

## Work after

- 1) I wasn't computing the TC correctly b/c in the last  $H+1$  periods, should feed in  $hx^j s_t$  for shock  $\{s_t\}_{T-H}^T$
  - 2) If I don't impuse TR, should impuse TR w/o res, or w/o a res but then have a res in A1 or A2.
  - 3) Should check the fsolve thing.
- 2) Adding TR as a residual eq doesn't change much at all for  $\{\pi_t, x_t, i_t\}_t$ ! (stopped prematurely.) Making the norm the loss concept however gets me much closer to the Taylor rule!
- Adding TR as a res did bring things closer to the TR-outcome for  $\{x_t, i_t\}_t$  as input. Here the norm doesn't make a big diff. (stopped prematurely).
- Adding TR for  $\{i_t\}$  as only input. (local min but VERY unstable) For norm unstable too, stopped premat.
- But it no longer blows up in my face :)

1) Computing the anch TC w/ correct  $E(\cdot)$  doesn't change things. Does the norm? It finds a local min, but a strange one. (Both for  $\{i\}$  only.)

For  $\{x_i, i_t\}$ , norm or no norm makes no diff, but I'm in general much closer to TR than before, if I initialize there. If not, then further.

for  $\{\bar{x}_i, x_i, i_t\}$  as input, I get the same as before.  
Not if I initialize at TR, I get a strange local min.

With norm top.

↳ Either the method is unusable or I'm not quite doing it correctly

→ Tomorrow: do fsolve thing

✓. Fix H

8 April 2020

✓. Fix too long fightless

✓. Fix tsolve for TR & RE-TC  $\rightarrow$  fixed! (resids are not defined for  $t=1$ )  
- String is f' up w/ the anch-TC.



↳ Or: I don't know if the sol. procedure is unstable, or if I'm doing sthg wrong.

Back to VFI: Eric Sims does it by sneak' plug in the constraint to rewrite the problem as a fct of the future state:

We have:  $V^{l+1}(k_t) = \max_{c_t} u(c_t) + \beta V^l(k_{t+1})$   
s.t.  $k_t^\alpha - c_t \geq k_{t+1}$   
 $\rightarrow c_t = k_t^\alpha - k_{t+1}$

$$V^{l+1}(k_t) = \max_{k_{t+1}} u(k_t^\alpha - k_{t+1}) + \beta V^l(k_{t+1})$$

Rewrite in terms of  $k, k'$

$$V^{l+1}(k) = \max_{k'} u(k^\alpha - k') + \beta V^l(k')$$

Ok that's fine but I'm still - we need the grid for  $k$ , nonetheless! In the nonstochastic world, we just need no evaluation of expectations and possibly no markov chains

Btw, Eric Sims calls the choice of  $k'$  the policy fit.  
But I guess that's isomorphic w/ choosing  $\pi$ .

I've found Collard's value function iteration notes.  
(value-function-collard-lectnotes.pdf)  
→ It seems to suggest that you use the grid for  
 $k$  and  $k'$

for  $i = 1 : n_{\text{grid}}$

$$k_+ = \text{kgrid}(i)$$

for  $j = 1 : n_{\text{grid}}$  \* there's a complication, but ignore

$$k_{++1} = \text{kgrid}(j) \quad \text{for now}$$

$$c(i,j) = c(k_+, k_{++1}) \rightarrow u(i,j) = u(c(i,j))$$

end

$$j^* = \max[u(c(i,j)), 2]$$

$$k_{++1}^* = \text{kgrid}(j^*)$$

$$v^{\text{new}}(i) = u(c(i,j^*)) + \beta \cdot v(j^*)$$

and

$$\text{err} = \max(\text{abs}(v^{\text{new}} - v))$$

Actually Collard does this

$c(i,:)$  is  $c$  when  $k_+ = \text{kgrid}(i)$ , and for all  $k_{++1}$

$$\xrightarrow{1 \times n_{\text{grid}}} u_{\text{tilde}}(i,:) = u(c(i,:)) \quad 1 \times n_{\text{grid}}$$

