

Work after

① proj. facility involved

i Projection facility isn't involved 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. Dam!

iii ↳ The above was for proj. facility w/o $\text{cig}(R)$.
(The "mini projection facility")

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

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

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

② $\det(\cdot)$

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

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

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

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

\Rightarrow the determinant is a good summary of the info a matrix comes.

crit-cusum with $\det[\tilde{\omega}' f f' - \theta_{t-1}]$

\rightarrow much more anchored. \uparrow this doesn't make sense b/c θ scalar
also $\det[\tilde{\omega}' f f'] - \theta_{t-1}$

\rightarrow much more anchored

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

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

scalar CUSUM:

initially anchoring \uparrow in Ψ_{π} , but
once $\Psi_{\pi} \geq 2$, anchoring \downarrow in $\Psi_{\pi} \uparrow$

CUSUM for scalar case: much more anchoring
(needs much lower $\bar{\theta} = 0.00005$)

\hookrightarrow 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 $\Psi_{\pi} \approx 1.5$.

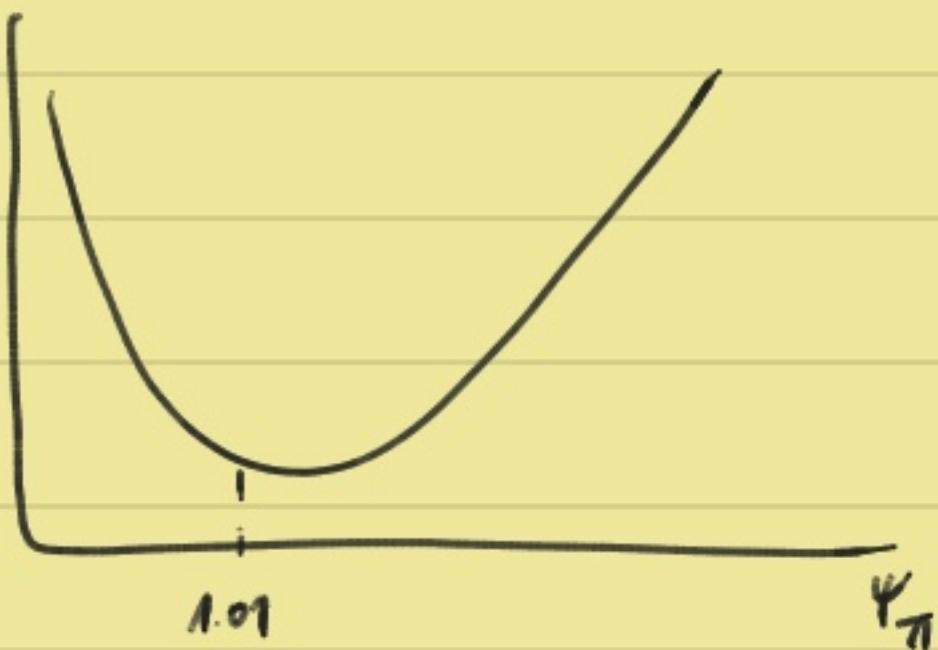
CUSUM for vector case: much more anchoring
(also needs like $\bar{\theta} = 0.00005$)

Γ seems to make it less monotonous.

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

\Rightarrow so the 12 seems to add monotonicity: w/o it,

anchoring



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

I wanna go back to the Woodford thing 30 Jun 2020
the "oni" := optimal nonindexed plan

So we have $\mathcal{L}^{\text{stat}} = \mathcal{L}_{\pi} + \gamma \alpha_x$

and we want $(f_{\pi}^{\text{on}}, f_x^{\text{on}}) = \underset{\text{argmin}}{\mathcal{L}^{\text{stat}}} \text{ st. (II) 8(IV)}$

Let's focus on \mathcal{L}_{π} :

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

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

$a' \times a$ where $f_a = a'$, x = longer uglier stuff (matrix)

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

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

so that suggests that

$$\begin{aligned} \frac{\partial \mathcal{L}_\pi}{\partial f_\pi} &= \frac{1}{1-\beta} \left[\left((I_{nx} - h_x h_x')^{-1} I_{nx} \right) + \left((I_{nx} - h_x h_x')^{-1} I_{nx} \right)' \right] f_\pi' \\ &- \left[\left((I_{nx} - h_x h_x')^{-1} (I_{nx} - \beta h_x h_x')^{-1} (h_x h_x') \right) + \left((I_{nx} - h_x h_x')^{-1} (I_{nx} - \beta h_x h_x')^{-1} (h_x h_x') \right)' \right] f_\pi \\ &= A_p \circ f_\pi' \quad \text{where} \\ &\quad 3 \times 3 \quad 3 \times 1 \end{aligned}$$

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

Since \mathbf{d}_x is symmetric to \mathbf{d}_{π} ,

$$\frac{\partial \mathbf{d}_x}{\partial f_x} = A_p \cdot f_x'$$

$$\Rightarrow \frac{\partial \mathbf{d}^{\text{stat}}}{\partial f_{\pi}} = A_p \cdot f_{\pi}' \quad \frac{\partial \mathbf{d}^{\text{stat}}}{\partial f_x} = \lambda \cdot A_p \cdot f_x'.$$

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

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

and (II) & (IV). → but then you only solve (II) & (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

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

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

$$\rightarrow f_x^{(0)} = \underset{x}{\operatorname{argmin}} \left(\frac{k}{1-\beta\rho} f_x + \frac{1}{1-\beta\rho} \right)^2 + \lambda f_x^2$$

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

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

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

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

Oh ok, so taking $\frac{\partial L^{\text{stat}}}{\partial f_i}$ 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 "a^{start}" conditional on each shock separately. And he doesn't consider shocks to the TR.

Let's work through this example.

Shocks:

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

$$u_t = p_u q_{t-1} + \epsilon_t^u \quad (2.18)$$

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

$$x_t = \bar{E}_t x_{t+1} - \beta [i_t - \bar{E}_t \pi_{t+1} - r_t^n] \quad (2.23)$$

Wait - I have two (maybe related) issues:

1.) What's the diff. between optimal policy and α_m ?

Woodford seems to suggest that α_m 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:

$$\pi_t = Kx_t + \beta E_t \bar{\pi}_{t+1} + u_t \quad (2.1)$$

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

You can sub in the TR and solve for $z = \bar{z} + f_2 s_t$

w/ $z = \begin{bmatrix} \pi \\ x \end{bmatrix}$ or don't sub it in & solve for $\bar{\pi} -$

w/ $z = \begin{bmatrix} \pi \\ x \\ i \end{bmatrix}$; it doesn't matter. Also it doesn't

matter if you write $z = \bar{z} + f_2 s_t$ or

$$\begin{array}{c} z = \bar{z} + f u_t + g r_t^n \left(\in h \cdot \bar{i}_t \right) \\ \text{m} \times 1 \quad n_y \times 1 \quad n_y \times 1 \quad \nearrow n_y \times 1 \end{array}$$

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 # egs is fine!

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

$$x_t = E_t x_{t+1} - \beta [\gamma_\pi \pi_t + \gamma_x x_t + \bar{i}_t - E_t \pi_{t+1} - r_t^n]$$

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

$$2\gamma_\pi \pi_t + (1 + 2\gamma_x) x_t = E_t x_{t+1} + \beta E_t \pi_{t+1} + \beta r_t^n - \beta \bar{i}_t \quad (2)$$

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

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

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,

$$(f_\pi, f_x, g_\pi, g_x, h_\pi, h_x)$$

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

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

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 $f(\text{eqs}) = 2$. (1) (1) (2) (2) (4)

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

→ Let $n_{eqs} := \# \text{ equations}$ (Addendum 31st fm)
 the number of equations to be solved in the end is
 the # FOCs once you've subbed in the constraints, i.e.

$$\text{Unknowns} - \text{Constraints} = n_y \cdot n_e - n_{eqs} \cdot n_e$$

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

$$\begin{aligned} \# \text{FOCs you'll use:} &= n_y \cdot n_e - (n_y - 1) \\ (n_y - n_{eqs}) \cdot n_e \Rightarrow \text{then} &= n_y(n_e - 1) + 1 \end{aligned}$$

$n_y = n_{eqs}$, the constraints fully determine the sol. $= 2(1-1) + 1 = 1 \text{ eq left.}$

$$\begin{aligned} \text{w/ the NKIS, } n_{eqs} = n_y \text{ so } &= n_y \cdot n_e - n_y \\ &= n_y(n_e - 1) \end{aligned}$$

If $n_e = 1$, thus = 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.

$$\begin{aligned} \text{If } n_{eqs} > n_y, \text{ e.g. } n_{eqs} = n_y + 1 &= n_y \cdot n_e - n_y - 1 \\ &= n_y(n_e - 1) - 1 \end{aligned}$$

→ Then we need $n_e \geq 2$ to have a solution at all!

