks:

$$r_t^n = (1-\rho_r)\bar{r} + \rho_r r_{t-1}^n + \varepsilon_t^r \qquad (2.27)$$

$$u_t = \rho_u u_{t-1} + \varepsilon_t^u \qquad (2.18)$$

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + u_t \qquad (2.1)$$

$$x_t = E_t x_{t+1} - \sigma [i_t - E_t \pi_{t+1} - r_t^n] \qquad (2.23)$$

Wait – I have two (maybe related) issues:

1.) What's the diff. between optimal policy and oni?
Woodford seems to suggest that oni is a restricted set of optimal policies that are purely forward-looking.

2) Does the presence of an NKIS relation necessitate an i-term in the CB's loss?

Ok – listen: these are the equations of the model:

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + u_t \qquad (2.1)$$

$$x_t = E_t x_{t+1} - \sigma\left[i_t - E_t \pi_{t+1} - r_t^n\right] \qquad (2.23)$$

You can sub in the TR and solve for $z = \bar{z} + f_z s_t$ w/ $z = \begin{bmatrix} \pi \\ x \end{bmatrix}$ or don't sub it in & solve for $\overbrace{\text{—"—}}$ w/ $z = \begin{bmatrix} \pi \\ x \\ i \end{bmatrix}$, it doesn't matter. Also it doesn't matter if you write $z = \bar{z} + \overset{3\times 3}{f_z} \overset{3\times 1}{s_t}$ or

$$\underset{n_y \times 1}{z} = \bar{z} + \underset{n_y \times 1}{f} u_t + \underset{n_y \times 1}{g} r_t^n \left(+ \underset{n_y \times 1}{h} \bar{i}_t\right)$$

and lastly it doesn't change anything conceptually whether there's a mon. pol. shock or not.

Very lastly: no, I don't think that the presence of an NKIS relation implies that the CB loss includes an i-term. Why should it? Does an IS-relation imply a concern for i-stabilization? I don't think a priori! Let's plug in the TR and see if the # eqs is fine!

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + u_t$$

$$x_t = E_t x_{t+1} - \sigma [\psi_\pi \pi_t + \psi_x x_t + \bar{i}_t - E_t \pi_{t+1} - r_t^n]$$

$$\Leftrightarrow \quad \pi_t - \kappa x_t = \beta E_t \pi_{t+1} + u_t \qquad (1)$$

$$\sigma\psi_\pi \pi_t + (1+\sigma\psi_x) x_t = E_t x_{t+1} + \sigma E_t \pi_{t+1} + \sigma r_t^n - \sigma \bar{i}_t \qquad (2)$$

Step 1 Postulate $\pi_t = \bar{\pi} + f_\pi r_t^n + g_\pi \bar{i}_t + h_\pi u_t$ (3)

$x_t = \bar{x} + f_x r_t^n + g_x \bar{i}_t + h_x u_t$ (4)

Step 2 This will give me 2 eqs in $(\bar{\pi}, \bar{x})$ which will be redundant.

And it will give me 2 constraints in the 6 unknowns, $(f_\pi, f_x, g_\pi, g_x, h_\pi, h_x)$.

If $\mathcal{L}^{stab} = f(var(\pi), var(x))$ only, then

$\mathcal{L}^{stab} = f(f_\pi, g_\pi, h_\pi, f_x, g_x, h_x)$

w/ 2 eqs 2 shocks

w/ 2 eqs 1 shock

w/ 1 eq 2 shocks

Ok, so let's do some dumb accounting: (Woodford:) (4) (2)

The number of unknowns is $n_y \times n_e = 2 \times 3 = 6$. (2)

The number of constraints is #(eqs) = 2. (1) (1) (2) (2) (4)

The number of FOCs from $\mathcal{L}^{stat}$ once you've subbed in the constraints is 4. (1) (3) (0) (2)

$\Rightarrow$ Let n_{eqs} := # equations (Addendum 31st Jan)

the number of equations to be solved in the end is the # FOCs once you've subbed in the constraints, i.e.

$n_{unknowns} - n_{constraints} = n_y \cdot n_e - n_{eqs} \cdot n_e$

With the NKIS, we have a sitn in which $n_y = n_{eqs}$, whereas Woodford had $n_{eqs} = n_y - 1$ so he got

FOCs you'll use: $(n_y - n_{eqs}) \cdot n_e \Rightarrow$ when $n_y = n_{eqs}$, the constraints fully determine the sol.

$$= n_y \cdot n_e - (n_y - 1)$$
$$= n_y (n_e - 1) + 1$$
$$= 2(1-1) + 1 = 1 \text{ eq left.}$$

W/ the NKIS, $n_{eqs} = n_y$ so $= n_y \cdot n_e - n_y$

$$= n_y (n_e - 1)$$

If $n_e = 1$, this $= 0$!

which shows that when $n_{eqs} = n_y$, we need at least 2 shocks in order not to have the constraints determine the coefficients to disturbances.

If $n_{eqs} > n_y$, e.g. $n_{eqs} = n_y + 1$ $\quad = n_y \cdot n_e - n_y - 1$

$$= n_y(n_e - 1) - 1$$

$\rightarrow$ then we need $n_e \geq 2$ to <u>have a solution at all</u>!

What's the economics behind this?

Supp. that the case where $n_{eqs} > n_y$ is not relevant

$\rightarrow$ an econ. model wouldn't give this.

But the case where $n_{eqs} = n_y$ is the usual case, which is why I was having trouble w/ Woodford's example where $n_{eqs} < n_y$.

But maybe this is the reason why Woodford adds r_t^n when he adds a second model equation, the NKIS, so now $n_{eqs} = 2 = n_y$, and $n_e \uparrow$ from 1 to 2.

Notice that if you don't sub in the TR, nothing changes b/c $n_{eqs} = 3 = n_y$ and you still need $n_e \geq 2$ for the $\mathcal{L}$ stat FOCs to matter.

So what's the intuition behind the statement that "when $n_{eqs} = n_y$, the model needs at least 2 disturbances for the ω-coefficients on disturbances not to be solely determined by constraints"?

These constraints say that in the LR, the econ shouldn't respond to shocks. If there's only one shock, then this restriction is sufficient to pin down state-contingent responses of σ_i to the shock. If there are 2 or more shocks, then an additional condition about minimizing variances is required (L^{stab} plays a role).

I'm wondering if it can (and does) happen that constraints are redundant? If it does, then maybe the conclusion will be that you need as many shocks as equations for any variance-minimization to be required.

So the intuition, roughly, seems to be this: if there are only a few sources of shocks in the model, then the requirement of a deterministic path for endog. variables is sufficient to pin down optimal policy responses. For more disturbances, variance-minimizing considerations are required in addition.

Ok so go back to model.

RE
$$\begin{cases} \pi_t - \kappa x_t = \beta E_t \pi_{t+1} + u_t & (1) \\ \sigma \psi_\pi \pi_t + (1+\sigma\psi_x) x_t = E_t x_{t+1} + \sigma E_t \pi_{t+1} + \sigma r_t^n - \sigma \bar{i}_t & (2) \end{cases}$$

Learning:

$$\pi_t - \kappa x_t = \hat{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \left\{ \kappa\alpha\beta x_{T+1} + (1-\alpha)\beta \pi_{T+1} + u_T \right\} \quad (3)$$

$$\sigma\psi_\pi \pi_t + (1+\sigma\psi_x) x_t = \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \left\{ (1-\beta-\sigma\beta\psi_x) x_{T+1} + \sigma(1-\beta\psi_\pi)\pi_{T+1} + \sigma r_T^n - \sigma \bar{i}_T \right\} \quad (4)$$

Step 1 Postulate

$$\pi_t = \bar{\pi} + f_\pi r_t^n + g_\pi \bar{i}_t + h_\pi u_t \quad (5)$$

$$x_t = \bar{x} + f_x r_t^n + g_x \bar{i}_t + h_x u_t \quad (6)$$

Step 2 Sub in (5) & (6) into (1) & (2). This will give the same eqs for the deterministic part $(\bar{\pi}, \bar{x})$ in the RE & learning model, but constraints on $(f_\pi, g_\pi, h_\pi, f_x, g_x, h_x)$ will be different.

Step 3. Derive $\mathcal{L}^{stab} = \sum_{T=t}^{\infty} \beta^{T-t} \left[var_t(\pi_t) + \lambda var_t(x_t) \right]$ given (5) & (6) will also yield the same $\mathcal{L}^{stab}$ for the RE & learning models.

Let's do RE in blue!

(1) $\bar{\pi} + f_\pi r_t^n + g_\pi \bar{i}_t + h_\pi u_t - \kappa\left(\bar{x} + f_x r_t^n + g_x \bar{i}_t + h_x u_t\right)$

$= \beta\left(\bar{\pi} + f_\pi r_{t+1}^n + g_\pi \bar{i}_{t+1} + h_\pi u_{t+1}\right) + u_t$

$\Leftrightarrow$

$\bar{\pi} - \kappa\bar{x} - \beta\bar{\pi} = -f_\pi r_t^n - g_\pi \bar{i}_t - h_\pi u_t - \kappa f_x r_t^n - \kappa g_x \bar{i}_t - \kappa h_x u_t$

$+ \beta\left(f_\pi \rho_r r_t^n + g_\pi \rho_i \bar{i}_t + h_\pi \rho_u \cdot u_t\right) + u_t$

So, ignoring the deterministic stuff, condition (I) will be $RHS \overset{!}{=} 0$,

i.e.

$\underbrace{(-f_\pi - \kappa f_x + \beta f_\pi \rho_r)}_{\overset{!}{=}0} r_t^n + \underbrace{(-g_\pi - \kappa g_x + \beta g_\pi \rho_i)}_{\overset{!}{=}0} \bar{i}_t + \underbrace{(-h_\pi - \kappa h_x + \beta h_\pi \rho_u + 1)}_{\overset{!}{=}0} u_t \overset{!}{=} 0$

and I already see that I was wrong b/c this gives me 3 instead of 1 constraint!

$\rightarrow$ So that means that # constraints $= n_{eq} \times n_e = 6$

So yes, if $n_{eq} = n_y$, then the constraints will fully determine the solution. $\Rightarrow$ won't even need to set up $\mathcal{L}^{stab}$!

$(\beta\rho_r - 1) f_\pi = \kappa f_x$ (I)

$(\beta\rho_i - 1) g_\pi = \kappa g_x$ (II)

$1 + (\beta\rho_u - 1) h_\pi = \kappa h_x$ (III)

Peter meeting 30 Jan 2020

CUSUM: the matrix to scalar

Lütkepohl's book on VARs → "Intro to Multiple Time Series Analysis" or Hamilton's book

"Multivariable CUSUM" or "New Intro to -11-"

or "Structural Vector Autoregr" w/ Kilian

check:

residuals-based test (CUSUM) vs. alternative stat. tests for parameter- or model instability (e.g. Chow-test)

linear vs. nonlinear

→ proj. facility: check whether it's the jump matrix g_x they check!