$$[v^{\text{new}}, j^*] = \max \left[ u(c(i,:)) + \beta v(:) \right]$$

$n_{\text{grid}} \times 1$

\* the additional complication is that  $c_+, k_{t+1} \geq 0$

$$c_+ > 0 \Rightarrow c_+ = b_t^\alpha - k_{t+1} \geq 0$$

$$k_{t+1} \geq 0 \Rightarrow \boxed{b_t^\alpha \geq k_{t+1} \geq 0}$$

The grid makes sure that  $k_{t+1} \geq 0$ , but I need to check that  $k_{t+1} \nmid b_t^\alpha > k_{t+1}^{\text{old}}(i)$  for each  $i$ .

Collard's VFI notes are helpful b/c 9 April 2020

they clarify what the maximization means and what  $V(k_{t+1})$  means.  $V^{\text{old}}(k_{t+1})$  just means  $V^{\text{old}}(:)$  for all values of the kgrid ( $k_{t+1}$ ). Similarly, you evaluate  $c(k_{t+1}(i), :)$  for all values of the kgrid ( $k_{t+1}$ ).

The maximization then is just to choose the index

$$j^* \text{ that } \max U(c(k_{t+1}(i), :), \beta V^{\text{old}}(:)),$$

and that for each  $i$ . So  $k_{t+1}(i) = \text{kgrid}(j)$ .

$$\text{and } V^{\text{new}}(i) = U(c(k_{t+1}(i), j^*), \beta V^{\text{old}}(j^*))$$

## FSOLVE / FMINCON

↳ Could try to attack from several, or smart initialization points.

A potential problem for the anchoring TC:  
overparameterized?

I input  $T+H = 60$  periods, but I can evaluate only  $T$  residuals.

① Comparing FSOLVE vs. FMINCON

(1.1) Taylor-rule

Fmincon usually gets more solutions, even when fsolve says no sol.

(1.2) RE-TC

Same

(1.3) Anch-TC

Same.

I think this just means that fmincon is a little more easy-going w/ the sol. crit., since fsolve seeks to get  $f(x) = 0$ , while fmincon just tries to min  $F(x)$

FSOLVE never finds a sol (it stalls once:  
 $\{x, i\}_+$  in RE-TC, and init.)

FMINCON always finds at least 1 or 2 local mins.

But they are of 3 groups

- nonsensical: extremely volatile & huge
- really close to TR

↳ for TR & RE-TC  $\rightarrow$  makes sense

b/c you're either implementing the TR  
or the TR is a good implementation of  
RE-TC

- less volatile for  $x \& i$  and more for  $\pi$  for anch-TC

↳ anch-TC calls for a less aggressive TR!

I actually think I should prioritize the "approximating the reaction function" approach of Peter over the value function iteration b/c it's more promising in terms of results.

(But first I try reoptimizing by inputting the fid  $\{i^*\}$  as an initial guess for  $\{i_t\}$ )

→ doesn't work either.

command-approx-reaction.m

Problems:

✓  $g_{\pi}$  is blowing up → a problem for the previous exercise too.

- Besides, in previous ex., was still evaluating the and TC wrong b/c at each t, need to account for CB's  $E(\cdot)$  of future shocks

- loss isn't decreasing in any direction.

✓ forgot  $g_{\bar{\pi}}$ .

$$\frac{\partial}{\partial \bar{\pi}} \left( \frac{1}{(dfc)^2} \right) = \frac{\partial f_c}{\partial \bar{\pi}} \left( -2 \frac{1}{(dfc)^2} f_c^{-1} \right)$$

$$= 2 \frac{1}{(dfc)^2} f_c^{-1} \quad \leftarrow = g_{\bar{\pi}}^{-1}$$

The problem w/ this is that  $f_c^{-3} \rightarrow \infty$   
if  $f_c$  is small.

I've inverted  $g_{\pi}$  and  $g_{\bar{\pi}}$  b/c in the  
materials, I've defined  $k_t^{-1} = g(\cdot)$ , while  
in the code I have  $k_t = g(\cdot)$ .