What's the economics behind this?

Sug. that the case where $n_{egs} > n_y$ is not relevant
→ an econ. model wouldn't give this.

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

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

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

So what's the intuition behind the statement that
"when $n_{egs} = n_y$, the model needs at least 2 disturbances
for the om-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 one 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.

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

$$3\pi_t + (1-\beta\kappa_x)x_t = E_t x_{t+1} + 3E_t \pi_{t+1} + 3r_t^n - 3\bar{i}_t \quad (2)$$

Learning:

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

$$3\pi_t \pi_t + (1-\beta\kappa_x)x_t = E_t \sum_{T=t}^{\infty} \beta^{T-t} \left\{ (1-\beta-3\beta\kappa_x)x_{T+1} - 3(1-\beta\kappa_x)\pi_{T+1} + 3r_t^n - 3\bar{i}_t \right\} \quad (4)$$

Step 1

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

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

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. Define $\mathcal{L}^{\text{stab}} = \sum_{T=t}^{\infty} \beta^{T-t} [\text{var}_t(\pi_t) + \lambda \text{var}_t(x_t)]$

Given (5) & (6) will also yield the same $\mathcal{L}^{\text{stab}}$ for the RE & learning models.

Let's do RE in blue!

$$\begin{aligned}
 (1) \quad & \bar{\pi} + f_{\pi} r_t^n + g_{\pi} \bar{i}_t + h_{\pi} u_t - k (\bar{x} + f_x r_t^n + g_x \bar{i}_t + h_x u_t) \\
 & = \beta (\bar{\pi} + f_{\pi} r_{t+1}^n + g_{\pi} \bar{i}_{t+1} + h_{\pi} u_{t+1}) + u_t \\
 \Leftrightarrow \quad & \bar{\pi} - k \bar{x} - \beta \bar{\pi} = -f_{\pi} r_t^n - g_{\pi} \bar{i}_t - h_x u_t - k f_x r_t^n - k g_x \bar{i}_t - k h_x u_t \\
 & \quad + \beta (f_{\pi} f_r r_t^n + g_{\pi} p_i \bar{i}_t + h_{\pi} f_u \cdot u_t) + u_t
 \end{aligned}$$

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

$$\underbrace{(-f_{\pi} - k f_x + \beta f_{\pi} p_r) r_t^n}_{=0} + \underbrace{(-g_{\pi} - k g_x + \beta g_{\pi} p_i) \bar{i}_t}_{=0} + \underbrace{(-h_{\pi} - k h_x + \beta h_{\pi} p_u + 1) u_t}_{=0} = 0$$

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

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

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

$$(\beta p_r - 1) f_{\pi} = k f_x \quad (I) \quad \text{if stab!}$$

$$(\beta p_i - 1) g_{\pi} = k g_x \quad (II)$$

$$1 + (\beta p_u - 1) h_{\pi} = k h_x \quad (III)$$

Peter meeting

30 Jun 2020
→ or Hamilton's book

CUSUM: the matrix to scatter

Lütkepohl's book on VARs → "Intro to Multiple Time
"Multivariate CUSUM"

"Series Analysis"
or "New Intro to -II"

check:

✓ Kilian
or "Structural Vector Autocorgr"

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 $\text{eig}(A)$ only makes sense
for dynamic systems such as $X_{t+1} = A X_t$
→ then $\text{eig}(A)$ will tell you about dynamics.

But for $Y_t = g \times X_t$, $\text{eig}(g)$ will only tell you about the scaling, or the units in what you're measuring the vars.

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

Work after

Projection facility issue

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

- Murat & Sargent 1989 also specify a projection facility as a set restricted to be "close enough" to another.
 - Branchi unpublished p. 5 $K(r) = \text{closed ball of radius } r$ around the REE. If $(\phi, r) \in K(c_2) \setminus \{ \}$, 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)

overshooting, like Ramey says 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 PM is overparameterized.

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

! Best discussion is on p. 41 !

In RE, the 2nd set of constraints comes from:

31 Jan 2020

$$2\gamma_{\pi}r_{+} + (1+2\gamma_x)x_{+} = \bar{r}_{+}x_{++1} + 2\bar{r}_{+\pi_{++1}} + 2r_{+}^n - 2\bar{i}_{+} \quad (2)$$

stuff in the constrained sol

$$\begin{aligned} & 2\gamma_{\pi}(\bar{\pi} + f_{\pi}r_{+}^n + g_{\pi}\bar{i}_{+} + h_{\pi}u_{+}) + (1+2\gamma_x)(\bar{x} + f_xr_{+}^n + g_x\bar{i}_{+} + h_xu_{+}) \\ &= (\bar{x} + f_xr_{++1}^n + g_x\bar{i}_{++1} + h_xu_{++1}) + 2(\bar{\pi} + f_{\pi}r_{++1}^n + g_{\pi}\bar{i}_{++1} + h_{\pi}u_{++1}) \\ &\quad + 2r_{+}^n - 2\bar{i}_{+} \end{aligned}$$

↔

$$\begin{aligned} & 2\gamma_{\pi}(\bar{\pi} + f_{\pi}r_{+}^n + g_{\pi}\bar{i}_{+} + h_{\pi}u_{+}) + (1+2\gamma_x)(\bar{x} + f_xr_{+}^n + g_x\bar{i}_{+} + h_xu_{+}) \\ &= (\bar{x} + f_xp_{\pi}r_{+}^n + g_xp_{\pi}\bar{i}_{+} + h_xp_{\pi}u_{+}) + 2(\bar{\pi} + f_{\pi}p_{\pi}r_{+}^n + g_{\pi}p_{\pi}\bar{i}_{+} + h_{\pi}p_{\pi}u_{+}) \\ &\quad + 2r_{+}^n - 2\bar{i}_{+} \end{aligned}$$

Take the deterministic stuff on LHS:

$$\begin{aligned} \bar{\pi}(2\gamma_{\pi} - 2) + \bar{x}(2\gamma_x) &= r_{+}^n [2\gamma_{\pi}f_{\pi} + (1+2\gamma_x)f_x + f_xp_{\pi} + 2f_{\pi}p_{\pi} - 2] \\ &\quad + \bar{i}_{+} [2\gamma_{\pi}g_{\pi} + (1+2\gamma_x)g_x + g_xp_{\pi} + 2g_{\pi}p_{\pi} - 2] \\ &\quad + u_{+} [2\gamma_{\pi}h_{\pi} + (1+2\gamma_x)h_x + h_xp_{\pi} + 2h_{\pi}p_{\pi}] \end{aligned}$$

$$\Rightarrow 2(\gamma_{\pi} + p_{\pi})f_{\pi} + (1+2\gamma_x + p_x)f_x + 2 = 0 \quad (IV)$$

$$2(\gamma_{\pi} + p_{\pi})g_{\pi} + (1+2\gamma_x + p_x)g_x - 2 = 0 \quad (V)$$

$$2(\gamma_{\pi} + p_{\pi})h_{\pi} + (1+2\gamma_x + p_x)h_x = 0 \quad (VI)$$

So then RE amounts to solving:

$$\begin{array}{l} (\beta p_r - 1) f_{\pi} = k f_x \quad (I) \\ (\beta p_i - 1) g_{\pi} = k g_x \quad (II) \\ (\beta p_u - 1) h_{\pi} = k h_x - 1 \quad (III) \end{array} \quad \left| \begin{array}{l} b(\gamma_{\pi} + p_r) f_{\pi} + (1 + b\gamma_x + p_r) f_x + b = 0 \quad (IV) \\ b(\gamma_{\pi} + p_i) g_{\pi} + (1 + b\gamma_x + p_i) g_x - b = 0 \quad (V) \\ b(\gamma_{\pi} + p_u) h_{\pi} + (1 + b\gamma_x + p_u) h_x = 0 \quad (VI) \end{array} \right.$$

The good news is that these eggs seem to come in pairs:

so for (f_{π}, f_x) , solve (I) & (IV)

$$(I) f_x = \frac{\beta p_r - 1}{k} f_{\pi} \rightarrow \text{in (IV):}$$

$$b(\gamma_{\pi} + p_r) f_{\pi} + (1 + b\gamma_x + p_r) \frac{\beta p_r - 1}{k} f_{\pi} + b = 0$$

$$\Rightarrow \frac{bK(\gamma_{\pi} + p_r) + (1 + b\gamma_x + p_r)(\beta p_r - 1)}{k} f_{\pi} = -b$$

$$\Rightarrow f_{\pi}^{\text{one, RE}} = - \frac{bK}{bK(\gamma_{\pi} + p_r) + (1 + b\gamma_x + p_r)(\beta p_r - 1)} \quad (1)$$

$$f_x^{\text{one, RE}} = - \left(\frac{\beta p_r - 1}{K} \right) \frac{bK}{bK(\gamma_{\pi} + p_r) + (1 + b\gamma_x + p_r)(\beta p_r - 1)} \quad (2)$$