The issue is this: checking eig(A) only makes sense for dynamic systems such as $X_{t+1} = A X_t$

→ then eig(A) will tell you about dynamics.

But for $Y_t = g_x X_t$, $eig(g_x)$ will only tell you about the scaling, or the units in which you're measuring the vars.

→ That's why the 1st step is to check EH (2001) & Graham to see whether they're really checking $eig(\phi)$, or what matrix's eig they're actually checking.

<u>Work after</u>

Projection facility issue

- <u>Graham</u> is checking $eig(\phi^s)$, but there, ϕ^s is $\hat{h}_x$ that agents are learning!
- <u>Evans & Honkapohja 2001</u> p. 36 says that the proj. facility is just an algorithm that restricts $\hat{\phi}$ to be in a neighborhood of ϕ^{RE} (i.e. of g_x). It's useful in case there are multiple eqba, so agents don't learn the wrong one.

- Marcet & Sargent 1989 also specify a projection facility as a set restricted to be "close enough" to another.
- Branch unpublished p. 5 $K(r)$ = closed ball of radius r around the REE. If $(\phi, R) \in K(c_2)$ ✓, else put $(\phi, R) \in K(c_1)$ for $0 < c_1 < c_2$

$\Rightarrow$ so the eig(.) thing was indeed a bluff!

overparameterization in EH (2001)

overfitting, like thinking that an AR(1) is an ARMA (p 213 middle).

See also p. 206 - 207 for this!

On p. 189 they discuss that dynamics/stability of the system depends on (how or if) the PLM is overparameterized.

$\Rightarrow$ so maybe overshooting is partly due to the fact that under cgain learn, (strong) E-stab. doesn't hold.

! Best discussion is on p. 41 !

In RE, the 2nd set of constraints comes from: **31 Jan 2020**

$$b\psi_\pi \pi_t + (1+b\psi_x) x_t = E_t x_{t+1} + b E_t \pi_{t+1} + b r_t^n - b \bar{i}_t \qquad (2)$$

Sub in the conjectured sol

$$b\psi_\pi(\bar{\pi} + f_\pi r_t^n + g_\pi \bar{i}_t + h_\pi u_t) + (1+b\psi_x)(\bar{x} + f_x r_t^n + g_x \bar{i}_t + h_x u_t)$$
$$= (\bar{x} + f_x r_{t+1}^n + g_x \bar{i}_{t+1} + h_x u_{t+1}) + b(\bar{\pi} + f_\pi r_{t+1}^n + g_\pi \bar{i}_{t+1} + h_\pi u_{t+1})$$
$$+ b r_t^n - b \bar{i}_t$$

$\Leftrightarrow$

$$b\psi_\pi(\bar{\pi} + f_\pi r_t^n + g_\pi \bar{i}_t + h_\pi u_t) + (1+b\psi_x)(\bar{x} + f_x r_t^n + g_x \bar{i}_t + h_x u_t)$$
$$= (\bar{x} + f_x \rho_r r_t^n + g_x \rho_i \bar{i}_t + h_x \rho_u u_t) + b(\bar{\pi} + f_\pi \rho_r r_t^n + g_\pi \rho_i \bar{i}_t + h_\pi \rho_u u_t)$$
$$+ b r_t^n - b \bar{i}_t$$

Take the deterministic stuff on LHS:

$$\bar{\pi}(b\psi_\pi - b) + \bar{x}(b\psi_x) = r_t^n\left[b\psi_\pi f_\pi + (1+b\psi_x) f_x + f_x \rho_r + b f_\pi \rho_r + b\right]$$
$$+ \bar{i}_t\left[b\psi_\pi g_\pi + (1+b\psi_x) g_x + g_x \rho_i + b g_\pi \rho_i - b\right]$$
$$+ u_t\left[b\psi_\pi h_\pi + (1+b\psi_x) h_x + h_x \rho_u + b h_\pi \rho_u\right]$$

$\Longrightarrow$

$$b(\psi_\pi + \rho_r) f_\pi + (1 + b\psi_x + \rho_r) f_x + b = 0 \qquad (\text{IV})$$

$$b(\psi_\pi + \rho_i) g_\pi + (1 + b\psi_x + \rho_i) g_x - b = 0 \qquad (\text{V})$$

$$b(\psi_\pi + \rho_u) h_\pi + (1 + b\psi_x + \rho_u) h_x = 0 \qquad (\text{VI})$$

So then RE amounts to solving:

$$(\beta \rho_r - 1) f_\pi = \kappa f_x \quad (\mathrm{I}) \qquad z(\psi_\pi + \rho_r) f_\pi + (1 + z\psi_x + \rho_r) f_x + z = 0 \quad (\mathrm{IV})$$

$$(\beta \rho_i - 1) g_\pi = \kappa g_x \quad (\mathrm{II}) \qquad z(\psi_\pi + \rho_i) g_\pi + (1 + z\psi_x + \rho_i) g_x - z = 0 \quad (\mathrm{V})$$

$$(\beta \rho_u - 1) h_\pi = \kappa h_x - 1 \quad (\mathrm{III}) \qquad z(\psi_\pi + \rho_u) h_\pi + (1 + z\psi_x + \rho_u) h_x = 0 \quad (\mathrm{VI})$$

the good news is that these eqs seem to come in pairs:

So for (f_π, f_x), solve (I) & IV

(I) $f_x = \frac{\beta \rho_r - 1}{\kappa} f_\pi \longrightarrow$ in (IV):

$$z(\psi_\pi + \rho_r) f_\pi + (1 + z\psi_x + \rho_r)\frac{(\beta \rho_r - 1)}{\kappa} f_\pi + z = 0$$

$$\Rightarrow \frac{z\kappa(\psi_\pi + \rho_r) + (1 + z\psi_x + \rho_r)(\beta \rho_r - 1)}{\kappa} f_\pi = -z$$

$$\Rightarrow f_\pi^{\,omi,RE} = -\frac{z\kappa}{z\kappa(\psi_\pi + \rho_r) + (1 + z\psi_x + \rho_r)(\beta \rho_r - 1)} \quad ①$$

$$f_x^{\,omi,RE} = -\left(\frac{\beta \rho_r - 1}{\kappa}\right) \frac{z\kappa}{z\kappa(\psi_\pi + \rho_r) + (1 + z\psi_x + \rho_r)(\beta \rho_r - 1)} \quad ②$$

Analogously, (II): $g_x = \frac{\beta \rho_i - 1}{\kappa} g_\pi \longrightarrow$ in (V)

$$z(\psi_\pi + \rho_i) g_\pi + (1 + z\psi_x + \rho_i)\frac{(\beta \rho_i - 1)}{\kappa} g_\pi = z$$

$$[z\kappa(\psi_\pi + \rho_i) + (1 + z\psi_x + \rho_i)(\beta \rho_i - 1)] g_\pi = \kappa z$$

$$g_\pi^{\,oni;RE} = \frac{\kappa b}{b\kappa(\psi_\pi + \rho_i) + (1 + b\psi_x + \rho_i)(\beta\rho_i - 1)} \qquad (3)$$

$$g_x^{\,oni;RE} = \left(\frac{\beta\rho_i - 1}{\kappa}\right) \cdot \frac{\kappa b}{b\kappa(\psi_\pi + \rho_i) + (1 + b\psi_x + \rho_i)(\beta\rho_i - 1)} \qquad (4)$$

And lastly, (III): $\frac{1}{\kappa} + \left(\frac{\beta\rho_u - 1}{\kappa}\right) h_\pi = h_x$

$$b(\psi_\pi + \rho_u) h_\pi + (1 + b\psi_x + \rho_u)\left[\frac{1}{\kappa} + \frac{\beta\rho_u - 1}{\kappa}\right] h_\pi = 0$$

$$\Leftrightarrow \left[b(\psi_\pi + \rho_u) + (1 + b\psi_x + \rho_u)\frac{\beta\rho_u - 1}{\kappa}\right] h_\pi = -\frac{(1 + b\psi_x + \rho_u)}{\kappa}$$

$$h_\pi^{\,oni;RE} = -\frac{(1 + b\psi_x + \rho_u)}{b\kappa(\psi_\pi + \rho_u) + (1 + b\psi_x + \rho_u)(\beta\rho_u - 1)} \qquad (5)$$

$$h_x^{\,oni;RE} = \frac{1}{\kappa} - \left(\frac{\beta\rho_u - 1}{\kappa}\right)\frac{(1 + b\psi_x + \rho_u)}{b\kappa(\psi_\pi + \rho_u) + (1 + b\psi_x + \rho_u)(\beta\rho_u - 1)} \qquad (6)$$

So that means that I have the oni for $(\pi, x)^{RE}$. I can obtain the oni for i_t^{RE} by subbing $i_t = \psi_\pi \pi^{oni} + \psi_x x^{oni} + \bar{i}_t$, to get

$$i_t^{oni} = \underbrace{(\psi_\pi f_\pi + \psi_x f_x)}_{f_i^{oni}} r_t^n + \underbrace{(\psi_\pi g_\pi + \psi_x g_x + 1)}_{g_i^{oni}} \bar{i}_t + \underbrace{(\psi_\pi h_\pi + \psi_x h_x)}_{h_i^{oni}} u_t$$

$$i_t^{oni} = f_i^{oni} r_t^n + g_i^{oni} \bar{i}_t + h_i^{oni} u_t$$

ok cool, but now what?

I'm doing something wrong. I can see on Woodford's oni, that there are 2 big diffs to mine:

1) λ_i $(i = x, i)$ are showing up, i.e. weights on output and interest rate gaps in the loss function

$\rightarrow$ $\mathcal{L}^{stat}$ <u>is</u> being used!

2) The Taylor-rule parameters (ψ_π, ψ_x) $(\rightarrow \phi_\pi, \phi_x)$ do <u>not</u> show up anywhere in the oni.

$\rightarrow$ the oni seems to be independent of the TR!

$\rightarrow$ I should work thru woodford's example in App F.4.

So back to Woodford's example p. 514 & App. F4.

He is saying that the objective is to find (ϕ_π^*, ϕ_x^*).