Next: need to be smart about anchoring for  
correct evaluation of anchTC in the positive/friction  
exercise. Regenerate.

$$f_c = \pi_t - (\bar{\pi}_{t-1} + b s_{t-1})$$

$$k_t = g(f_c) = \frac{1}{dfc^2}$$

$$k_t^{-1} = g(f_c)^{-1} = dfc^2$$

$$\frac{\partial k_t^{-1}}{\partial \pi_t} = 2dfc = : g_{\pi} \quad \text{vs. } g_{\bar{\pi}}^{-1}$$

$$\frac{\partial k_t^{-1}}{\partial \bar{\pi}_{t-1}} = -2dfc = : g_{\bar{\pi}} \quad \text{vs. } g_{\bar{\pi}}^{-1}$$

- Still loss isn't decreasing
- Still TC eval is wrong for previous.  
 ↳ I think I fixed that now - it takes way longer to run → 25-30 min?  
 → Took 17 minutes! (47 iter)  
 Or up to 30!

While it's running, let's work thru Bellard's VPI

$$u(c) = \frac{c^{1-\beta} - 1}{1-\beta}$$

$$k' = k^\alpha - c + (1-\delta)k$$

$$\Rightarrow V(k) = \max_c \frac{c^{1-\beta} - 1}{1-\beta} + \beta V(k'), \text{ or,}$$

plugging in  $u_m(k)$  for  $c$ ,

$$V(k) \max_{k'} \frac{(k^\alpha + (1-\delta)k - k')^{1-\beta} - 1}{1-\beta} + \beta V(k')$$

Let's make sure that for no grid value of  $k(i), k'(j)$   
 is  $c < 0 \rightarrow$  for any  $k(i), k(i)^\alpha + (1-\delta)k(i) - k'(j) \geq 0$

I think I'm still evaluating the TC wrong 10 April 2020  
- in all exercises, b/c if you right zero innovations,  
that's gonna change  $f_A$  &  $f_B$  too. No - it's  
correct b/c those are  $f_B$  &  $f_A$  that the CB expects  
people to have. Pew!

Now I have the problem that `sim_learnLTI.m`  
doesn't work. It doesn't produce IRFs, and the  
simulation doesn't seem to converge to RE either.

Sorry - it was b/c I turned  $\beta$  &  $\rho$  of shocks to 0.  
IRFs are working.

And sorry, it is converging, I plotted  $\epsilon_{t+1}$   
and it does.

I'm cleaning up!

`sim_learnLTI.m` is no longer touched!

`sim_learnLTI-clean.m` = `simLearnLTI.m` and is used  
as a basis for subsequent work. (`matrices24.m`)

(combines the two.)

sim-lancLH-cleanSmooth.m is meant for  
the smooth anchoring function only!

Ryan meeting

10 April 2020

→ should get a resid of zero if init at Taylor  
rule sequence

→ feed fsolve the T residuals and not the  
square but the level  
whereas fmincon will need the square.

Don't use fmincon b/c it takes lots of info and  
since it uses a scalar

lsqnonlin

↳ similar to fmincon

Fodre is the way to go when you know a  
resid of 0 exists. Else lsqnonlin b/c  
it won't hit 0.

fsolve

1<sup>st</sup> thing to check: # of exog vars off equations  
→ you can check that mistake in EE or  
in the last period

A brute force: if  $T=200$ ,  $H=100$ ,  
then at  $t=150$ , the  $\Delta L$  in 2<sup>nd</sup> line of (B.1)  
is fixed.

An implementable TC sounds interesting for a  
discussion. It can also be an option for  
this optimization. But it's not a concern  
for the CB unless communicating w/ the  
public is a concern.

## Work after

- So back to `fsolve`: need to find in the "implemented the TR" setup where the mistake lies.
- just some thoughts:

Ryan seems to think that Peter's approach is kinder "as-backwards" (he said "circular", I think)

- why not max the (-loss) directly?
- ⇒ which makes me wonder if I can solve for the optimal  $\{i_t\}$  w/o value fit iteration?