Analogously, (II): $g_x = \frac{\beta p_i - 1}{K} g_{\pi} \rightarrow \text{in (V)}$

$$b(\gamma_{\pi} + p_i) g_{\pi} + (1 + b\gamma_x + p_i) \frac{\beta p_i - 1}{K} g_{\pi} = b$$

$$[bK(\gamma_{\pi} + p_i) + (1 + b\gamma_x + p_i)(\beta p_i - 1)] g_{\pi} = K b$$

$$g_{\pi}^{\text{oni,RE}} = \frac{Kb}{2K(\gamma_{\pi} + p_i) + (1+b\gamma_x + p_i)(\beta p_i - 1)} \quad (3)$$

$$g_x^{\text{oni,RE}} = \left(\frac{\beta p_i - 1}{K} \right) \cdot \frac{Kb}{2K(\gamma_{\pi} + p_i) + (1+b\gamma_x + p_i)(\beta p_i - 1)} \quad (4)$$

And lastly, (4): $\frac{1}{K} + \left(\frac{\beta p_n - 1}{K} \right) h_{\pi} = h_x$

$$2(\gamma_{\pi} + p_n)h_{\pi} + (1+b\gamma_x + p_n) \left[\frac{1}{K} + \frac{\beta p_n - 1}{K} \right] h_{\pi} = 0$$

$$\Leftrightarrow \left[2(\gamma_{\pi} + p_n) + (1+b\gamma_x + p_n) \frac{\beta p_n - 1}{K} \right] h_{\pi} = - \frac{(1+b\gamma_x + p_n)}{K}$$

$$h_{\pi}^{\text{oni,RE}} = \frac{-(1+b\gamma_x + p_n)}{2K(\gamma_{\pi} + p_n) + (1+b\gamma_x + p_n)(\beta p_n - 1)} \quad (5)$$

$$h_x^{\text{oni,RE}} = \frac{1}{K} - \left(\frac{\beta p_n - 1}{K} \right) \frac{(1+b\gamma_x + p_n)}{2K(\gamma_{\pi} + p_n) + (1+b\gamma_x + p_n)(\beta p_n - 1)} \quad (6)$$

So that means that I have the Oni for $(\pi, x)^{\text{RE}}$. I can obtain

the Oni for i^{RE} by substituting $i^{\text{RE}} = \gamma_{\pi}\pi^{\text{oni}} + \gamma_x x^{\text{oni}} + i_+$,

to get

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

Ok cool, but now what?

I'm doing something wrong. I can see on Woodford's
omi, 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
 $\rightarrow L^{\text{stat}}$ is being used!
- 2) The Taylor-rule parameters (ψ_{π}, ψ_x) ($\rightarrow \phi_{\pi}, \phi_x$)

do not show up anywhere in the omi.
 \rightarrow the omi 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_t^*, ϕ_x^*) .

Model equations are

$$r_t^n = (1-p_t) \bar{r} + p_t r_{t-1}^n + \epsilon_t^{rx} \quad (2.27) \quad \left. \begin{array}{l} \text{2 exog} \\ n_e=2 \end{array} \right\}$$

$$u_t = p_u q_{t-1} + \epsilon_t^{uq} \quad (2.18) \quad \left. \begin{array}{l} \text{2 endog} \\ n_e=2 \end{array} \right\}$$

$$\pi_t = \alpha x_t + \beta \bar{\pi}_t \pi_{t+1} + u_t \quad (2.1) \quad \left. \begin{array}{l} \text{2 endog} \\ n_e=2 \end{array} \right\}$$

$$x_t = \bar{x}_t x_{t+1} - \beta [i_t - E_t \pi_{t+1} - r_t^n] \quad (2.23) \quad \left. \begin{array}{l} \text{2 eqs} \\ n_e=2 \end{array} \right\}$$

Conjoined LOM

$$y_t = \bar{y} + f_y u_t + g_y r_t^n \quad (2.6)$$

$$\text{where } y_t = \begin{pmatrix} \pi_t \\ x_t \\ i_t \end{pmatrix} \rightarrow n_y = 3$$

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

$$\# \text{unknowns} = n_y \cdot n_e = 3 \cdot 2 = 6$$

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 $L^{\text{stat}, r}$ & $L^{\text{stat}, u}$

(3.7) & (3.8) (p. 513) as given!

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

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

$$+ \beta (\bar{\pi} + f_\pi u_{t+1} + g_\pi r_{t+1}^n) + u_t$$

$$\bar{\pi} - \kappa \bar{x} - \beta \bar{\pi} = \underbrace{\kappa f_x u_t + \kappa g_x r_t^n}_{\text{from } \pi_t} - \underbrace{\beta f_\pi p_n u_t}_{\text{from } \pi_t} + \underbrace{\beta g_\pi [(1-p_r)\bar{r} + p_r r_t^n]}_{-f_n u_t - g_\pi r_t^n}$$

$$\text{RHS: } (\beta p_n f_\pi - f_\pi + \kappa f_x + 1) u_t$$

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

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

$$(1 - \beta p_n) f_\pi = \kappa f_x + 1 \quad (1)$$

$$(1 - \beta p_r) g_\pi = \kappa g_x \quad (2)$$

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

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

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

$$\begin{aligned}\bar{i} - \bar{\pi} &= -\underline{f_x u_t} - \underline{g_x r_t^n} + \underline{f_x p_t u_t} + g_x \left[(1-p_r) \bar{r} + \underline{p_r r_t^n} \right] \\ &= \underline{-b f_i u_t} - \underline{b g_i r_t^n} + \underline{b f_n p_t u_t} + b g_\pi \left[(1-p_r) \bar{r} + \underline{p_r r_t^n} \right] \\ &\quad + \underline{b r_t^n}\end{aligned}$$

$$\begin{aligned}3(\bar{i} - \bar{\pi}) &= (-f_x + f_x p_t - b f_i + b f_n p_t) u_t \\ &\quad + (-g_x + g_x p_r - b g_i + b g_\pi p_r + b) r_t^n \\ &\quad + g_x (1-p_r) \bar{r} + b g_\pi (1-p_r) \bar{r}\end{aligned}$$

$$3[\bar{i} - \bar{\pi} - g_\pi (1-p_r) \bar{r}] - g_x (1-p_r) \bar{r} = \text{RHS}.$$

Another WTF-term.

$$(u-1) f_x - b f_i + b p_t f_n = 0 \quad (3)$$

$$(p_r-1) g_x - b g_i + b p_r g_\pi + b = 0 \quad (4)$$

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

But somehow that means:

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

$$(1-\beta p_u)f_{\bar{\pi}} = \kappa f_x + 1 \quad (C1)$$

$$(1-\beta p_r)g_{\bar{\pi}} = \kappa g_x \quad (C2)$$

$$3\left[\bar{i} - \bar{\pi} - \underline{g_{\bar{\pi}}(1-p_r)\bar{r}}\right] - \underline{g_x(1-p_r)\bar{r}} = 0$$

$$(p_u - 1)f_x - 3f_i + 3p_u f_{\bar{\pi}} = 0 \quad (C3)$$

$$(p_r - 1)g_x - 3g_i + 3p_r g_{\bar{\pi}} + 3 = 0 \quad (C4)$$

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

Ok: do the following: given these (Cs), try to derive $f_{\bar{\pi}}^{\text{oni}}$

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

(C3):

$$(p_u - 1)\frac{(1-\beta p_u)f_{\bar{\pi}}}{\kappa} - \frac{(p_u - 1)}{\kappa} - 3f_i + 3p_u f_{\bar{\pi}} = 0$$

$$\frac{\kappa\beta p_u - (1-p_u)(1-\beta p_u)}{\kappa} f_{\bar{\pi}} - \frac{1-f_i}{\kappa} = \frac{3\kappa f_i}{\kappa}$$

$$f_i = \frac{(\kappa\beta p_u - (1-p_u)(1-\beta p_u))}{3\kappa} f_{\bar{\pi}} + \frac{1-p_u}{3\kappa}$$