Model equations are

$$r_t^n = (1-\rho_r)\bar{r} + \rho_r r_{t-1}^n + \varepsilon_t^r \quad (2.27)$$

$$u_t = \rho_u u_{t-1} + \varepsilon_t^u \quad (2.18)$$

exog $n_e = 2$

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + u_t \quad (2.1)$$

$$x_t = E_t x_{t+1} - \sigma[i_t - E_t \pi_{t+1} - r_t^n] \quad (2.23)$$

endog $n_{eqs} = 2$

Conjectured LOM

$$y_t = \bar{y} + f_y u_t + g_y r_t^n \quad (3.6)$$

where $y_t = \begin{pmatrix} \pi \\ x \\ i \end{pmatrix} \rightarrow n_y = 3$

$\Rightarrow$ ok: so $(n_y - n_{eqs}) n_e = (3-2)\cdot 2 = 2$ FOCs

unknowns $= n_y \cdot n_e = 3 \cdot 2 = 6$

constraints $= n_{eqs} \cdot n_e = 2 \cdot 2 = 4$ $\rightarrow$ and that's what Woodford says too on p. 512: "2 restrictions on f_y and 2 on g_y"

Ok, fine, I give in, let's do it, taking $L^{stab,r}$ & $L^{stab,u}$ (3.7) & (3.8) (p. 513) as given!

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + u_t$$

$$\bar{\pi} + f_\pi u_t + g_\pi r_t^n = \kappa(\bar{x} + f_x u_t + g_x r_t^n)$$
$$+\beta(\bar{\pi} + f_\pi u_{t+1} + g_\pi r_{t+1}^n) + u_t$$

$$\bar{\pi} - \kappa\bar{x} - \beta\bar{\pi} = \underline{\kappa f_x u_t} + \kappa g_x r_t^n + \underline{\beta f_\pi \rho_u u_t} \; \underline{+ u_t}$$
$$+ \beta g_\pi [(1-\rho_r)\bar{r} + \rho_r r_t^n] \; \underline{- f_\pi u_t} - g_\pi r_t^n$$

RHS: $(\beta\rho_u f_\pi - f_\pi + \kappa f_x + 1) u_t$

$$+ [\beta\rho_r g_\pi - g_\pi + \kappa g_x] r_t^n + \underbrace{\beta g_\pi (1-\rho_r)\bar{r}}_{\text{WTF to do w/ this?}} \stackrel{!}{=} 0$$

$\Rightarrow$ I feel it's gonna be $\bar{i}$ somehow...

$$(1-\beta\rho_u) f_\pi = \kappa f_x + 1 \qquad (C1)$$
$$(1-\beta\rho_r) g_\pi = \kappa g_x \qquad (C2)$$

$$x_t = E_t x_{t+1} - \sigma[i_t - E_t \pi_{t+1} - r_t^n]$$

$$\cancel{\bar{x}} + f_x u_t + g_x r_t^n = \cancel{\bar{x}} + f_x u_{t+1} + g_x r_{t+1}^n - \sigma(\bar{i} + f_i u_t + g_i r_t^n)$$
$$+ \sigma[\bar{\pi} + f_\pi u_{t+1} + g_\pi r_{t+1}^n] + \sigma r_t^n$$

$$\sigma\bar{i} - \sigma\bar{\pi} = -f_x u_t - g_x r_t^n + f_x \rho_u u_t + g_x[(1-\rho_r)\bar{r} + \rho_r r_t^n]$$
$$- \sigma f_i u_t - \sigma g_i r_t^n + \sigma f_\pi \rho_u u_t + \sigma g_\pi [(1-\rho_r)\bar{r} + \rho_r r_t^n]$$
$$+ \sigma r_t^n$$

$$\beta \bar{i} - \beta \bar{\pi} = -f_x u_t - g_x r_t^n + f_x \rho_u u_t + g_x \left[(1-\rho_r)\bar{r} + \rho_r r_t^n \right]$$
$$- \beta f_i u_t - \beta g_i r_t^n + \beta f_\pi \rho_u u_t + \beta g_\pi \left[(1-\rho_r)\bar{r} + \rho_r r_t^n \right]$$
$$+ \beta r_t^n$$

$$\beta(\bar{i} - \bar{\pi}) = (-f_x + f_x \rho_u - \beta f_i + \beta f_\pi \rho_u) u_t$$
$$+ (-g_x + g_x \rho_r - \beta g_i + \beta g_\pi \rho_r + \beta) r_t^n$$
$$+ g_x (1-\rho_r)\bar{r} + \beta g_\pi (1-\rho_r)\bar{r}$$

$$\underbrace{\beta\left[\bar{i} - \bar{\pi} - g_\pi (1-\rho_r)\bar{r}\right] - g_x (1-\rho_r)\bar{r}}_{\text{Another WTF-term}} = RHS.$$

$$(\rho_u - 1) f_x - \beta f_i + \beta \rho_u f_\pi = 0 \qquad (C3)$$
$$(\rho_r - 1) g_x - \beta g_i + \beta \rho_r g_\pi + \beta = 0 \qquad (C4)$$

Ok just realized sthg. LHS $\stackrel{!}{=} 0$ and RHS $\stackrel{!}{=} 0$ like before, but we don't assume $\bar{y} = 0$ b/c for a concern for i-rate-stabilization, $(i_t - i^*)^2$ in $\mathcal{L}$, LR-values may not be zero, in fact, will be 0 only if $\lambda_i = 0$

But somehow that means:

$$(1-\beta)\bar{\pi} - \kappa\bar{x} \underline{- \beta g_\pi (1-\rho_r)\bar{r}} = 0$$

$$(1-\beta\rho_u) f_\pi = \kappa f_x + 1 \qquad (C1)$$

$$(1-\beta\rho_r) g_\pi = \kappa g_x \qquad (C2)$$

$$\sigma\left[\bar{i} - \bar{\pi} \underline{- g_\pi(1-\rho_r)\bar{r}}\right] \underline{- g_x(1-\rho_r)\bar{r}} = 0$$

$$(\rho_u - 1) f_x - \sigma f_i + \sigma\rho_u f_\pi = 0 \qquad (C3)$$

$$(\rho_r - 1) g_x - \sigma g_i + \sigma\rho_r g_\pi + \sigma = 0 \qquad (C4)$$

I don't know how to treat the underlined WTF-terms.

Ok: do the following: given these (Cs), try to derive f_π^{opt}

$$(C1): \quad f_x = \frac{(1-\beta\rho_u) f_\pi}{\kappa} - \frac{1}{\kappa}$$

(C3):

$$(\rho_u - 1)\frac{(1-\beta\rho_u)}{\kappa} f_\pi - \frac{(\rho_u - 1)}{\kappa} - \sigma f_i + \sigma\rho_u f_\pi = 0$$

$$\frac{\kappa\sigma\rho_u - (1-\rho_u)(1-\beta\rho_u)}{\kappa} f_\pi + \frac{1-\rho_u}{\kappa} = \frac{\sigma\kappa f_i}{\kappa}$$

$$f_i = \frac{(\kappa\sigma\rho_u - (1-\rho_u)(1-\beta\rho_u))}{\sigma\kappa} f_\pi + \frac{1-\rho_u}{\sigma\kappa}$$

Ok now plug these into $\mathcal{L}^{\text{stab},u} = f_\pi^2 + \lambda_x f_x^2 + \lambda_i f_i^2$

$$\mathcal{L}^{stab,u} = f_\pi^2 + \lambda_x \left(\frac{(1-\beta\rho_u)}{\kappa} f_\pi - \frac{1}{\kappa} \right)^2$$
$$+ \lambda_i \left(\frac{(\kappa\sigma\rho_u - (1-\rho_u)(1-\beta\rho_u))}{\sigma\kappa} f_\pi + \frac{1-\rho_u}{\sigma\kappa} \right)^2$$

Let's use Woodford's simplifying notation $\gamma_j := (1-\rho_j)(1-\beta\rho_j)$

$$\rightarrow \mathcal{L}^{stab,u} = f_\pi^2 + \lambda_x \left[\left(\frac{1-\beta\rho_u}{\kappa}\right)^2 f_\pi^2 + \left(\frac{1}{\kappa}\right)^2 - \frac{2}{\kappa} \frac{(1-\beta\rho_u)}{\kappa} f_\pi \right]$$

$$+ \lambda_i \left[\left(\frac{\kappa\sigma\rho_u - \gamma_u}{\sigma\kappa}\right)^2 f_\pi^2 + \left(\frac{1-\rho_u}{\sigma\kappa}\right)^2 + \frac{2(\kappa\sigma\rho_u - \gamma_u)(1-\rho_u)}{(\sigma\kappa)^2} f_\pi \right]$$

FOC for f_π:

$$2f_\pi + 2\lambda_x \left(\frac{1-\beta\rho_u}{\kappa}\right)^2 f_\pi - 2\lambda_x \frac{(1-\beta\rho_u)}{\kappa^2} + 2\lambda_i \left(\frac{\kappa\sigma\rho_u - \gamma_u}{\sigma\kappa}\right)^2 f_\pi$$

$$+ 2\lambda_i \frac{(\kappa\sigma\rho_u - \gamma_u)(1-\rho_u)}{(\sigma\kappa)^2} = 0$$

$\Leftrightarrow$

$$f_\pi \left[1 + \lambda_x \frac{(1-\beta\rho_u)^2}{\kappa^2} + \lambda_i \frac{(\kappa\sigma\rho_u - \gamma_u)^2}{\sigma^2\kappa^2} \right] - \lambda_x \frac{(1-\beta\rho_u)}{\kappa^2} + \lambda_i \frac{(\kappa\sigma\rho_u - \gamma_u)(1-\rho_u)}{\sigma^2\kappa^2} = 0$$

$$f_\pi \left[\sigma^2\kappa^2 + \lambda_x \sigma^2 (1-\beta\rho_u)^2 + \lambda_i (\kappa\sigma\rho_u - \gamma_u)^2 \right] = \lambda_x \sigma^2 (1-\beta\rho_u) - \lambda_i (\kappa\sigma\rho_u - \gamma_u)(1-\rho_u)$$

$$f_\pi = \frac{\lambda_x \sigma^2 (1-\beta\rho_u) - \lambda_i (\kappa\sigma\rho_u - \gamma_u)(1-\rho_u)}{\sigma^2\kappa^2 + \lambda_x \sigma^2 (1-\beta\rho_u)^2 + \lambda_i (\kappa\sigma\rho_u - \gamma_u)^2}$$

$$f_\pi = \frac{\lambda_x \quad (1-\beta\rho_u) - \lambda_i(\kappa\sigma\rho_u - \gamma_u)(1-\rho_u)\sigma^{-2}}{\kappa^2 + \lambda_x \quad (1-\beta\rho_u)^2 + \lambda_i(\kappa\sigma\rho_u - \gamma_u)^2\sigma^{-2}}$$

Call the denominator h_π

$$h_\pi = \lambda_i \sigma^{-2}(\gamma_u - \kappa\sigma\rho_u)^2 + \lambda_x(1-\beta\rho_u)^2 + \kappa^2 \qquad \checkmark = \text{Woodford}$$

the numerator:

$$\lambda_i \sigma^{-2}(\gamma_u - \rho_u\kappa\sigma)(1-\rho_u) + \underbrace{\lambda_x(1-\beta\rho_u)}_{= \xi_u \text{ for him}} \qquad \checkmark = \text{Woodford}$$

Yes!