I mean,  $\min \sum (\bar{a}_t^2 + \gamma_x x_t^2)$

s.t. model equations?

→ Why bother w/ VFI???

Hmm, hmm, hmm m...!

## Back to forth

It's the NKIS that isn't satisfied if you input  $\{T_t, x_t, i_t\}$  or  $\{x_t, i_t\}$ .

If you input only  $\{i_t\}$ , it's the Taylor rule residuals that explode at the end, just as  $\{x_t, \pi_t\}$  explode.

$\hookrightarrow$  I mean, in this case A9 & A10 have to hold, b/c they're used to generate  $\{x_t, T_t\}$ . And they do. But they lead to  $\{x_t, \pi_t\}$  exploding, which is why the Taylor-rule doesn't hold.

$$\{i_t^{\text{seq}}\} = \gamma_\pi \pi_t + \gamma_x x_t \quad \hookrightarrow \infty \quad \hookrightarrow \infty$$

given  $\neq \infty$

$\Rightarrow$  I'm returning to thinking that it's the expectations...

$\hookrightarrow$  that's why it's NKIS that doesn't hold! B/c it's a function of  $f_b$ , which is the expectation that explodes a lot more!

But now I'm in the world where  $b = g_x \cdot h_x = RE$ , and  $a = \begin{bmatrix} \bar{\pi} \\ 0 \\ 0 \end{bmatrix}$  so  $f_a(2) = f_a(3) = f_\beta(2) = f_\beta(3) = 0$

Therefore the  $-2\beta f_b(3)$  shouldn't cause headaches. But if that's so, why are  $E(\cdot)$  exploding when we input the  $\{i_t^{TR}\}$ ?

So I'm really back to the idea that the model somehow doesn't think that the  $i_t^{TR}$  is in place when it's not actively used, so it's not  $E$ -stable.  
 $\rightarrow$  But this is kinda at odds w/  $f_a, f_b$  changing nicely the moment you input  $x_t$ , too.

Is it? Maybe not.

Input only  $i_t$ :  $\text{exp. } x_t = -2i_t + EC \quad \text{exp.}$

$$\text{exp. } \pi_t = K x_t + E(\cdot)$$

Input  $x_t, i_t$ :  $\text{here expectations seem disrupted somehow. (They converge slowly but yeah)}$

$$\text{res } x_t = -2i_t + EC \quad /$$

$$\pi_t = K x_t + E(\cdot)$$

$$i_t = K_a \pi_t + K_x x_t$$

Indeed, as soon as you add  $x_t$ ,  $\bar{\pi}$  doesn't explode anymore. What happens to  $i_t$ ?

$$\rightarrow i_t^{-1} \rightarrow 0.14475 \text{ (not zero) when input } \{i\}.$$

Hmm.  $\text{res}(\text{NKIS})$  seems to mirror  $\bar{\pi}$  in shape

Yes. Quite closely, except  $\text{res}(\text{NKIS}) \approx 100 \cdot \bar{\pi}$

$$\text{Err} \rightarrow f_b(1) \approx 100 \cdot \bar{\pi} = \frac{1}{1-\beta} \bar{\pi} (!)$$

Of course! I've set  $h_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$  ( $p_i = 0 \quad i=r^n, i, h$ )

$$\text{So } f_{a+} = \frac{1}{1-\alpha\beta} a_{t-1} = \frac{1}{1-\alpha\beta} \begin{bmatrix} \bar{\pi}_{t-1} \\ 0 \\ 0 \end{bmatrix} = 9.1743 \begin{bmatrix} \bar{\pi}_{t-1} \\ 0 \\ 0 \end{bmatrix}$$

$$f_b = \frac{1}{1-\beta} a_{t-1}$$

$$\Rightarrow \text{So } x_t = -i_t + \beta f_b(1) \leftarrow 0 \quad (\text{again b/c } h_x = 0)$$

and since  $\beta = 1$ ,

$$x_t = -i_t + 100 \cdot \bar{\pi}_{t-1}$$

$$\Rightarrow \text{which is why } \text{res}(\text{NKIS}) = \underline{x_t + i_t} = 100 \bar{\pi}_{t-1}$$