Ok now plug these into $J^{\text{stab}, n} = f_{\bar{\pi}}^2 + \lambda_x f_x^2 + \lambda_i f_i^2$

$$f_{\pi}^{stab, n} = f_{\pi}^2 + \lambda_x \left(\frac{(1-\beta p_n)}{K} f_{\pi} - \frac{1}{K} \right)^2 + \lambda_i \left(\frac{(K \beta p_n - (1-p_n)(1-\beta p_n))}{B K} f_{\pi} + \frac{1-p_n}{B K} \right)^2$$

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

$$\rightarrow f_{\pi}^{stab, n} = f_{\pi}^2 + \lambda_x \left[\left(\frac{1-\beta p_n}{K} \right)^2 f_{\pi}^2 + \left(\frac{1}{K} \right)^2 - 2 \frac{(1-\beta p_n)}{K} f_{\pi} \right]$$

$$+ \lambda_i \left[\left(\frac{K \beta p_n - \delta_n}{B K} \right)^2 f_{\pi}^2 + \left(\frac{1-p_n}{B K} \right)^2 + 2 \frac{(K \beta p_n - \delta_n)(1-p_n)}{(B K)^2} f_{\pi} \right]$$

FOC for f_{π} :

$$2f_{\pi} + 2\lambda_x \left(\frac{1-\beta p_n}{K} \right)^2 f_{\pi} - 2\lambda_x \frac{(1-\beta p_n)}{K^2} + 2\lambda_i \left(\frac{K \beta p_n - \delta_n}{B K} \right)^2 f_{\pi}$$

$$+ 2\lambda_i \frac{(K \beta p_n - \delta_n)(1-p_n)}{(B K)^2} = 0$$

\Leftrightarrow

$$f_{\pi} \left[1 + 2\lambda_x \left(\frac{1-\beta p_n}{K} \right)^2 + 2\lambda_i \left(\frac{K \beta p_n - \delta_n}{B^2 K^2} \right)^2 \right] - 2\lambda_x \frac{(1-\beta p_n)}{K^2} + 2\lambda_i \frac{(K \beta p_n - \delta_n)(1-p_n)}{B^2 K^2} = 0$$

$$f_{\pi} \left[B^2 K^2 + \lambda_x B^2 (1-\beta p_n)^2 + \lambda_i (K \beta p_n - \delta_n)^2 \right] = \lambda_x B^2 (1-\beta p_n) - \lambda_i (K \beta p_n - \delta_n)(1-p_n)$$

$$f_{\pi} = \frac{\lambda_x B^2 (1-\beta p_n) - \lambda_i (K \beta p_n - \delta_n)(1-p_n)}{B^2 K^2 + \lambda_x B^2 (1-\beta p_n)^2 + \lambda_i (K \beta p_n - \delta_n)^2}$$

$$f_\pi = \frac{\lambda_x (1-\beta p_n) - \lambda_i (K \beta p_n - \gamma_n) (1-p_n) \beta^{-2}}{K^2 + \lambda_x (1-\beta p_n)^2 + \lambda_i (K \beta p_n - \gamma_n)^2 \beta^{-2}}$$

Call the denominator h_π

$$h_\pi = \lambda_i \beta^{-2} (\gamma_n - K \beta p_n)^2 + \lambda_x (1-\beta p_n)^2 + K^2 \quad \checkmark = \text{Woodford}$$

the numerator:

$$\lambda_i \beta^{-2} (\gamma_n - \beta p_n K \beta) (1-p_n) + \underbrace{\lambda_x (1-\beta p_n)}_{= \xi_n \text{ for } \text{Wm C}} \quad \checkmark = \text{Woodford}$$

Yes!

This $f_\pi = \text{Woodford's } \pi_n$ in App. F.4. (p. 713). \checkmark

Let's try to analogously solve for $g_\pi (= \pi_r)$ from $\Delta^{\text{stat}, r}$

$$\Delta^{\text{stat}, r} \propto g_\pi^2 + \lambda_x g_x^2 + \lambda_i g_i^2$$

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

$$(1) \quad (p_r - 1) g_x - 2g_i + 2p_r g_\pi + b = 0$$

$$\Leftrightarrow (p_r - 1) \left(\frac{1-\beta p_r}{\kappa} \right) g_\pi + 2p_r g_\pi - 2g_i + b = 0$$

$$(bK p_r - (1-p_r)(1-\beta p_r)) g_\pi + bK = bK g_i$$

$$g_i = 1 + (\beta K p_r - (1-p_r)(1-\beta p_c)) g_{\pi}$$

$$\text{So } \frac{\partial \text{stab}_i}{\partial r} = g_{\pi}^2 + \lambda_x \left(\frac{1-\beta p_r}{K} \right) g_{\pi}^2 + \lambda_i \left(1 + (\beta K p_r - \gamma_r) g_{\pi} \right)^2 \\ = \left[1 + \lambda_x \left(\frac{1-\beta p_r}{K} \right)^2 \right] g_{\pi}^2 + \lambda_i \left[(\beta K p_r - \gamma_r)^2 g_{\pi}^2 + 2(\beta K p_r - \gamma_r) g_{\pi} + 1 \right]$$

$$\text{FOC } g_{\pi}: 2 \left[1 + \lambda_x \left(\frac{1-\beta p_r}{K} \right)^2 + \lambda_i (\beta K p_r - \gamma_r)^2 \right] g_{\pi} + 2 \lambda_i (\beta K p_r - \gamma_r) \\ = 0$$

$$g_{\pi} = \frac{\lambda_i (\beta K p_r - \gamma_r)}{1 + \lambda_x \left(\frac{1-\beta p_r}{K} \right)^2 + \lambda_i (\beta K p_r - \gamma_r)^2} \quad \begin{matrix} \text{this doesn't seem} \\ \text{to be quite what} \\ \text{Woodford gets.} \end{matrix}$$

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:

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

At the same time, we have

$$i_r = \phi_{\pi} \pi_r + \phi_x x_r$$

so maybe we can do coeff comparison

$$\begin{aligned} i_r &= \phi_{\pi} (\bar{\pi} + \pi_r \cdot r_r^n + \pi_n \cdot u_r) + \phi_x (\bar{x} + x_r \cdot r_r^n + x_n \cdot u_r) \\ &= \underbrace{(\phi_{\pi} \bar{\pi} + \phi_x \bar{x})}_{\bar{i}} + \underbrace{(\phi_{\pi} \pi_r + \phi_x x_r)}_{i_r \cdot r_r^n} + \underbrace{(\phi_{\pi} \pi_n + \phi_x x_n)}_{i_n \cdot u_r} u_r \end{aligned}$$

→ I'm just surprised b/c if we only had i_r & i_n , then we could solve the following system for (ϕ_x, ϕ_{π})

$$\begin{array}{lcl} \phi_{\pi} \pi_r + \phi_x x_r = i_r & (1) & \\ \phi_{\pi} \pi_n + \phi_x x_n = i_n & (2) & \end{array} \quad \left. \begin{array}{l} 2 \text{ eqs in 2 unknowns} \end{array} \right\}$$

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

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

\uparrow

If this is also a param, then

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

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

$$\begin{aligned} (1-\beta) \bar{\pi} - \kappa \bar{x} &= 0 \\ 3[\bar{i} - \bar{\pi}] &= 0 \end{aligned} \quad \begin{cases} \bar{x} = \frac{1-\beta}{\kappa} \bar{\pi} \\ \bar{i} = \bar{\pi} \end{cases}$$

\hookrightarrow and so given (ϕ_{π}, ϕ_x) ^{ori} we can determine \bar{i}^{ori} .
(maybe)

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

$$\hat{r}_t^n = \rho r_{t-1}^n + \epsilon_{rt} \quad \text{w/ } r_t^n = (r_t^n - \bar{r}) \quad \Rightarrow E_t = \begin{pmatrix} r_t^n \\ u_t \end{pmatrix}'.$$

An idea: where is \hat{r}_+^n and where r_+^n ?