This f_π = Woodford's π_u in App. F.4. (p. 713).

Let's try to analogously solve for g_π $(=\pi_r)$ from $\mathcal{L}^{stab,r}$

$$\mathcal{L}^{stab,r} \propto g_\pi^2 + \lambda_x g_x^2 + \lambda_i g_i^2$$

(12) $g_x = \frac{1-\beta\rho_r}{\kappa} g_\pi$

(14) $(\rho_r - 1) g_x - \sigma g_i + \sigma\rho_r g_\pi + \sigma = 0$

$$\Leftrightarrow (\rho_r - 1)\left(\frac{1-\beta\rho_r}{\kappa}\right) g_\pi + \sigma\rho_r g_\pi - \sigma g_i + \sigma = 0$$

$$\left(\sigma\kappa\rho_r - (1-\rho_r)(1-\beta\rho_r)\right) g_\pi + \sigma\kappa = \sigma\kappa g_i$$

$$g_i = 1 + (\sigma\kappa\rho_r - (1-\rho_r)(1-\beta\rho_r))g_\pi$$

So $\mathcal{L}^{stab,r} = g_\pi^2 + \lambda_x \left(\frac{1-\beta\rho_r}{\kappa}\right)^2 g_\pi^2 + \lambda_i \left(1 + (\sigma\kappa\rho_r - \gamma_r)g_\pi\right)^2$

$$= \left[1 + \lambda_x\left(\frac{1-\beta\rho_r}{\kappa}\right)^2\right] g_\pi^2 + \lambda_i\left[(\sigma\kappa\rho_r - \gamma_r)^2 g_\pi^2 + 2(\sigma\kappa\rho_r - \gamma_r)g_\pi + 1\right]$$

FOC g_π: $2\left[1 + \lambda_x\left(\frac{1-\beta\rho_r}{\kappa}\right)^2 + \lambda_i(\sigma\kappa\rho_r - \gamma_r)^2\right]g_\pi + 2\lambda_i(\sigma\kappa\rho_r - \gamma_r)$

$= 0$

$$g_\pi = \frac{\lambda_i(\sigma\kappa\rho_r - \gamma_r)}{1 + \lambda_x\left(\frac{1-\beta\rho_r}{\kappa}\right)^2 + \lambda_i(\sigma\kappa\rho_r - \gamma_r)^2}$$

this doesn't seem to be quite what Woodford gets.

But stop for a moment: maybe I don't need to resolve the $\bar{r}$-issue if I can take his result and understand how he gets the Taylor-rule coeffs from it.

He has the one:

$\pi = \bar{\pi} + \pi_r \cdot r + \pi_u \cdot u$, $x = \bar{x} + x_r \cdot r + x_u \cdot u$

$i = \bar{i} + i_r \cdot r + i_u \cdot u$

At the same time, we have

$i_t = \phi_\pi \pi_t + \phi_x x_t$

so maybe we can do coeff comparison

$i_t = \phi_\pi(\bar{\pi} + \pi_r \cdot r_t^n + \pi_u \cdot u_t) + \phi_x(\bar{x} + x_r \cdot r_t^n + x_u \cdot u_t)$

$= \underbrace{(\phi_\pi \bar{\pi} + \phi_x \bar{x})}_{\bar{i}} + \underbrace{(\phi_\pi \pi_r + \phi_x x_r)}_{i_r} r_t^n + \underbrace{(\phi_\pi \pi_u + \phi_x x_u)}_{i_u} u_t$

$= \bar{i} + i_r \cdot r_t^n + i_u \cdot u_t$

$\rightarrow$ I'm just surprised b/c if we only had i_r & i_u, then we could solve the following system for (ϕ_x, ϕ_π)

$$\left.\begin{array}{ll} \phi_\pi \pi_r + \phi_x x_r = i_r & (1) \\ \phi_\pi \pi_u + \phi_x x_u = i_u & (2) \end{array}\right\} \text{ 2 eqs in 2 unknowns}$$

But now having the extra equation

$\phi_\pi \bar{\pi} + \phi_x \bar{x} = \bar{i}$

feels like we were "overdetermined" since I think we know $\bar{i}$ (as well as $\bar{\pi}$ and $\bar{x}$). Ok maybe we don't know $\bar{i}$ yet b/c we only have 2 model equations in

3 variables: $\bar{x}$, $\bar{\pi}$ and $\bar{i}$.

OR: you know what: the TR may have 3 parameters

$$i_t = \phi_\pi \pi_t + \phi_x x_t + \bar{i}$$

if this is also a param, then

$$\phi_\pi \bar{\pi} + \phi_x \bar{x} + \bar{i} \stackrel{!}{=} \bar{i}^{\,oni} \leftarrow \text{the LR value from oni.}$$

Ok, here's the deal: I think we aren't able to solve for $(\bar{\pi}, \bar{x}, \bar{i})$ fully from oni b/c the two model eqs (ass-ing an AR(1) for r_t^n) give us 2 LHSs:

$$\left.\begin{aligned} (1-\beta)\bar{\pi} - k\bar{x} &= 0 \\ \sigma[\bar{i} - \bar{\pi}] &= 0 \end{aligned}\right\} \quad \begin{aligned} \bar{x} &= \frac{1-\beta}{k}\bar{\pi} \\ \bar{i} &= \bar{\pi} \end{aligned}$$

↳ and so given $(\phi_\pi, \phi_x)^{oni}$ we can determine $\bar{i}^{\,oni}$.

(maybe)

Note: in Giannoni & Woodford (2002b NBER WP) they write

$$\hat{r}_t^n = \rho_r \hat{r}_{t-1}^n + \varepsilon_{rt} \quad w/ \quad \hat{r}_t^n = (r_t^n - \bar{r}) \quad \Rightarrow \varepsilon_t = (\hat{r}_t^n, u_t)'.$$

An idea: where is $\hat{r}_t^n$ and where r_t^n?

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + u_t$$

$$x_t = E_t x_{t+1} - \sigma \left[i_t - E_t \pi_{t+1} - r_t^n \right]$$

$$\bar{\pi} + f_\pi u_t + g_\pi \hat{r}_t^n = \kappa(\bar{x} + f_x u_t + g_x \hat{r}_t^n) + \beta(\bar{\pi} + f_\pi \rho_u u_t + \underbrace{g_\pi \hat{r}_{t+1}^n}_{g_\pi \rho_r \hat{r}_t^n}) + u_t$$

$$\Rightarrow \underbrace{(1-\beta)\bar{\pi} - \kappa \bar{x}}_{\equiv M1} = \underbrace{(-f_\pi + \kappa f_x + \beta f_\pi \rho_u + 1)}_{\equiv C1} u_t + \underbrace{(-g_\pi + \kappa g_x + \beta g_\pi \rho_r)}_{\equiv C2} \hat{r}_t^n$$

$$\bar{x} + f_x u_t + g_x \hat{r}_t^n = \bar{x} + f_x \rho_u u_t + g_x \rho_r \hat{r}_t^n - \sigma(\bar{i} + f_i u_t + g_i \hat{r}_t^n) + \sigma(\bar{\pi} + f_\pi \rho_u u_t + g_\pi \rho_r \hat{r}_t^n) + \sigma(\underline{\hat{r}_t} + \bar{r})$$

$$\Leftrightarrow$$

$$\underbrace{\sigma(-\bar{\pi} - \bar{r} + \bar{i})}_{\equiv M2} = \overbrace{(-f_x + f_x \rho_u - \sigma f_i + \sigma f_\pi \rho_u)}^{\equiv C3} u_t + \underbrace{(-g_x + g_x \rho_r - \sigma g_i + \sigma g_\pi \rho_r + \sigma)}_{\equiv C4} \hat{r}_t^n$$

$M2$: $\bar{\pi} = \bar{i} - \bar{r}$ $\Rightarrow$ So can solve for f_j, g_j $j = \pi, x, i$ as I did,

$M1$: $\bar{x} = \frac{1-\beta}{\kappa} \bar{\pi}$ and then obtain $(\bar{i}, \phi_\pi, \phi_x)$ by coeff-comparison.

Let's pause Woodford there. Let's turn to tests of structural change

- Lütkepohl, Introduction to Multiple Time Series Analysis
  "Multiple TS"
- Killian & Lütkepohl, "SVAR Analysis" (→ pdf)
- Hamilton → doesn't seem to be anything in it.

This only says about cusum & cusum-sq: → "prone to rejecting H_0: no break in small samples even when H_0 is true when the DGP involves large transitory dynamics." p. 72

Lütkepohl, "Multiple TS"

4.6. Tests for Structural Change p. 159

Let y_t be a Gaussian VAR(p), K-dimensional, stationary: $y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t$ (4.1.1)

The optimal h-step ahead fcst at time T is $y_T(h)$.

The corresponding fcst error is

$$e_T(h) = y_{T+h} - y_T(h) = \sum_{i=0}^{h-1} \phi_i u_{T+h-i} = [\phi_{h-1}, \dots, \phi_1, I_K] u_{T,h}$$

$(K \times 1)$ ↳ MA-representation

Since $u_{T,h} \sim N(0, I_h \otimes \Sigma_u)$, the fcst error is a linear transformation of a multivariate normal distrib and is thus also MVN:

$e_T(h) \sim N(0, \Sigma_y(h))$

where $\Sigma_y(h) = \sum_{i=0}^{h-1} \phi_i \Sigma_u \phi_i'$

is the fcst MSE matrix (i.e. the FEV).

So since the FE ~ MVN w/ VC matrix $\Sigma_y(h)$, the statistic $\tau_h := e_T(h)' \Sigma_y(h)^{-1} e_T(h)' \sim \chi^2(k)$

i.e. the multiplication of two MVNs (scaled by VC) is distrib as χ^2.

$\Rightarrow$ So τ_h can be used to test the following H_0:

H_0: y_{T+h} is generated by the same VAR(p) (Gaussian) as $y_1, \ldots, y_T$.

if $\tau_h \geq$ critical value $\chi^2(k)$.

That's really cool but not feasible b/c it involves unknown quantities: the FE $\hat{e}_T(h)$ and the FEV $\Sigma_y(h)$.

I'm not gonna go thru these in detail b/c I have them.