What about  $\bar{\pi}_t$ ?  $\underbrace{0.099}$

$$\bar{\pi}_t = \kappa x_t + (1-\kappa)\beta f_a(1) = \kappa x_t + 0.9083 \bar{\pi}_{t-1}$$

Supp we input  $\{x_t, i_t\}$ :

$$i_t = \{i_t^{TR}\} = \gamma_\pi \pi_t + \gamma_x x_t \rightarrow res = 0$$

$$x_t = -i_t + 100 \cdot \bar{\pi}_{t-1} \rightarrow res_{AG} = 100 \bar{\pi}_{t-1}$$

$$\pi_t = \kappa x_t + 0.9083 \bar{\pi}_{t-1} \rightarrow res_{AD} = 0.$$

Supp we input  $\{\bar{a}_t, x_t, i_t\}$

$$i_t = \{i_t^{TR}\} = \gamma_\pi \pi_t + \gamma_x x_t \rightarrow res = 0$$

$$x_t = -i_t + 100 \cdot \bar{\pi}_{t-1} \rightarrow res_{AG} = 100 \bar{\pi}_{t-1}$$

$$\pi_t = \kappa x_t + 0.9083 \bar{\pi}_{t-1} \rightarrow res_{AD} \text{ should be}$$

$$0.9083 \bar{\pi}_{t-1}$$

But it's not.

Moreover,  $res_{AG} \neq 100 \bar{\pi}_{t-1}$  exactly.

Inputting  $\{x_t, i_t\}$  vs.  $\{\bar{a}_t, x_t, i_t\}$  doesn't seem to make a diff? Can it be b/c  $\gamma_x < 0$ ? No. Setting  $\gamma_x = 0.5$  doesn't change things a lot.

→ Why isn't there a diff b/w inputting  $\{x_t, i_t\}$  and  $\{\bar{a}_t, x_t, i_t\}$ ?

Maybe the reason inputting  $\{\pi_t, x_t, i_t\}$  isn't diff from  $\{x_t, i_t\}$ , b/c we input the same  $\{x_t, i_t\}$  which is the case when  $\pi_t$  is endog. determined pins down the TR  $\pi_t$ , which coincides w/ the one I'm endog. inputting. So this should break if I initialize at random sequences?

→ Yes it does!

1) TR doesn't hold anymore.

2) NKPC holds if you input  $\{x_t, i_t\}$ , but no longer does when you input  $\{\pi_t, x_t, i_t\}$ . Which I think is exactly as it should be.

Now I'm back to inputting the Taylor-rule sequences.

All equations are fulfilled, except the TR.

If I input  $\{\pi_t, x_t, i_t\}$  or  $\{x_t, i_t\}$ , then the TR is fulfilled (by design), NKPC is either fulfilled by design or due to  $x_t$  being given, so it's always the eq. That's not

unfilled where  $E(\cdot)$  show up, or find an outlet.  
 The  $E(\cdot)$  are clearly not the same as in the TR -  
 simulation.

$\Rightarrow$  So somehow the same sequence isn't generating  
 the same expectations.

In particular,  $k$  is basically constant in the exog  
 sequence, while in the Taylor-mle sequence it's  
 basically decreasing.

So why is this happening?

$$\bar{\pi}_t = \bar{\pi}_{t+1} + k_t^{-1} \underbrace{\left( \pi_t - \bar{\pi}_{t+1} - b_1 s_{t+1} \right)}_{f\epsilon_{t|t-1}^{\text{eve}}}$$

$$k_t^{-1} = k_{t-1}^{-1} + g(f\epsilon_{t|t-1}^{\text{eve}})$$

$$\text{When } t=1, \quad k_1 = k_1^{\text{TR}} = \bar{g}^{-1}$$

$$t=2, \quad k_2 = k_2^{\text{TR}}$$

$$t=3, \quad k_3 > k_3^{\text{TR}} \Rightarrow \text{It must be that } f\epsilon$$

are different! I am suspecting that even a small  
 difference in  $f\epsilon$  will be amplified by  $g(\cdot)$ .