$$\pi_+ = Kx_+ + \beta \bar{\pi}_+ \pi_{++1} + u_+$$

$$x_+ = \bar{x}_+ x_{++1} - \beta [i_+ - E_+ \pi_{++1} - r_+^n]$$

$$\bar{\pi} + f_{\bar{\pi}} u_+ + g_{\bar{\pi}} \hat{r}_+^n = K(\bar{x} + f_x u_+ + g_x \hat{r}_+^n)$$

$$+ \beta (\bar{\pi} + f_{\bar{\pi}} p_u u_+ + \underbrace{g_{\bar{\pi}} \hat{r}_{++1}^n}_{g_{\bar{\pi}} p_r \hat{r}_+^n}) + u_1$$

$$\Rightarrow \underbrace{(1-\beta)\bar{\pi} - K\bar{x}}_{\equiv M1} = \underbrace{(-f_{\bar{\pi}} + Kf_x + \beta f_{\bar{\pi}} p_u + 1)u_+}_{\equiv C1} + \underbrace{(-g_{\bar{\pi}} + Kg_x + \beta g_{\bar{\pi}} p_r)}_{\equiv C2} \hat{r}_+^n$$

$$\bar{x} + f_x u_+ + g_x \hat{r}_+^n = \bar{x} + f_x p_u u_+ + g_x p_r \hat{r}_+^n - \beta (\bar{i} + f_i u_+ + g_i \hat{r}_+^n) \\ + \beta (\bar{\pi} + f_{\bar{\pi}} p_u u_+ + g_{\bar{\pi}} p_r \hat{r}_+^n) \\ + \beta (\hat{r}_+^n + \bar{r})$$

\Leftrightarrow

$$\underbrace{\beta(-\bar{\pi} - \bar{r} + \bar{i})}_{\equiv M2} = \underbrace{(-f_x + f_x p_u - \beta f_i + \beta f_{\bar{\pi}} p_u)}_{\equiv C3} u_+ \\ + \underbrace{(-g_x + g_x p_r - \beta g_i + \beta g_{\bar{\pi}} p_r + \beta)}_{\equiv C4} \hat{r}_+^n$$

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

M1: $\bar{x} = \frac{1-\beta}{K} \bar{\pi}$ and then obtain $(\bar{i}, \phi_{\bar{\pi}}, \phi_x)$ 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"
- Kilian & Lütkepohl, "SVAR Analysis" (\rightarrow pdf)
- Hamilton \rightarrow doesn't seem to be anything in it. \downarrow

Lütkepohl, "Multiple TS" this only says about Cusum & cusum-sq:
 \rightarrow "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

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 test at time T is $y_T(h)$.

The corresponding test error is

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

$(k \times 1)$ \hookrightarrow MA-representation

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

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

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

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

So since the FE \sim MVN w/ VC matrix $\Sigma_y(h)$, the statistic $\gamma_h := e_T(h)' \Sigma_y(h)^{-1} e_T(h) \sim \chi^2(k)$
i.e. the multiplication of two MVNs (scaled by VC)
is distib as χ^2 .

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

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

If $\gamma_h \geq$ critical value $\chi^2(k)$.

That's really cool but not feasible b/c it involves unknown quantities: the FE $e_T(h)$ 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 Littoral estimates:

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

$$\text{FEV} = \sum_{i=0}^{n-1} (\text{MA-coeff}_i) \sum_{\text{errors}} (\text{MA-coeff}_i)'$$

Ok, then you can use: where I'm wondering if he
 $\hat{\tau}_h = \hat{e}_T(h)' \hat{\Sigma}_y(h) \hat{e}_T(h)$ forgot the (-1) here
(very likely! I think it's a typo!)

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

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

$$\tilde{\tau}_h := \hat{e}_T(h)' \hat{\Sigma}_y(h)^{-1} \hat{e}_T(h) \cdot \frac{1}{k} \stackrel{\text{approx}}{\sim} F(k, T-kp-1)$$

where 2 changes:

- 1) for $\hat{\Sigma}_y(h)$ 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 adpated 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-sy test is like a mix of Cusum and Litterpol 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, Litterpol is describing a kind of "in spirit Cusum-test" for VARs.

My concern is that technically my \hat{Z}_t vector is not a VAR ... although in a sense it is b/c $\hat{Z}_t = \hat{g}_x^T S_t$ and S_t is a VAR(1). So I think the MNormality of $\hat{\epsilon}$ should still hold, and thus, my statistic $\hat{\chi}^2$ should also $\rightarrow \chi^2(k)$

Ok so compare then the critical values of

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

$|$
 $= n_y$

vs.

$$\tilde{\tau} = f' \tilde{\omega}^{-1} f / K \approx F(K, T - K - p - 1)$$

$|$
 $= n_y$ $|$
 $= n_y$ $|$
order of VIFR, = 1

and I'm not changing my estimate of the FEV
bc I already have it.

increases as agents' sample grows.

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

$$0 \rightarrow \infty$$

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

$$\chi^2(3) \quad 7.815$$

$$F(3, 1) \quad 2.157073$$

$$F(3, 6) \quad 4.7571$$

$$F(3, 120) \quad 2.6802$$

$$F(3, \infty) \quad 2.6045$$

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

But I still obtain that Cusum anchors more as $\gamma_0 \uparrow$
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 = $f(\gamma_0)$ non-monotonic:
 - for low values of γ_0 , lots of anchoring
 - for intermediate values of γ_0 , less anchoring
 - for high values of γ_0 , more anchoring again

- 2) Squaring makes the #anchoring = $f(\gamma_0)$ 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 $\gamma_0 \uparrow$, but for f , you lose this monotonicity.

Ok and why else square? b/c sum of squared norms

$$\text{is } \chi^2 : z_i \sim N(0, 1) \rightarrow Q = \sum_{i=1}^k z_i^2 \sim \chi^2(k)$$

↪ so if you don't square

goodness knows how θ_f is distributed then.

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

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

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

or math.hkbu.edu.hk (<http://math.hkbu.edu.hk/Math3806/Lecture-note3.pdf>)

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

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

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

But why does this behave opposite to Comp's?

- 3) A 3rd observation: for the Littlepohl-style criterion,
while $\chi_n \uparrow$ leads to more anchoring mandatorily, it
actually leads to less anchoring early in the sample
→ no analogy for the Comp criterion for this!