I'm just noting in passing that Lütkepohl estimates:

$$\hat{\Sigma}_y(h) = \sum_{i=0}^{h-1} \hat{\phi}_i \hat{\Sigma}_u \hat{\phi}_i'$$

$$\hat{FEV} = \sum_{i=0}^{h-1} (\text{MA-coeffs}_i) \Sigma_{\text{errors}} (\text{MA-coeffs}_i)'$$

Ok, then you can use:

$$\hat{\tau}_h = \hat{e}_T(h)' \hat{\Sigma}_y(h) \hat{e}_T(h)$$

where I'm wondering if he forgot the (-1) here

(very likely! I think it's a typo!)

$$\hat{\tau}_h \xrightarrow{d} \chi^2(k)$$

He adds however that $\hat{\tau}_h \to \chi^2(k)$ likely won't hold in small samples, in which case the statistic

$$\bar{\tau}_h := \hat{e}_T(h)' \hat{\Sigma}_{\hat{y}}(h)^{-1} \hat{e}_T(h) \cdot \frac{1}{k} \approx F(\underset{\underset{3}{\uparrow}}{k}, \underset{\underset{\uparrow}{\uparrow}}{T} - \underset{\underset{3 \cdot 1}{\uparrow}}{\widetilde{kp}^{2}} - 1)$$

where 2 changes:

1) for $\hat{\Sigma}_{\hat{y}}(h)$ we use a different estimator (Sect 3.5.2)
2) we divide by d.o.f. k in order to adjust for the fact that we're using an adjusted FEV-matrix.

This is all cool but I'm not sure I see the link to the cusum test.

Ok - here's some more thinking:

CEMP's version of the cusum-sq test is like a mix of cusum and Lütkepohl b/c the statistic that is computed is a squared FE, normalized by an estimated FEV. The relation is that the standardized residuals in the cusum-test are also FE's, divided by an estimate of the FEV.

So in that sense, Lütkepohl is describing a kind of "in spirit cusum-test" for VARs.

My concern is that technically my z_t vector is not a VAR... although in a sense it is b/c $z_t = \hat{g}_x s_t$ and s_t is a VAR(1). So I think the MVNormality of of FEV should still hold, and thus, my statistic $\hat{\tau}$ should also $\longrightarrow \chi^2(K)$

Ok so compare then the critical values of

$$\hat{\tau} = f' \omega^{-1} f \sim \chi^2(\underbrace{K}_{=n_y}) = \chi^2(3)$$

vs.

$$\tilde{\tau} = \underbrace{f' \omega^{-1} f}_{\text{and I'm not changing my estimate of the FEV b/c I already have it.}} / \underbrace{K}_{=n_y} \quad \tilde{\sim} F(\underbrace{K}_{=n_y}, T - \underbrace{K}_{=n_y} \cdot \underbrace{p}_{\text{order of VAR, }=1} - 1)$$

T: increases as agents' sample grows.

$$F(3, T-3)$$

$0 \to \infty$

For $\alpha = 0.05$ one-sided test (upper tail)

$\chi^2(3)$	7.815
$F(3,1)$	215.7073
$F(3,6)$	4.7571
$F(3,120)$	2.6802
$F(3,\infty)$	2.6049

The funny thing is that for $\hat{\tau}$, $\tilde{\theta} = 2.5$ seems to be doing well.

But I still obtain that cusum anchors more as $\psi_\pi \uparrow$ and I don't feel that I'm closer to understanding why.

1) I had the observation before that not squaring things made the #anchoring $= f(\psi_\pi)$ non-monotonic:
- for low values of ψ_π, lots of anchoring
- for intermediate values of ψ_π, less anchoring
- for high values of ψ_π, more anchoring again

2) Squaring makes the #anchoring $= f(\psi_\pi)$ monotonically increasing

→ why this difference?

I think I know why: b/c squaring is like an abs. value

→ it kinda makes "errors have the same sign"

I've confirmed w/ the old -cusum code that when you take f^2 or $\sqrt{f^2}$ you get monotonically ↑ #anchoring in $\psi_\pi \uparrow$, but for f, you lose this monotonicity.

Ok and why else square? b/c sum of squared normals is χ^2 : $Z_i \sim N(0,1) \rightarrow Q = \sum_{i=1}^{k} Z_i^2 \sim \chi^2(k)$

↳ so if you don't square goodness knows how θ_t is distributed then.

$\theta_t = \frac{FE^2}{\omega}$ and we're dividing by the variance so we have standard normals.

Wiki: Let $z \sim MVN(0, B)$

then $X = z' A z \sim$ generalized $\chi^2(A, B)$.

or math.hkbu.edu.hk (hpeng/Math3806/Lecture_note3.pdf)

$X \sim N_p(\mu, \Sigma)$ then

$(X-\mu)' \Sigma^{-1} (X-\mu) \sim \chi^2_p$

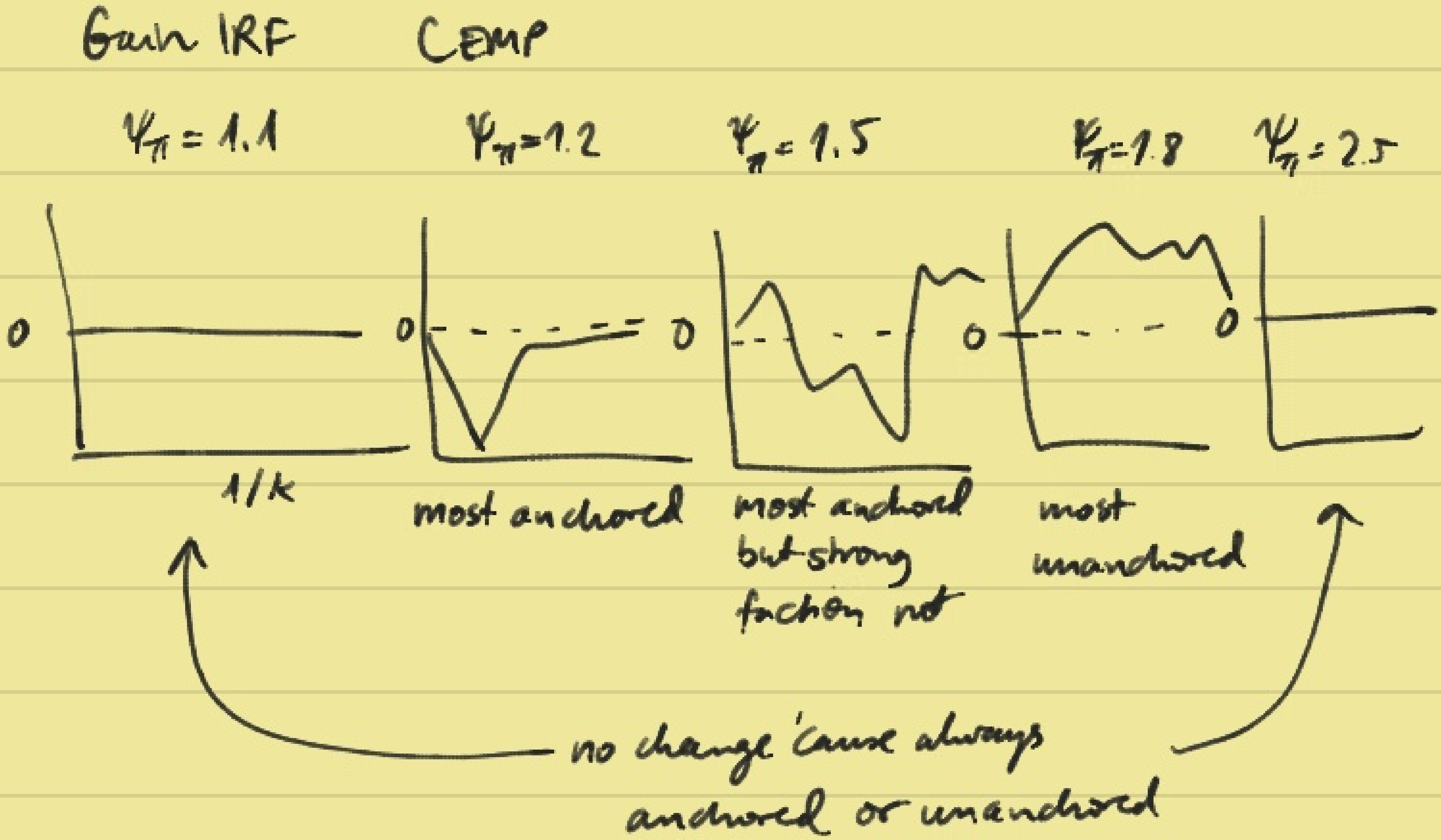
Look so we're constructing a χ^2-statistic, very consciously.

But why does this behave opposite to CEMP's?

3) A 3rd observation: for the Lütkepohl-style cusum, while $\chi_\alpha \uparrow$ leads to more anchoring monotonously, it actually leads to less anchoring early in the sample

→ no analogy for the CEMP criterion for this!

→ CEMP's criterion is kind of smoother: its relationship to anchoring is smoother / more monotonic.



→ they are telling the same story as the simulation: a too high or low ψ_π won't change the anchoring situation in response to a shock b/c it's set already; however, a low ψ_π tendentially can lead to more anchoring after a shock than a high one.

Gain IRF   CUSUM

$\psi_\pi = 1.1$ $\psi_\pi = 1.2$ $\psi_\pi = 1.5$ $\psi_\pi = 1.8$ $\psi_\pi = 2$

You know, I have a hard time reading these also b/c they are conditional on how many were anchored at that point

→ but it seems like a shock can have different effects on impact vs in later periods

Whatevs.

CEMP: $\left|\phi - [\hat{F}, \hat{G}]\right|_{\Sigma^{-1}} \overset{?}{>} \bar{\theta}$ CUSUM: $f' \omega^{-1} f \overset{?}{>} \tilde{\theta}$

Ok: here's the big diff between the two:

$\hat{F}, \hat{G}$ are functions of LH expectations (as a function of ϕ)

→ that's why they dislike ψ_π high b/c that makes

$\hat{F}, \hat{G}$ move a lot, making $\phi - [\hat{F}, \hat{G}]$ big!

However, when ψ_π is high, current π moves less in response to shocks, decreasing 1-period-ahead fcst errors, so the CUSUM criterion becomes small!

→ got it!

Back to Woodford's ooi $\qquad$ 1 Feb 2020

The equations of RE were:

$$\underbrace{(1-\beta)\bar{\pi} - \kappa\bar{x}}_{\equiv M1} = \underbrace{(-f_\pi + \kappa f_x + \beta f_\pi \rho_u + 1)}_{\equiv C1} u_t + \underbrace{(-g_\pi + \kappa g_x + \beta g_\pi \rho_r)}_{\equiv C2} \hat{r}_t^n$$

$$\underbrace{\sigma(-\bar{\pi} - \bar{r} + \bar{i})}_{\equiv M2} = \overbrace{(-f_x + f_x\rho_u - \sigma f_i + \sigma f_\pi \rho_u)}^{\equiv C3} u_t$$
$$+ \underbrace{(-g_x + g_x\rho_r - \sigma g_i + \sigma g_\pi \rho_r + \sigma)}_{\equiv C4} \hat{r}_t^n$$

M1: $\bar{x} = \frac{1-\beta}{\kappa}\bar{\pi}$

M2: $\bar{\pi} = \bar{i} - \bar{r}$

C1: $(1-\beta\rho_u) f_\pi = 1 + \kappa f_x$

C3: $(1-\rho_u) f_x + \sigma f_\pi \rho_u = \sigma f_i$

C2: $(1-\beta\rho_r) g_\pi = \kappa g_x$

C4: $(1-\rho_r) g_x + \sigma g_\pi \rho_r + \sigma = \sigma g_i$

I solved $\min \mathcal{L}^{stab,u}$ st. C1 & C3 $\rightarrow f_\pi^{ooi}, f_x^{ooi}, f_i^{ooi}$

and $\min \mathcal{L}^{stab,r}$ st. C2 & C4 $\rightarrow g_\pi^{ooi}, g_x^{ooi}, g_i^{ooi}$

on Mathematica (materials15) and got the same as Woodford.