Strange - setting a dgain doesn't change anything, except things look quantitatively diff, but qualitatively, they're the same: they blow up.

So for dgain, if I input  $\{x_t, i_t\}$ ,  $\bar{\pi} = \bar{\pi}^R$   
 $k^{-1} = k^{-1}TR$ , and  $fe_{H+1}^{eve} = fe_{H+1}^{eve TR}$ .

and  $y = y^R$ !

But NKIS still isn't fulfilled!

Same for cgain!

If I only input  $\{i_t\}$ , then  $\begin{pmatrix} \bar{\pi} \\ k^{-1} \\ fe \end{pmatrix} \neq \begin{pmatrix} \bar{\pi} \\ k^{-1} \\ fe \end{pmatrix}^{TR} \rightarrow y \neq y^R$   
and TR is the one that  
isn't fulfilled.

You guys.. this is true for the smooth endog gain too!

But if E are the same, and y are the same, then  
why are NKIS residuals  $\neq 0$ ?

$\rightarrow$  Now I'm thinking again that it might be the A-

matrices. Let me check if the ones from Mathematica  
 (materials13 - time-baseline.nb) equal the ones  
 in materials2 (and 3&4):

$$A_a = \begin{bmatrix} \frac{(1-\alpha)\beta(1+b\gamma_x)}{w} & \frac{(1+b\gamma_x)ka\beta}{w} & 0 \\ -\frac{(1-\alpha)\beta b\gamma_\pi}{w} & -\frac{b\gamma_\pi ka\beta}{w} & 0 \\ \frac{(1-\alpha)\beta\gamma_\pi}{w} & \frac{\gamma_\pi ka\beta}{w} & 0 \end{bmatrix}$$

Should be

$$\frac{\gamma_\pi(1-\alpha)\beta(1+b\gamma_x) - \gamma_x(1-\alpha)\beta b\gamma_\pi}{w}$$

$$\text{num} = (1-\alpha)\beta\gamma_\pi [1+b\gamma_x - \cancel{\gamma_x b}]$$

should be

$$\frac{\gamma_\pi(1+b\gamma_x)ka\beta - \gamma_x b\gamma_\pi ka\beta}{w} = \frac{\gamma_\pi ka\beta(1+2\gamma_x - b\gamma_x)}{w}$$

$A_a$  is fine! What about  $A_b$ ?

$$A_B = \begin{bmatrix} \frac{\kappa b(1-\beta\gamma_\pi)}{\omega} & \frac{\kappa(1-\beta-\beta b\gamma_x)}{\omega} & 0 \\ \frac{\beta - \beta b\gamma_\pi}{\omega} & \frac{1-\beta-\beta b\gamma_x}{\omega} & 0 \\ \frac{\beta(1-\beta\gamma_\pi)(\gamma_x + \kappa\gamma_\pi)}{\omega} & \frac{(1-\beta-\beta b\gamma_x)(\gamma_x + \kappa\gamma_\pi)}{\omega} & 0 \end{bmatrix}$$

↑ num =  $\gamma_\pi \kappa b(1-\beta\gamma_\pi) + \gamma_x \beta(1-\beta\gamma_\pi)$

↑ num =  $\gamma_\pi \kappa(1-\beta-\beta b\gamma_x) + \gamma_x (1-\beta-\beta b\gamma_x)$

Ok and at this point I guess I'll just trust

that As is correct too.

Well, what I can do is if  $hx = 0$ , then  $ia = I_3 = ib$

then, according to materials<sup>2</sup>, As becomes

$$A_S = \left[ \begin{array}{cc} \frac{1+2\gamma x}{w} & \frac{Kb}{w} \\ -\frac{2\gamma q}{w} & \frac{b}{w} \end{array} \right]$$

are these switched?

$$1 - \frac{k_b Y_\pi}{w} = \frac{w - k_b Y_\pi}{w} = \frac{1 + b Y_x + \cancel{k_b Y_\pi} - k_b Y_\pi}{w}$$