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

Gauth IRF

CAMP

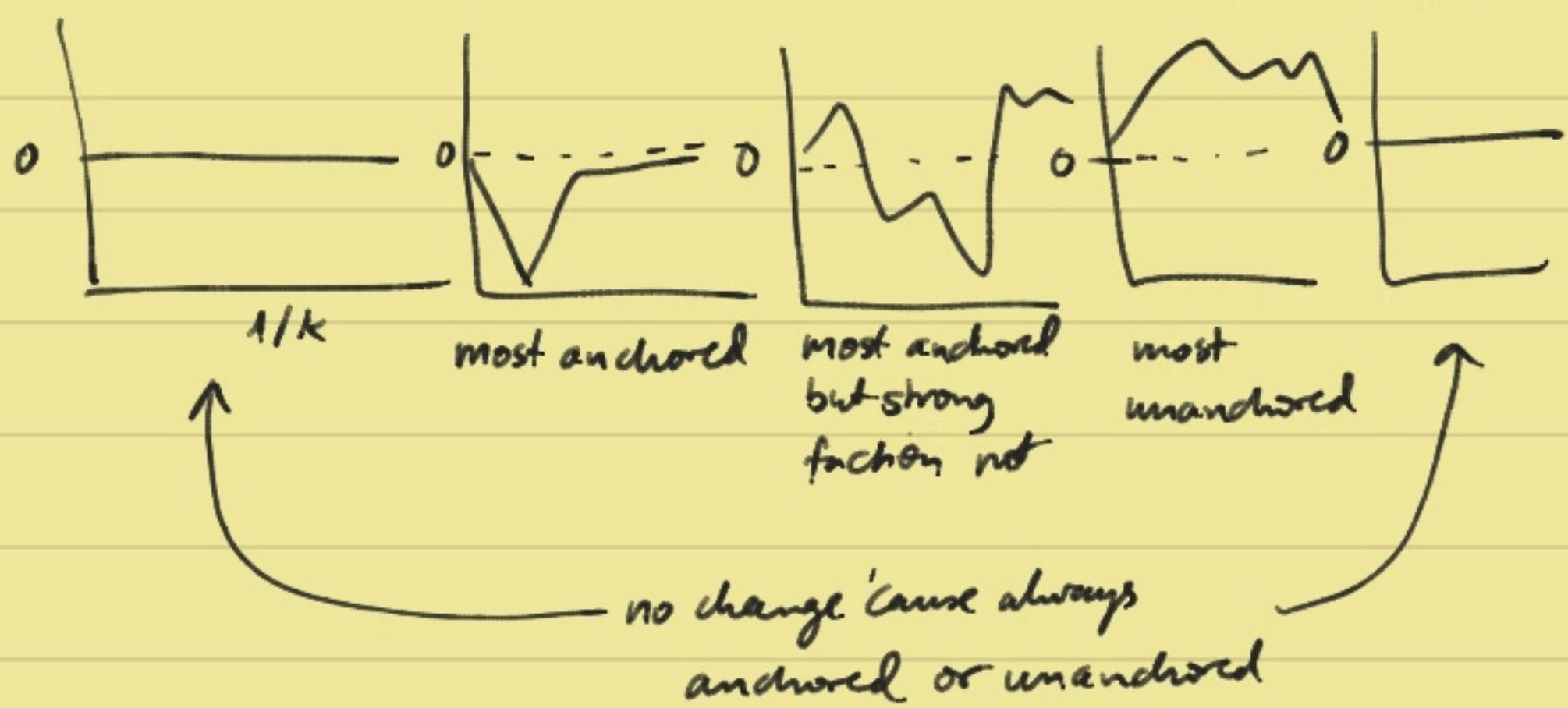
$$\Psi_{\pi} = 1.1$$

$$\Psi_{\pi} = 1.2$$

$$\Psi_{\pi} = 1.5$$

$$\Psi_{\pi} = 1.8$$

$$\Psi_{\pi} = 2.5$$



→ 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

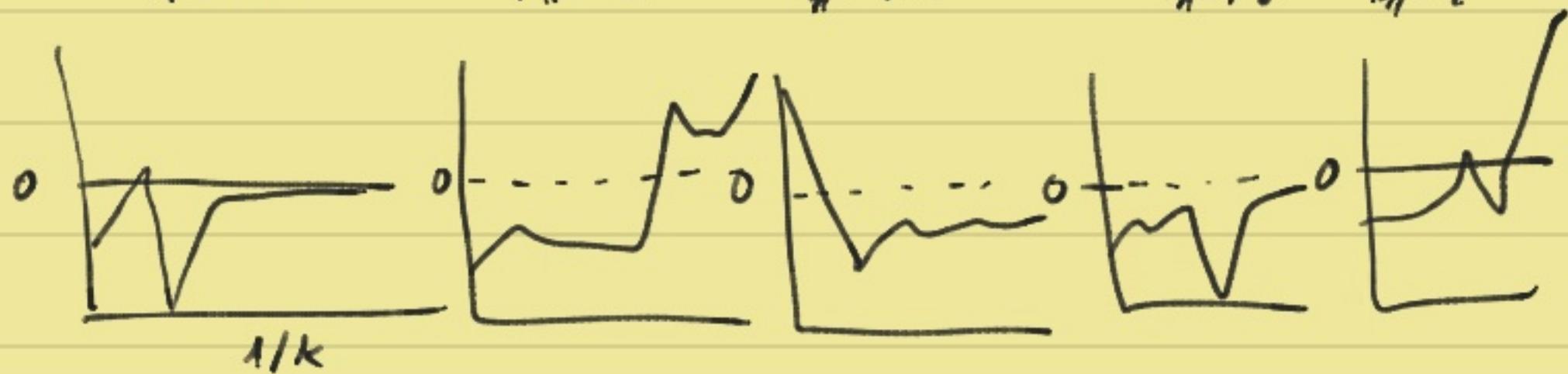
$$\Psi_{\pi} = 1.1$$

$$\Psi_{\pi} = 1.2$$

$$\Psi_{\pi} = 1.5$$

$$\Psi_{\pi} = 1.8$$

$$\Psi_{\pi} = 2$$



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.

$$\text{CEMP: } |\phi - [\hat{F}, \hat{G}]| > \bar{\theta} \quad \text{CUSUM: } f' \tilde{\omega}^{-1} f > \tilde{\theta}$$

$\sum_{t=1}^T$

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

\hat{F}, \hat{G} are functions of π_t expectations (as a function of ϕ)
 \rightarrow 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
forecast errors, so the CUSUM criterion becomes small!

\rightarrow got it!

Back to Woodford's ori

1 Feb 2020

The equations of RE were:

$$\underbrace{(1-\beta)\bar{i} - \kappa\bar{x}}_{\equiv M1} = \underbrace{(-f_{\pi} + \kappa f_x + \beta f_{\pi} p_n + 1)u_+}_{\equiv C1} + \underbrace{(-g_{\pi} + \kappa g_x + \beta g_{\pi} p_r)}_{\equiv C2}\hat{r}_+^n$$

$$\underbrace{3(-\bar{\pi} - \bar{r} + \bar{i})}_{\equiv M2} = \underbrace{(-f_x + f_x p_n - 3f_i + 3f_{\pi} p_n)}_{\equiv C3} u_+ + \underbrace{(-g_x + g_x p_r - 3g_i + 3g_{\pi} p_r + 3)}_{\equiv C4}\hat{r}_+^n$$

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

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

$$C1: (1-\beta p_n)f_{\pi} = 1 + \kappa f_x$$

$$C3: (1-p_n)f_x + 3f_{\pi}p_n = 3f_i$$

$$C2: (1-\beta p_r)g_{\pi} = \kappa g_x$$

$$C4: (1-p_r)g_x + 3g_{\pi}p_r + 3 = 3g_i$$

I solved $\min J^{stab, n}$ st. C1 & C3 $\rightarrow f_{\pi}^{ori}, f_x^{ori}, f_i^{ori}$

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

on Mathematica (matlab's) and got the same as Woodford.

What I still don't get is how Wooldridge 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!

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

(p. 50), slightly modified to include i)

Let's plug in the conjectures and $M1$ & $M2$

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

(shocks are mean zero)

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

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

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

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

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

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

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

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

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

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

$$\bar{\pi} = \frac{\lambda_x k^{-1} (1-\beta) \cancel{x} + \lambda_i (i^* - \bar{r})}{1 + (1-\beta)^2 k^{-2} \lambda_x + \lambda_i} \quad \checkmark \text{ Woodford yeah !!}$$

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 $\frac{\lambda_i (i^* - \bar{r})}{1 + \lambda_i} \dots$

2) If we were able to solve for $\bar{\pi}$ and thus for \bar{x} and \bar{r}

then coeff.-comparison in the Taylor-rule is over determined!

Somewhat Woodford's TR is:

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

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

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

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

$$\phi_{\pi} f_{\pi} + \phi_x f_x = f_i \quad T1$$

$$\phi_{\pi} g_{\pi} + \phi_x g_x = g_i \quad T2$$

Out... in Mathematica, solving the system $\begin{bmatrix} T1 \\ T2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

and setting $p_r = p_u = p$, I obtain EXACTLY what

Woodford gets for ϕ_{π}^* and ϕ_x^* in (3.12).

(onaterrah15.nb)

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

Step 1 Postulate convexities $z_t = \bar{z} + f_j u_t + g_j \hat{r}_t^n$

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

Step 2 Use the model equations (NK(S, NKPC)

and the LOMs of the two shocks to derive

M_1 & M_2 as constraints on $(\bar{\pi}, \bar{x}, \bar{i})$, and

C_1, C_2, C_3 and C_4 as constraints on $f_j, g_j \quad j = \pi, x, i$.

by plugging the conjectures into the model equations and using the LOMs of shocks to obtain expected future values of shocks.

Step 3. Solve 3 sets of optimizations

$$1. \min L^{\text{stat}} \text{ s.t. } M_1 \& M_2 \rightarrow \text{get } \bar{\pi} \rightarrow \bar{x}, \bar{i}$$

$$2. \min Y^{\text{stat}, \pi} \text{ s.t. } C_1 \& C_3 \rightarrow \text{get } f_{\pi}^{\text{ori}} \rightarrow f_x^{\text{ori}}, f_i^{\text{ori}}$$

$$3. \min L^{\text{stat}, x} \text{ s.t. } C_2 \& C_4 \rightarrow \text{get } g_x^{\text{ori}} \rightarrow g_{\pi}^{\text{ori}}, g_i^{\text{ori}}$$

Step 4. Compare coeffs of TR to $i_t = \bar{i} + f_i u_t + g_i \hat{r}_t^n$ to
Solve $T_1 = 0$ & $T_2 = 0$ for (ϕ_{π}^*, ϕ_x^*)

What changes / stays the same for the learning model.