What I still don't get is how Woodford is able to solve for $\bar{z} = (\bar{\pi}, \bar{x}, \bar{i})$ at this stage. From his expressions it's clear that $\bar{x} = \frac{1-\beta}{\kappa}\bar{\pi}$ and $\bar{i} = \bar{\pi} + \bar{r}$, but he seems to be able to solve for $\bar{\pi}$. Is it from $\mathcal{L}^{det}$?

↳ Yes!

$$\mathcal{L}^{det} = \sum_{T=t}^{\infty} \beta^{T-t} \left[(E_t \pi_T)^2 + \lambda_x (E_t x_T - x^*)^2 + \lambda_i (E_t i_t - i^*)^2 \right]$$

(p. 501, slightly modified to include i)

Let's plug in the conjectures and M1 & M2

$$\Rightarrow \mathcal{L}^{det} = \sum_{T=t}^{\infty} \beta^{T-t} \left[(\bar{\pi})^2 + \lambda_x (\bar{x} - x^*)^2 + \lambda_i (\bar{i} - i^*)^2 \right]$$

(shocks are mean zero)

$$\Rightarrow \mathcal{L}^{det} = \sum_{T=t}^{\infty} \beta^{T-t} \left[\bar{\pi}^2 + \lambda_x \left(\frac{1-\beta}{\kappa}\bar{\pi} - x^*\right)^2 + \lambda_i (\bar{\pi} + \bar{r} - i^*)^2 \right]$$

$$\Rightarrow \mathcal{L}^{det} = \frac{1}{1-\beta} \left[\bar{\pi}^2 + \lambda_x \left(\frac{1-\beta}{\kappa}\right)^2 \bar{\pi}^2 - 2\lambda_x \frac{1-\beta}{\kappa}\bar{\pi}x^* + \lambda_x (x^*)^2 + \lambda_i \left(\bar{\pi}^2 + \bar{r}^2 + i^{*2} + 2\bar{\pi}\bar{r} - 2\bar{\pi}i^* - 2\bar{r}i^*\right) \right]$$

FOC: $\underline{\cancel{2}\bar{\pi}} + \underline{\cancel{2}\lambda_x \left(\frac{1-\beta}{\kappa}\right)^2 \bar{\pi}} - \cancel{2}\lambda_x \frac{1-\beta}{\kappa}x^* + \underline{\cancel{2}\lambda_i \bar{\pi}} + \cancel{2}\lambda_i \bar{r} - \cancel{2}\lambda_i i^* = 0$

$$\left(1 + \lambda_x \frac{(1-\beta)^2}{\kappa^2} + \lambda_i\right)\bar{\pi} = \lambda_x \frac{(1-\beta)}{\kappa} x^* - \lambda_i \bar{r} + \lambda_i i^*$$

$$(\bar\pi + \bar r - i^*)^2 = (\bar\pi + \bar r - i^*)\ (\bar\pi + \bar r - i^*)$$

$$= \cancel{\bar\pi^2} + \cancel{\bar\pi \bar r} - \cancel{\bar\pi i^*} + \cancel{\bar\pi \bar r} + \cancel{\bar r^2} - \cancel{\bar r i^*} - \cancel{\bar\pi i^*} - \cancel{\bar r i^*} + \cancel{i^{*2}}$$

$$= \bar\pi^2 + \bar r^2 + i^{*2} + 2\bar\pi \bar r - 2\bar\pi i^* - 2\bar r i^*$$

$$\rightarrow \left(1 + \lambda_x \frac{(1-\beta)^2}{k^2} + \lambda_i\right)\bar\pi = \lambda_x \frac{(1-\beta)}{k} x^* - \lambda_i \bar r + \lambda_i i^*$$

$$\Leftrightarrow (k^2 + \lambda_x (1-\beta)^2 + \lambda_i k^2)\bar\pi = \lambda_x k (1-\beta) x^* - \lambda_i k^2 \bar r + \lambda_i k^2 i^*$$

$$\Leftrightarrow \bar\pi = \frac{\lambda_x k (1-\beta) x^* - \lambda_i k^2 \bar r + \lambda_i k^2 i^*}{k^2 + \lambda_x (1-\beta)^2 + \lambda_i k^2} \qquad | : k^2$$

$$\bar\pi = \frac{\lambda_x k^{-1} (1-\beta) x^* + \lambda_i (i^* - \bar r)}{1 + (1-\beta)^2 k^{-2} \lambda_x + \lambda_i}$$

✓ Woodford yeah!!

What I now don't get is 1) in the opt. TR section (p. 574), why does Woodford have $\bar\pi = \frac{\lambda_i}{\lambda_i + \beta}(i^* - \bar r)$?

Even if I set $\lambda_x = 0$ I'd get $\frac{\lambda_i (i^* - \bar r)}{1 + \lambda_i}$ ...

2) If we were able to solve for $\bar\pi$ and thus for $\bar x$ and $\bar i$ then coeff-comparison in the Taylor-rule is overdetermined!

Sorry. Woodford's TR is:

$$i_t = \bar{i} + \phi_\pi(\pi_t - \bar{\pi}) + \phi_x(x_t - \bar{x})/4$$

$$= \bar{i} + \phi_\pi\left(\bar{\pi} - \bar{\pi} + f_\pi u_t + g_\pi \hat{r}_t^n\right) + \phi_x/4\left(\bar{x} - \bar{x} + f_x u_t + g_x \hat{r}_t^n\right)$$

$$= \bar{i} + (\phi_\pi f_\pi + \phi_x f_x)u_t + (\phi_\pi g_\pi + \phi_x g_x)\hat{r}_t^n$$

has to be coefficient-compared w/

$$i_t = \bar{i} + f_i u_t + g_i \hat{r}_t^n$$

whew! so $\bar{i} = \bar{i}$

and ϕ_π and ϕ_x solve

$$\phi_\pi f_\pi + \phi_x f_x = f_i \qquad T1$$

$$\phi_\pi g_\pi + \phi_x g_x = g_i \qquad T2$$

OMG ... in Mathematica, solving the system $\begin{bmatrix} T1 \\ T2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and setting $\rho_r = \rho_u = \rho$, I obtain EXACTLY what Woodford gets for ϕ_π^* and ϕ_x^* in (3.12).

(materials15.nb)

Ok, so recap how to get optimal Taylor rule coeffs using the optimal noninertial plan:

<u>Step 1</u> Postulate conjectures $z_t = \bar{z} + f_j u_t + g_j \hat{r}_t^n$

where $z = \begin{bmatrix} \pi \\ x \\ i \end{bmatrix}$ $j = \pi, x, i$, $\hat{r}_t^n = r_t - \bar{r}$

<u>Step 2</u> Use the model equations (NKIS, NKPC) and the LOMs of the two shocks to derive M1 & M2 as constraints on $(\bar{\pi}, \bar{x}, \bar{i})$, and C1, C2, C3 and C4 as constraints on f_j, g_j $j = \pi, x, i$. by plugging the conjectures into the model equations and using the LOMs of shocks to obtain expected future values of shocks.

<u>Step 3.</u> Solve 3 sets of optimizations

1. $\min \mathcal{L}^{det}$ s.t. M1 & M2 $\rightarrow$ get $\bar{\pi} \rightarrow \bar{x}, \bar{i}$
2. $\min \mathcal{L}^{stab,u}$ s.t. C1 & C3 $\rightarrow$ get $f_\pi^{oni} \rightarrow f_x^{oni}, f_i^{oni}$
3. $\min \mathcal{L}^{stab,r}$ s.t. C2 & C4 $\rightarrow$ get $g_\pi^{oni} \rightarrow g_x^{oni}, g_i^{oni}$

<u>Step 4.</u> Compare coeffs of TR to $i_t = \bar{i} + f_i u_t + g_i \hat{r}_t^n$ to solve $T1 = 0$ & $T2 = 0$ for (ϕ_π^*, ϕ_x^*)

What changes / stays the same for the learning model.

- M1 & M2, $\mathcal{L}^{det}$ are the same, so $\bar{\pi}, \bar{x}, \bar{i}$ are the same.
- I'm going to assume identical shock processes to Woodford's, that is $\hat{r}_t^n = \rho_r \hat{r}_{t-1}^n + \varepsilon_t^r \quad \hat{r}_t = r_t - \bar{r}$

(For simplicity, I'll impose $\rho_u = \rho_r = \rho$ too!) $\quad u_t = \rho_u u_{t-1} + \varepsilon_t^u$

- I'm going to assume identical CB loss

$$\mathcal{L}^{CB} = E_t \sum_{T=t}^{\infty} \beta^{T-t} \left[\pi_t^2 + \lambda_x (x_t - x^*)^2 + \lambda_i (i_t - i^*)^2 \right] \quad (2.29)$$

as Woodford, so $\mathcal{L}^{det}$, $\mathcal{L}^{stab,u}$, $\mathcal{L}^{stab,r}$ are the same

(I don't even need to verify that his $\mathcal{L}^{stab}$ is correct b/c the part I suspect is wrong is just a multiplicative constant and thus doesn't matter.)