1st d of last row:

$$Y_\pi \frac{1 + b Y_x}{w} + Y_x \left( -\frac{b Y_\pi}{w} \right) = \frac{Y_\pi + b Y_\pi Y_x - b Y_\pi Y_x}{w}$$

$$= \underline{\underline{\frac{Y_\pi}{w}}}$$

2nd d of last row

$$Y_\pi \frac{k_b}{w} + Y_x \left( \frac{b}{w} \right) \textcircled{+1} = \frac{k_b Y_\pi + b Y_x + w}{w} ?$$

$\hookrightarrow = 0$  if there's no mono. pol. shock.  $\rightarrow \frac{k_b Y_\pi + b Y_x}{w}$

3rd d of last row

$$Y_\pi \left( -\frac{k_b}{w} \right) + Y_x \left( -\frac{b}{w} \right) = -\frac{(k_b Y_\pi + b Y_x)}{w}$$

$\hookrightarrow$  then

$$A_S(s; \cdot) = \left[ \frac{Y_\pi}{w}, \frac{k_b Y_\pi + b Y_x}{w}, -\frac{(k_b Y_\pi + b Y_x)}{w} \right]$$

$A_S$  isn't quite right.

→ could this be the reason?

All right - let's assume iid shocks ( $h_x = 0$ ) and pencil & paper solve for  $A_a, A_b, A_s$ .

(A.9) & (A.10) becomes

$$x_+ = -bi_+ + [b \ 1-\beta \ -b\beta]f_b + b r_r^n$$

$$\text{depth } \pi_+ = K x_+ + [(1-\alpha)\beta, \kappa\alpha\beta, 0]f_a + u_+$$

$$=: \gamma i_+ = \gamma_\pi \pi_+ + \gamma_x x_+$$

$$\begin{bmatrix} 0 & 1 & b \\ 1 & -K & 0 \\ -\gamma_\pi & -\gamma_x & 1 \end{bmatrix} \begin{bmatrix} \pi_+ \\ x_+ \\ i_+ \end{bmatrix} = \begin{bmatrix} \text{stuff}_1 f_b + b r_r^n \\ \text{stuff}_3 f_a + u_+ \\ 0 \end{bmatrix}$$

maybe I should  
try replacing it  
w/mis

$$\begin{bmatrix} \pi_+ \\ x_+ \\ i_+ \end{bmatrix} = \begin{bmatrix} \frac{K}{w} & \frac{1+b\gamma_x}{w} & b \\ \frac{1}{w} & \frac{-b\gamma_\pi}{w} & \frac{b}{w} \\ \frac{\gamma_x + K\gamma_\pi}{w} & \frac{\gamma_\pi}{w} & \frac{1}{w} \end{bmatrix} \begin{bmatrix} \text{stuff}_1 f_b + b r_r^n \\ \text{stuff}_3 f_a + u_+ \\ 0 \end{bmatrix}$$

in mind!

$$\pi_+ = \frac{K}{w} \text{stuff}_1 f_b + \left( \frac{1+b\gamma_x}{w} \right) \text{stuff}_3 f_a + \frac{b}{w} b r_r^n + \left( \frac{1+b\gamma_x}{w} \right) u_+$$

$$x_+ = \frac{1}{w} \text{stuff}_1 f_b + -\frac{b\gamma_\pi}{w} \text{stuff}_3 f_a + \frac{b}{w} b r_r^n + -\frac{b\gamma_\pi}{w} u_+$$

$$i_+ = \frac{\gamma_x + K\gamma_\pi}{w} \text{stuff}_1 f_b + \frac{\gamma_\pi}{w} \text{stuff}_3 f_a + \frac{\gamma_x + K\gamma_\pi}{w} b r_r^n + \frac{\gamma_\pi}{w} u_+$$

Tomorrow: try to implement the TR-sequence using  
 ← that system in Matlab instead of the A-matrices!