- M_1 & M_2 , λ^{det} are the same, so $\bar{\pi}, \bar{x}, \bar{i}$ are the same.
- I'm going to assume identical stochastic processes to Woodford's, that is $\hat{r}_t^n = p_r \hat{r}_{t-1}^{n-1} + E_t^r \quad \hat{r}_t^1 = r_t - \bar{r}$

(For simplicity, I'll impose $p_u = p_r = p$ too!) $u_t = p_u u_{t-1} + E_t^u$

- I'm going to assume identical CB loss

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

as Woodford, so $L^{\text{det}}, L^{\text{stab}, n}, L^{\text{stab}, 1}$ are the same

(I don't even need to verify that his 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 \downarrow doesn't matter
 $i_t = \bar{i} + \gamma_\pi (\pi_t - \bar{\pi}) + \gamma_x (x_t - \bar{x}) \quad (14) \quad (3.1)$

- What changes is C_1, C_2, C_3, C_4 and thus the sols to

f_π, f_x, f_i and $g_{\pi i}, g_{x i}, g_i$

→ therefore the sols to $T1=0$ & $T2=0$ will involve
 $(\gamma_\pi^*, \gamma_x^*) \neq (\phi_\pi^*, \phi_x^*)$
 $\uparrow \text{Planning} \quad \uparrow \text{RE}$

Ok so learning C1-C4.

$$\pi_{t+1} - \alpha x_t = E_t \sum_{T=t+1}^{\infty} (\alpha \beta)^{T-t} \left\{ k \alpha \beta x_{T+1} + (1-\alpha) \beta \pi_{T+1} + u_T \right\}$$

$$x_{t+1} - \beta u_t = E_t \sum_{T=t+1}^{\infty} \beta^T \left\{ (1-\beta) x_{T+1} - \beta \beta u_{T+1} + \beta \pi_{T+1} + \beta r_T^n \right\}$$

Ignore the deterministic part to save space.

$$f_\pi u_t + g_\pi r_T^n - k f_x u_t - k g_x r_T^n = E_t \sum_{T=t+1}^{\infty} (\alpha \beta)^{T-t} \left\{ k \alpha \beta [f_x u_{T+1} + g_x r_{T+1}^n] + (1-\alpha) \beta [f_\pi u_{T+1} + g_\pi r_{T+1}^n] + u_T \right\} \quad (1)$$

$$f_x u_t + g_x r_T^n + \beta f_i u_t + \beta g_i r_T^n = E_t \sum_{T=t+1}^{\infty} \beta^{T-t} \left\{ (1-\beta) [f_x u_{T+1} + g_x r_{T+1}^n] - \beta [f_i u_{T+1} + g_i r_{T+1}^n] + \beta [f_\pi u_{T+1} + g_\pi r_{T+1}^n] + \beta r_T^n \right\} \quad (2)$$

\uparrow note: I'm

turning this into \hat{r}

b/c $r_T = \hat{r}_T + \bar{r}$,
take \bar{r} to LHS

(deterministic part)

$$(f_\pi - k f_x) u_t + (g_\pi - k g_x) r_T^n = E_t \sum_{T=t+1}^{\infty} (\alpha \beta)^{T-t} \left\{ [k \alpha \beta f_x + (1-\alpha) \beta f_\pi] u_{T+1} + [k \alpha \beta g_x + (1-\alpha) \beta g_\pi] r_{T+1}^n + \alpha \beta u_{T+1} \right\} + u_t \quad (1)$$

$$(f_x + \beta f_i) u_t + (g_x + \beta g_i) r_T^n = E_t \sum_{T=t+1}^{\infty} \beta^{T-t} \left\{ [(1-\beta) f_x - \beta \beta f_i + \beta f_\pi] u_{T+1} + [(1-\beta) g_x - \beta \beta g_i + \beta g_\pi] r_{T+1}^n + \beta \beta r_{T+1}^n \right\} + \beta r_T^n \quad (2)$$

\Rightarrow

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

$$(f_x + b f_i) u_t + (g_x + b g_i - b) \hat{r}_t^n = \frac{(1-\beta) h_x - b \beta f_i - b f_{\pi}}{1 - \beta p_u} u_t + \frac{(1-\beta) g_x - b \beta g_i - b g_{\pi} + b \beta}{1 - \beta p_r} \hat{r}_t^n \quad (2)$$

$$\Leftrightarrow \left[f_{\pi} - \kappa f_x - 1 - \frac{\kappa \alpha \beta f_x + (1-\alpha)\beta f_{\pi} + \alpha \beta}{1 - \alpha \beta p_u} u_t \right] =: C_1$$

$$+ \left[g_{\pi} - \kappa g_x - \frac{\kappa \alpha \beta g_x + (1-\alpha)\beta g_{\pi}}{1 - \alpha \beta p_r} \hat{r}_t^n \right] =: C_2 \stackrel{!}{=} 0 \quad (1)$$

and

$$\left[f_x + b f_i - \frac{(1-\beta) h_x - b \beta f_i - b f_{\pi}}{1 - \beta p_u} u_t \right] + \left[g_x + b g_i - b - \frac{(1-\beta) g_x - b \beta g_i - b g_{\pi} + b \beta}{1 - \beta p_r} \hat{r}_t^n \right] =: C_3 \stackrel{!}{=} 0 \quad (2)$$

$$=: C_4$$

just thinking in preparation
for meeting on Wed

4 Feb 2020

why do the CEMF vs. USUM criterion� have opposite ways? \Rightarrow is it really the case that the CEMF criterion internalizes U_1 expectations and USUM doesn't? I don't think so right now.

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

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

This isn't a first error.

This is the non-rational
first error.

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

and rational parts.

$$\hookrightarrow \phi_{t-1} - [F, G]$$

 $\phi^{\text{CEMF}} =$
 \uparrow
shouldn't I

re-evaluate next

wrong period? I do thought!

$$\begin{aligned} & \phi_{t-1} \begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix} - [F, G] \begin{bmatrix} 1 \\ s_{t-1} + \varepsilon_t \end{bmatrix} \\ &= \phi_{t-1} \begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix} - [F, G] \begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix} \\ &\quad - [F, G] \begin{bmatrix} 0 \\ \varepsilon_t \end{bmatrix} \\ &= \phi^{\text{CEMF}} - [F, G] \begin{bmatrix} 0 \\ \varepsilon_t \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
 & \text{So since } \theta^{\text{casum}} \approx f' \omega^{-1} f \\
 & = (\theta^{\text{comp}} - [F, G] \begin{bmatrix} 1 \\ \varepsilon_+ \end{bmatrix})' \omega^{-1} (\theta^{\text{comp}} - [F, G] \begin{bmatrix} 1 \\ \varepsilon_+ \end{bmatrix}) \\
 & \approx \frac{\uparrow (\theta^{\text{comp}})^2 + [F, G]^2 \begin{bmatrix} 1 \\ \varepsilon_+ \end{bmatrix}^2}{\omega} \quad \begin{array}{l} \text{var of shrubs,} \\ \Sigma, \text{constant} \end{array} \\
 & \quad - 2 \text{ cross-term}
 \end{aligned}$$

When $\gamma_\alpha \uparrow, \theta^{\text{comp}} \uparrow$

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

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

Instead of $\Sigma^{-1}(\phi - [F, G])$

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

Problem: $\tilde{\Sigma}$ is 4×4 .

Why? B/c $(\phi - [F, G])$ is 3×4

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

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

1) $\phi - [F, b]$ "discrepancy between RE & PMA coming from learning"

2) ϵ_+ noise

↳ having noise enter there is undesirable, but it's worse when it enters in loads

b/c it can dampen fast errors

(depending on the sign)

⇒ squaring is like biasing your results

up (\hat{f}_{OLSUM} is too high) but at least you know that it's not low due to the wrong relas.

⇒ you can measure the size of the bias

as the FER, ω , and then descale by that.

→ which is what OLSUM does!

(→ but that also means that the squared error is

being cleaned out. \rightarrow but does the squaring
of whatever's left make a diff?

Ok, looking at CEMP, what I've understood

\Rightarrow that $\sigma_{CEMP} = |\text{expected fcst error}|$
 $\qquad \qquad \qquad \swarrow$ knowing the model

i.e. $|\text{fcst error} - \text{noise}|$
 $\qquad \qquad \qquad \swarrow E(\text{noise}) = 0.$

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

But trying a version of CEMP squared (σ_{CEMP}^2)
still yields the same CEMP dynamics ($Y_n \uparrow$ and \downarrow)

\hookrightarrow It doesn't seem to be the squaring that makes
a difference. It's the concept of FE's.

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

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

\nearrow should at least
More in the same
direction as 0 \nearrow we took (w_{-1})
 \nearrow of this by w

OK: what I've tried is to write $\Sigma = \text{Var}(\epsilon_t) \stackrel{=}{\sim} \eta^2$
 into CUSUM and now $\psi_n \uparrow \rightarrow \text{anchoring!}$

↳ 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 FER?

Timmer: Wild:

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

So would my analogy to Littlepol's $\sum_y = \sum_{i=0}^{n-1} \phi_i \sum_{j=0}^{n-1} \phi_j$
 be $\sum_{i=0}^{n-1} h x_i \sum_{j=0}^{n-1} h x_j$?

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

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 $\hat{F}\hat{E}$.

- For CEMP: $\gamma_1 \uparrow$ raises FE b/c more expectation mistakes.