- I'm going to assume identical Taylor rule

$$i_t = \bar{i} + \psi_\pi (\pi_t - \bar{\pi}) + \psi_x (x_t - \bar{x}) \quad (14) \quad (3.1)$$

doesn't matter

- What changes is C1, C2, C3, C4 and thus the sols to f_π, f_x, f_i and g_π, g_x, g_i

$\rightarrow$ therefore the sols to $T1 \overset{!}{=} 0$ & $T2 \overset{!}{=} 0$ will involve

$$(\psi_\pi^*, \psi_x^*) \neq (\phi_\pi^*, \phi_x^*)$$

learning $\quad$ RE

Ok so learning C1-C4.

$$\pi_t - \kappa x_t = \hat{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \{ \kappa\alpha\beta x_{T+1} + (1-\alpha)\beta\pi_{T+1} + u_T \}$$

! TR not plugged !

$$x_t + \sigma i_t = \hat{E}_t \sum_{T=t}^{\infty} \beta^T \{ (1-\beta) x_{T+1} - \sigma\beta i_{T+1} + \sigma\pi_{T+1} + \sigma r_T^n \}$$

Ignore the deterministic part to save space.

$$f_\pi u_t + g_\pi \hat{r}_t^n - \kappa f_x u_t - \kappa g_x \hat{r}_t^n = \hat{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \{ \kappa\alpha\beta [f_x u_{T+1} + g_x \hat{r}_{T+1}^n]$$
$$+ (1-\alpha)\beta [f_\pi u_{T+1} + g_\pi \hat{r}_{T+1}^n] + u_T \} \quad (1)$$

$$f_x u_t + g_x \hat{r}_t^n + \sigma f_i u_t + \sigma g_i \hat{r}_t^n = \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \{ (1-\beta) [f_x u_{T+1} + g_x \hat{r}_{T+1}^n]$$
$$- \sigma\beta [f_i u_{T+1} + g_i \hat{r}_{T+1}^n] + \sigma [f_\pi u_{T+1} + g_\pi \hat{r}_{T+1}^n] + \sigma \hat{r}_T^n \} \quad (2)$$

note: I'm turning this into $\hat{r}$ b/c $r_t = \hat{r}_t + \bar{r}$, take $\bar{r}$ to LHS (deterministic part)

$\Leftrightarrow$

$$(f_\pi - \kappa f_x) u_t + (g_\pi - \kappa g_x) \hat{r}_t^n = \hat{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \{ [\kappa\alpha\beta f_x + (1-\alpha)\beta f_\pi] u_{T+1}$$
$$+ [\kappa\alpha\beta g_x + (1-\alpha)\beta g_\pi] r_{T+1}^n + \alpha\beta u_{T+1} \} + u_t \quad (1)$$

$$(f_x + \sigma f_i) u_t + (g_x + \sigma g_i) \hat{r}_t^n = \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \{ [(1-\beta) f_x - \sigma\beta f_i + \sigma f_\pi] u_{T+1}$$
$$+ [(1-\beta) g_x - \sigma\beta g_i + \sigma g_\pi] \hat{r}_{T+1}^n + \sigma\beta \hat{r}_{T+1}^n] + \sigma \hat{r}_t^n \quad (2)$$

$\Longrightarrow$

$$(f_\pi - \kappa f_x - 1)u_t + (g_\pi - \kappa g_x)\hat{r}_t^{\,n} = \frac{\kappa\alpha\beta f_x + (1-\alpha)\beta f_\pi + \alpha\beta}{1-\alpha\beta\rho_u}\,u_t$$

$$+ \frac{\kappa\alpha\beta g_x + (1-\alpha)\beta g_\pi}{1-\alpha\beta\rho_r}\,\hat{r}_t^{\,n} \qquad (1)$$

$$(f_x + z f_i)u_t + (g_x + z g_i - z)\hat{r}_t^{\,n} = \frac{(1-\beta)f_x - z\beta f_i + z f_\pi}{1-\beta\rho_u}\,u_t$$

$$+ \frac{(1-\beta)g_x - z\beta g_i + z g_\pi + z\beta}{1-\beta\rho_r}\,\hat{r}_t^{\,n} \qquad (2)$$

$$\Longleftrightarrow \overbrace{\left[f_\pi - \kappa f_x - 1 - \frac{\kappa\alpha\beta f_x + (1-\alpha)\beta f_\pi + \alpha\beta}{1-\alpha\beta\rho_u}\right]}^{=:C1} u_t$$

$$+ \overbrace{\left[g_\pi - \kappa g_x - \frac{\kappa\alpha\beta g_x + (1-\alpha)\beta g_\pi}{1-\alpha\beta\rho_r}\right]}^{=:C2} \hat{r}_t^{\,n} \overset{!}{=} 0 \qquad (1)$$

$$\text{and } \overbrace{\left[f_x + z f_i - \frac{(1-\beta)f_x - z\beta f_i + z f_\pi}{1-\beta\rho_u}\right]}^{=:C3} u_t$$

$$+ \underbrace{\left[g_x + z g_i - z - \frac{(1-\beta)g_x - z\beta g_i + z g_\pi + z\beta}{1-\beta\rho_r}\right]}_{=:C4} \hat{r}_t^{\,n} \overset{!}{=} 0 \qquad (2)$$

Just thinking in preparation for meeting on Wed

4 Feb 2020

Why do the CEMP vs. CUSUM criterion behave opposite ways? ⇒ is it really the case that the CEMP criterion internalizes LH expectations and CUSUM doesn't? I don't think so right now.

$\hat{E}_{t-1} z_t - E_{t-1} z_t$ vs. $f = \hat{E}_{t-1} z_t - z_t$

$\phi_{t-1}\begin{bmatrix}1\\ s_{t-1}\end{bmatrix} - [F, G]\begin{bmatrix}1\\ s_{t-1}\end{bmatrix}$

This isn't a fcst error. It's the discrepancy between non-rational and rational fcsts.

↳ $\phi_{t-1} - [F, G]$

$\theta^{CEMP} :=$

shouldn't I reevaluate these every period? I do though!

This is the non-rational fcst. error.

$= \phi_{t-1}\begin{bmatrix}1\\ s_{t-1}\end{bmatrix} - [F, G]\begin{bmatrix}1\\ s_t\end{bmatrix}$

!

$\phi_{t-1}\begin{bmatrix}1\\ s_{t-1}\end{bmatrix} - [F, G]\begin{bmatrix}1\\ s_{t-1} + \varepsilon_t\end{bmatrix}$

$= \phi_{t-1}\begin{bmatrix}1\\ s_{t-1}\end{bmatrix} - [F, G]\begin{bmatrix}1\\ s_{t-1}\end{bmatrix}$

$- [F, G]\begin{bmatrix}0\\ \varepsilon_t\end{bmatrix}$

$= \theta^{CEMP} - [F, G]\begin{bmatrix}0\\ \varepsilon_t\end{bmatrix}$

So since $\theta^{cusum} \approx f' \omega^{-1} f$

$$= \left(\theta^{comp} - [F, G]\begin{bmatrix}1\\ \varepsilon_t\end{bmatrix}\right)' \omega^{-1} \left(\theta^{comp} - [F, G]\begin{bmatrix}1\\ \varepsilon_t\end{bmatrix}\right)$$

var of shocks, Σ, constant

$$\approx \frac{\uparrow(\theta^{comp})^2 + [F, G]^2 \begin{bmatrix}1\\ \varepsilon_t\end{bmatrix}^2 - 2\,\text{cross-term}}{\omega}$$

When $\psi_\pi \uparrow$, $\theta^{comp} \uparrow$

I'm sorry, I don't see anything from here.

What instead I'm thinking to try is an analogue of cusum for comp

instead of $\Sigma^{-1} |\phi - [F, G]|$

call $\tilde{\Sigma} := (\phi - [F, G])' \Sigma^{-1} (\phi - [F, G])$

Problem: $\tilde{\Sigma}$ is 4x4.

Why? B/c $(\phi - [F, G])$ is 3x4

Ok, I see why we need to square for CUSUM and why it makes things smoother:

b/c the FE, f, consists of 2 parts:

1) $\phi - [F, G]$ "discrepancies between RE & PLM coming from learning"

2) ε_t noise.

↳ having noise enter there is undesirable, but it's worse when it enters in levels b/c it can dampen fcst errors (depending on the sign)

⇒ squaring is like biasing your results up (θ^{cusum} is too high) but at least you know that it's not low due to the wrong reason.

⇒ you can measure the size of the bias as the FEV, ω, and then descale by that.

→ which is what θ^{cusum} does!

↳ but that also means that the squared error is being cleaned out. → but does the squaring of whatever's left make a diff?

Ok, looking at CEMP, what I've understood is that θ^{CEMP} = |expected fcst error|

↳ knowing the model

i.e. | fcst error − noise |

$E(\text{noise}) = 0$.

$$|E[FE]| \quad \text{vs.} \quad \frac{FE^2}{FEV}$$

But trying a version of CEMP squared (θ^{CEMP^2}) still yields the same CEMP dynamics ($\psi_\pi \uparrow$ anchoring $\downarrow$)

↳ it doesn't seem to be the squaring that makes a difference. It's the concept of FEs.

$E[PLM - ALM]$ $\qquad (PLM - ALM)^2 = (\theta^{CEMP} + \text{shocks})^2$

$$= (\theta^{CEMP})^2 + \text{shocks}^2 + 2\theta^{CEMP}\text{shocks}$$

should at least move in the same direction as θ^{CEMP}

we took care of this by $\hat{W}^{-1}$

Ok: what I've tried is to input $\Sigma \overset{=\eta\eta'}{=} Var(\varepsilon_t)$ into Θ^{CUSUM} <u>and NOW</u> $\psi_\pi \uparrow \Rightarrow \downarrow$ anchoring!

↳ so it's the fact that ω^{-1} is learned (that is, estimated) which causes the discrepancy

↳ But is it the fact that it's estimated, or is it that it's the FEV?

Timeout: Wold:

$$
\begin{aligned}
y_t &= \rho y_{t-1} + e_t \\
&= \rho[\rho y_{t-2} + \varepsilon_{t-1}] + \varepsilon_t \\
&= \rho^2 y_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t \\
&= \sum_{k=0}^{t} \rho^k \varepsilon_{t-k}
\end{aligned}
$$

So would my analogy to Lütkepohl's $\Sigma_y = \sum_{i=0}^{h-1} \phi_i \Sigma_u \phi_i'$ be $\sum_{i=0}^{h-1} hx_i \Sigma hx_i'$?

$$\frac{E[FE]\uparrow}{Var(noise)} \quad vs \quad \frac{FE^2 \uparrow}{\hat{FEV} \uparrow}$$

I see. Imagine a shock that raises FE $\uparrow$

I think I see. CEMP's criterion focuses really on the part of the FE that comes from learning. For CUSUM, agents can't distinguish between this part of the FE and one due to noise: both raise the $\hat{FEV}$.

- For CEMP: $\psi_\pi \uparrow$ raises FE b/c more expectational mistakes.

For CUSUM too, but it also raises $\hat{FEV}$.

- If you $\uparrow$ Var(noise), $\theta^{CEMP} \downarrow$ while θ^{CUSUM} $\uparrow$ constant?

→ It has something to do w/ the fact that a strong reaction to π calms things down a bit on impact so CUSUM-agents think they're in a less volatile world.

If you try $f'f$ instead of $f'W^{-1}f$ — 5 Feb 2020

you also replicate the θ^{CEMP} behavior!

↳ the difference is <u>really</u> due to $\frac{1}{\widehat{FEV}}$.

Ignoring the constant,

$E[FE] = (\phi - [F,G])s_{t-1}$ $\quad FE = \phi s_{t-1} - [F,G]s_t$

$$= \phi s_{t-1} - [F,G][s_{t-1} + \varepsilon_t]$$

$$= (\phi - [F,G])s_{t-1} - [F,G]\varepsilon_t$$

$$= \theta^{CEMP} - [F,G]\varepsilon_t$$

$$\rightarrow FE^2 = \theta^{CEMP^2} + [F,G]^2\varepsilon_t^2 - 2[F,G]\theta^{CEMP}\varepsilon_t$$

We're saying that this behaves similarly to θ^{CEMP}.

Only once you divide by $\widehat{FEV}$ do you change behavior.

$$FEV = E[FE^2] = E[\theta^{CEMP^2} + [F,G]^2\varepsilon_t^2] \quad \text{b/c} \quad E[\varepsilon_t] = 0$$

$$= EE[FE] + [F,G]^2 Var(noise)$$

$$\rightarrow \widehat{FEV} = \theta^{CEMP^2} + [F,G]^2 \widehat{Var}(noise)$$

Supposing that $\widehat{Var}(noise)$ is an accurate estimate, we get

$$\theta^{CUSUM} = 1 - \frac{2[F,G]\theta^{CEMP}\varepsilon_t}{\theta^{CEMP^2} + [F,G]^2\widehat{Var}(noise)}$$

$$\theta^{CUSUM} = 1 - \frac{2\,[F,G]\,\theta^{CEMP}\,\varepsilon_t}{(\theta^{CEMP})^2 + [F,G]^2\,Var(\widehat{noise})}$$

Disregard the noise b/c if there is more ($\varepsilon_t \uparrow$), $Var(\widehat{noise}) \uparrow$ and $\theta^{CEMP} \downarrow$ both in det & num.

The point is that whatever the sign of θ^{CEMP}, this point goes in the opposite direction

$$\frac{1+2+5}{1+2} \begin{cases} \frac{8}{3} = 2\frac{2}{3} \\ \frac{1+2}{1+2} + \frac{5}{3} = 1 + 1\frac{2}{3} \end{cases}$$

The cross-term $-[F,G]\,\theta^{CEMP}\,\varepsilon_t$ is the discrepancy between FE^2 and $E[FE^2] = FEV$ (or its sample analog, $\widehat{FEV}$).

The point is that agents aren't able to evaluate the expectation $E[FE]$ or $E[FE^2]$ b/c they don't know the model. The next-best thing is FE^2 and $\widehat{FEV}$ that they can use. But FE^2 has the feature that it doesn't distinguish between errors coming from shocks vs errors coming from learning: this cross-term screws up inference for CUSUM-agents!

You'd ideally say:

$$E[FE^2] = FE^2 - \text{crossterm}$$

$$= \text{learning errors} + \text{noise errors} - \text{interaction}$$

$$= \text{variation in learning} + \text{var in noise} - \text{cov(learn, noise)}$$

<u>Questions for the analytical (ψ^*, ϕ^*)</u>

1.) What's the diff between ours and optimality?

2) Shocks to TR? (mon. pol. shocks)

3.) $\mathcal{L}^{CB} = \ldots \lambda_i (i_t - i^*)$ necessary?

4) ψ^* doesn't depend on <u>form</u> of learning?

↳ entire approach for learning

1) p. 466–467 lays out nicely why discretionary optimal policy is not first best, and, in fact, why purely forward-looking policies are also

not first best. Why? b/c

1. dynamic programming only yields the first best in situations where E(future policy) do not determine today's outcomes
2. optimal rules are therefore history-dependent b/c it must take into account the "advantages of anticipating the policy at earlier dates" (p. 467)

⇒ timeless-perspective will be this

Disadvantages of discretion

- inflationary bias (Kydland & Prescott (1977), Barro & Gordon (1983))
- suboptimal responses to unexpected shocks (p. 468)
  This is a less well-understood issue and cannot be solved w/ a purely fwd-looking policy. ("stabilization bias" p. 493)

- It seems like optimality from a timeless perspective only pins down uniquely the LR average (i.e. target) values, but not uniquely the state-contingent responses to shocks (p. 492 top)

- p. 495: optimal (commitment) policy is <u>history-dependent</u> in a way optimal discretion isn't: b/c under discretion, you have a clean slate: you just set whatever you have to set now. E.g. price level gets a new, higher st.st. value (fig. 7.3). But under commitment, the price level returns to its previous st.st. after a cost-push shock b/c past states matter for expectations about the future.

p. 495:

"Optimal policy is history dependent [..] b/c the <u>anticipation</u> by the private sector that future policy will be different as a result of conditions at date t"

- p. 503 the "responses to disturbances under a timelessly optimal policy are of the same kind as under a t_0-optimal commitment."

- p. 507 suggests that (t_0-) optimal commitment is about choosing a plan for (π, x) (and i), as is timeless optimality. This doesn't answer the question of what rule to use.

- it then seems that once you're given a rule of the form of a TR, you've already constrained yourself to purely forward-looking rules, so you've already "only 2nd best".
- but at least you can still be timelessly optimal and time-consistent within this class of rules ($\to L^{det}$ & L^{stab}) p. 507-8

and p. 510: they are "not fully optimal" (b/c not history-dependent)

2) Right now I think you could add shocks to the TR b/c it would add 2 more constraints and $\int^{\text{stab}, i}$

3) I don't think it's necessary to have $\lambda_i(\bar{i} - i^*)$ in the loss if one can find a way to determine $\bar{i}$ (of that I'm not sure).

4) E-formation doesn't matter for optimal policy beyond the fact that it's not rational E. But maybe this isn't optimal once learning has converged?
Potentially there's a superior TR which is history dependent $\rightarrow$ conditions on evolution of E?

Ryan meeting 5 Feb 2020

Is there a way to subtract out shocks?

⇒ should be able to get rid of shocks that are now realized & thus in info set.

→ is there a happy middle ground?

Check in lit if there's a default endog gain choice b/c you need a baseline w/ arguments for it.

Could still have a nonmechanical rule that depends on how expectations evolve

vs. a rule that takes into account a stance of expectations

send Clough - reminder. ✓

Work after

Some interpretation:

1) CEMP vs. CUSUM criterion:

Ryan is wondering if there's some way to reconcile the two given that when the FE is realized, time t shocks become known since s_{t-1} was already in the infoset and now s_t is in the infoset too. So couldn't you get a FE that subtracts ε_t and thus you're more aligned with CEMP's criterion?

$\Rightarrow$ in fact I think this is like saying

$$\theta^{cusum} - \left(-2\,\theta^{cemp}\,[F,G]\,\varepsilon_t \right)$$

i.e. like an error-corrected θ^{cusum}.

As for literature on endog. gain, there was sthg...

- Branch & Evans (2011) is related in a weird way.
- Marcet & Nicolini (2003) (p. 10 Mac)

2) Optimal TR - coeffs

It's a different thing to say that there's an oni where the LOM - Expectations are taken into account [i.e. a purely fwd-looking (std TR) policy, but one which encompasses how expectations evolve (→ the gain!)]

vs. a more optimal class of policy rule that conditions on the state variable E(·).
(i.e. one that ain't oni no more!)

The bad news is I have no idea how to go about the latter, and Ryan doesn't know how to go about the former. He seemed to suggest that the conjecture cannot be just $\bar{z} + F s_t$.

The other thing he said is to use simulation on the computer to get a sense of how the solution behaves.

Ok do some planning of how to proceed:
I have 5 days until I need to submit (rough) draft.

1.) I will need to choose a baseline criterion for that draft.

2.) I need to make a decision about what kind of result to present.

- It'd be great to present something analytical but it's unlikely I'll have something in 5 days.
- That would imply that I should prepare some simulation-based quasi-results.

  E.g. a set of CB-losses for LCMP- & CUSUM-criterion for different "internal parameters", potentially compared to RE would be good.

Ok so short-run to do:

① Compare Branch & Evans & Marcet & Nicolini for cod/log gains.

② Do fmincon w/ figures of loss for RE & learning for several configs of "internal parameters"

③ om w/ LOM-E(·).

① 1.1) Marcet & Nicolini 2003

Perceived inflation $=: \beta_t$

$$\beta_t = \beta_{t-1} + \underbrace{\frac{1}{\alpha_t}}_{\text{gain}} \underbrace{\Big(\underbrace{\frac{P_{t-1}}{P_{t-2}}}_{=:\pi_{t-1}} - \beta_{t-1} \Big)}_{=: FE_{t-1}}$$

$$\alpha_t = \begin{cases} \alpha_{t-1} + 1 & \text{if } \left| \dfrac{\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}}{\beta_{t-1}} \right| < \underbrace{\nu}_{= \bar{\theta} \text{ or } \tilde{\theta}} \\ \bar{\alpha} \ (\text{cgain}) & \text{else.} \end{cases}$$

$$\left| \frac{FE}{\text{last belief}} \right|$$

⇒ so this is an absolute FE scaled by the last belief

→ so it's closer to CUSUM, but it's not the squared CUSUM-type, and the scaling is extremely simple.

1.2.) <u>Branch & Evans (2011)</u>

Agents choose whether to predict the endog vars from u_t or from r_t^n. The metric they use is the MSE of their fcst:

$$EU^j = -E(y_{t+j} - E_t^j y_{t+j})' W (y_{t+j} - E^j y_{t+j})$$

weighting mat, they use I

Then relative fcst performance, $F(n)$ is:

$$F(n) = EU^1 - EU^2$$

So EU is like FEV, or $fE' \cdot W \cdot fE$;

again the criterion is closer to CUSUM than COMP.

1.3) Milani (2014)

$$g_{t,y} = \begin{cases} \left(g_{t-1,y}^{-1} + 1\right)^{-1} & \text{if } \frac{\sum_{j=0}^{J}\left(|y_{t-j} - E_{t-j-1} y_{t-j}|\right)}{J} < v_t^y \\ \bar{g}_y & \text{else.} \end{cases}$$

$y = \pi, x, i$

→ i.e. dgain if the avg. of past absolute FEs J periods back is smaller than a threshold

The threshold v_t^y = mean absolute deviation of historical FEs, which is recursively updated.

So it feels like $v_t^y = \frac{\sum_{j=0}^{t_0}\left(|y_{t-j} - E_{t-j-1} y_{t-j}|\right)}{t_0}$

→ so again this is like the non-squared CUSUM

1.4) Cho & Kasa (2015)

Note that in previous models and I think my own work, agents are estimating parameters. When setting a gain, they believe there might have been a regime change, i.e. coefficients might have changed. But here, in Cho & Kasa (2015), agents are trying to choose the best-performing model (as in Branch & Evans actually!). So in a sense - at least from my standpoint right now -, this is less relevant - although this question of model uncertainty may become more relevant later.

They have a Lagrange-Multiplier (LM) score test, but it looks quite similar to the CUSUM test, in fact, they refer to it too!

$\Rightarrow$ so CUSUM (or FE-based tests) are the standard!

② Do fmincon w/ figures of loss for RE & learning for several configs of "internal parameters"

6 Feb 2020