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

- If you \uparrow Var(noise), $\theta^{\text{CEMP}} \downarrow$ while $\theta^{\text{CUSUM}} \xrightarrow[\uparrow \text{ constant?}]{} \uparrow$
- 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 σ^{COMP} behavior!

↪ the difference is really due to $\frac{1}{\hat{F}\hat{E}^V}$.

Ignoring the constant,

$$\begin{aligned} E[FE] &= (\phi - [F, G]) s_{t-1} \quad FE = \phi s_{t-1} - [F, G] s_t \\ &= \phi s_{t-1} - [F, G] [e_{t-1} + \varepsilon_t] \\ &= (\phi - [F, G]) s_{t-1} - [F, G] \varepsilon_t \\ &= \sigma^{\text{COMP}} - [F, G] \varepsilon_t \end{aligned}$$

$$\rightarrow FE^2 = \sigma^{\text{COMP}}^2 + [F, G]^2 \varepsilon_t^2 - 2[F, G] \sigma^{\text{COMP}} \varepsilon_t$$

We're saying that this behaves similarly to σ^{COMP} .

Only once you divide by $\hat{F}\hat{E}^V$ do you change behavior.

$$\begin{aligned} F\hat{E}^V &= E[FE^2] = E[\sigma^{\text{COMP}}^2 + [F, G]^2 \varepsilon_t^2] \quad \text{b/c } E[\varepsilon_t] = 0 \\ &= E[FE] + [F, G]^2 \text{Var}(\text{noise}) \\ \rightarrow \hat{F}\hat{E}^V &= \sigma^{\text{COMP}}^2 + [F, G]^2 \text{Var}(\text{noise}) \end{aligned}$$

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

$$\phi_{\text{casum}} = \frac{1 - 2[F, G]\sigma^{\text{COMP}}\varepsilon_t}{\sigma^{\text{COMP}}^2 + [F, G]^2 \hat{\text{Var}}(\text{noise})}$$

$$\theta_{\text{census}} = 1 - \frac{2[F_{14}] \theta^{\text{COMP}} \varepsilon_t}{(\theta^{\text{COMP}})^2 + [F_{14}]^2 \text{Var}(\hat{\text{noise}})}$$

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

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

$$\begin{array}{ccc} \frac{1+2 \times 5}{1+2} & \xrightarrow{\quad \quad} & \frac{8}{3} = 2\frac{2}{3} \\ & \searrow & \\ & \frac{1+2}{1+2} + \frac{5}{3} = 1 + 1\frac{2}{3} & \end{array}$$

The cross-term $-[F_{14}] \theta^{\text{COMP}} \varepsilon_t$ is the discrepancy between FE^2 and $E[\text{FE}^2] = \text{FEV}$ (or its sample analog, $\hat{\text{FEV}}$). The point is that agents aren't able to evaluate the expectation $E[\text{FE}]$ or $E[\text{FE}^2]$ b/c they don't know the model. The next-best thing is FE^2 and $\hat{\text{FEV}}$ 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:

$$E[F\epsilon^2] = F\epsilon^2 - \text{cross term}$$

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

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

Questions for the analytical (γ^* , ϕ^*)

1) What's the diff between on and off-optimal?

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

3) $\Delta^{CB} = \dots \lambda(i_t - i^*)$ necessary?

4) γ^* doesn't depend on form 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 $E(\text{future policy})$ does 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)

⇒ timeline-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 first-best 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 history-dependent 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 anticipation 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 timellessly 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 timless optimality. This doesn't answer the question of what rule to use.
- it then seems that once you've given a rule of the form of a TR, you've already constrained yourself to purely forward-looking rules, so you're already "only 2nd best".
- but at least you can still be timellessly optimal and time-consistent within this class of rules

($\rightarrow L^{det} \& L^{stat}$) 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 $\{s_{t+1}, i\}$
- 3) I don't think it's necessary to have $\beta_i(i - i^*)$ in the loss if we can find a way to determine 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 its 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 nominal rule that
depends on how expectations evolve

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

sand 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 infset and now s_t is in the infset too. So couldn't you get a FE that subtracts ϵ_t and thus you're more aligned with CEMP's criterion?

⇒ in fact I think this is like saying

$$\theta^{\text{cusum}} - \left(-2 \theta^{\text{CEMP}} [F, b] \epsilon_t \right)$$

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

As for literature on endog. gain, there was 5%g...

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

2) Optimal TR - weiffs

It's a different thing to say that there's
an one where the LM-Expectations are taken
into account [i.e. a purely fwd-looking
(std TR) policy, but one which encompasses
how expectations evolve (\rightarrow the gain!)]
is a more optimal class of policy rule
that conditions on the state variable $E(\cdot)$.
(i.e. one that ain't one 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 $\pi_t^* = \bar{\epsilon} + F_{S_t}$,
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 results 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 CEMF- & CUSUM-criterion for different "internal parameters", potentially compared to RE would be good.

Ok so short-run to do:

① Compare Branch & Evans & Harrit & Nielson for endog gains.

② Do finincon w/ figures of loss for RE & learning for several configz of "internal parameters"

③ om w/ Lom-E(·).

① 11) Maret & Nicohi 2003

Perceived inflation = β_t

$$\beta_t = \beta_{t-1} + \frac{1}{\alpha_t} \left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right)$$

\uparrow
 gain \uparrow
 π_{t-1} $\curvearrowright =: FE_{t-1}$

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

\uparrow
 $= \bar{\theta} \text{ or } \tilde{\theta}$

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

\Rightarrow so this is an absolute FE scaled by the last belief

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

1.2.) Broadberry & Branson (2011)

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

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

\uparrow
weighting mat,
they use I

Then relative forecast performance, $F(u)$ is:

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

so EU is like FEV , or $FE^1 \cdot w \cdot fE$;

again the criterion is closer to causum than comp.

1.3) Milani (2014)

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

$y = \pi, X, i$

→ i.e. again if the avg. of past \check{V} FEs of periods back
is smaller than a threshold

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

$$\text{So it feels like } v_t^{\text{abs}} = \frac{\sum\limits_{j=0}^{t_0} (|y_{t-j} - E_{t-j-1} y_{t-j}|)}{t_0}$$

→ so again this is like the non-squared CUSUM

14) Ho & Kasa (2015)

Note that in previous models and I think my own work, agents are estimating parameters when setting a claim, they believe there might have been a regime change, i.e. coefficients might have changed. But here, in Ho & Kasa (2015), agents are trying to choose the best-performing model (as in Brandt & 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!

⇒ so CUSUM (or FE-based tests) are the standard!

② Do finincon w/ figures of loss for
RE & learning for several configs of "internal parameters" 6 Feb 2020

Done that and written a first very rough 10 Feb 2020
draft for Abogad. Now I'll talk to Rosen.

The next step anyway is to try to get a numerical
solution for the learning model.

I still think that for a property specified on,
which will look something like this

$$z_t = \bar{z} + F s_t + G \text{ expectations},$$

\bar{z} will still coincide w/ RE. In particular,
for $z_t = \bar{z} + f u_t + g r_t^n + h i_t + m \cdot k_t$
↑ gain

in the LR $(u_t, r_t^n, i_t, k_t) \rightarrow 0$, so $z_t = \bar{z} = \bar{z}^{\text{RE}}$

So maybe the only thing is to add the gain and try
to solve like that. So leave \bar{z} as it is.

Rosen meeting

10 Feb 2020

- Is $L^{\text{RE}} > L^{\text{team}}$ at $\gamma_t = \gamma_t^*$.
CB can exploit ppl's mistakes to achieve a better outcome
- Interpretation: they know the target but they have doubts whether the CB can implement it!

Optimal mon pol in Woodford's sense

this not to surprise agents

→ does the divine command still hold?

It should if the CB knows V , π d. $E(\cdot)$.

But what if it doesn't know $E(\cdot)$

↪ The function still goes in that direction

⇒ What loss can it achieve knowing $E(\cdot)$

vs. when it doesn't & just uses TR?

If there's a big gap, is there sth inbetween?

→ Survey of expectations should be included
people have argued this.

Would it be enough to know $LR-E(\cdot)$ instead of
 $E(\pi_{t+1})$?

① First ask

$$TR : \pi_t, x_t \rightarrow \gamma_1 \left(\hat{E}_t^i(\pi_t) - E_t(\pi_t) \right)$$

↑
what's γ_1^* ? $\neq 0$?

② "How feasible is this?"

Start w/ assuming the CB knows everything. Then
take a step away and ask what if the CB's
measure (of k , say) is imperfect? What can
it implement? What does it do optimally?
Can it approximate it using a simple TR, given
that it doesn't know $E(\cdot)$?