Also check the inverse

$$\begin{bmatrix} 0 & 1 & b \\ 1 & -k & 0 \\ -\gamma_x & -\gamma_x & 1 \end{bmatrix}^{-1} = ? \quad \begin{bmatrix} \frac{k}{w} & \frac{1+b\gamma_x}{w} & -\frac{k^2 b}{w} \\ \frac{1}{w} & -\frac{b\gamma_x}{w} & \frac{b}{w} \\ \frac{\gamma_x + k\gamma_n}{w} & \frac{\gamma_n}{w} & \frac{1}{w} \end{bmatrix}$$

Now it's good.

No! Those two els are diff.

11 April 2020

$$\begin{bmatrix} \pi_+ \\ x_+ \\ i_+ \end{bmatrix} = \begin{bmatrix} \frac{k}{w} & \frac{1+b\gamma_x}{w} & -\frac{k^2 b}{w} \\ \frac{1}{w} & -\frac{b\gamma_x}{w} & \frac{b}{w} \\ \frac{\gamma_x + k\gamma_n}{w} & \frac{\gamma_n}{w} & \frac{1}{w} \end{bmatrix} \begin{bmatrix} \text{stuff}_1 f_b + \text{stuff}_2 \cdot s_t \\ \text{stuff}_3 f_n + \text{stuff}_4 \cdot s_t \\ 0 \cdot \text{stuff}_5 \cdot s_t \end{bmatrix}$$

$$A_a = \begin{bmatrix} \frac{1+b\gamma_x}{w} \\ -\frac{b\gamma_x}{w} \\ \frac{\gamma_n}{w} \end{bmatrix} [\text{stuff}_3]$$

almost equal to old Matlab stuff.

$$A_b = \begin{bmatrix} k/w \\ 1/w \\ \gamma_x + k\gamma_n \end{bmatrix} [\text{stuff}_1]$$

Not equal.  
 ↘ minus  
 ↗  $\frac{k^2 b}{w}$  stuffs

$$A_s = \begin{bmatrix} \frac{k}{w} \text{stuff}_2 + \frac{1+b\gamma_x}{w} \text{stuff}_4 \\ 1/w \text{stuff}_2 + -\frac{b\gamma_x}{w} \text{stuff}_4 \\ \frac{\gamma_x + k\gamma_n}{w} \text{stuff}_2 + \frac{\gamma_n}{w} \text{stuff}_4 \end{bmatrix}$$

↙  $\frac{b}{w}$  stuffs  
 ↗  $\frac{1}{w}$  stuffs

↙ Equal except dependence on it

Good news and a bad:

Good: w/ max "agril"-A-matrices, the TR sim & the exeq. sim coincide. The bad news is even the TR sim diverges  $\rightarrow$  of course! Blc w/o expectations internalizing a Taylor-principle, the model is not E-stable!

$\Rightarrow$  But overall this means that I need a single, unified way to derive the A-matrices, and I think my best bet is to do as few steps analytically as possible.

For both implementations, (A.9.) & (A.10) describes the DGP for  $x_t, i_t$

$$x_t = -\beta i_t + \text{stuff}_1 f_b + \text{stuff}_2 \cdot s_t$$

$$\pi_t = \kappa x_t + \text{stuff}_3 f_a + \text{stuff}_4 \cdot s_t$$

and for the TR implementation, it is given by a TR

$$i_t = \gamma_\alpha \pi_t + \gamma_x x_t \quad (\text{w/o a shock})$$

This is already what AGA10 uses to generate  $x_+$ ,  $\pi_+$  given inputs.

→ it's clearly saying that w/ this configuration of the model, even a Taylor-mle system isn't E-stable.

$$x_+ = -\beta i_+ + \text{stuff}_1 f_b + \text{stuff}_2 \cdot s_+$$

$$\pi_+ = k x_+ + \text{stuff}_3 f_a + \text{stuff}_4 \cdot s_+$$

$$i_+ = \gamma_\pi \pi_+ + \gamma_x x_+$$

$$\underbrace{\begin{bmatrix} 0 & 1 & 2 \\ 1 & -k & 0 \\ -\gamma_\pi & -\gamma_x & 1 \end{bmatrix}}_{=: \mathcal{G}} \begin{bmatrix} \pi_+ \\ x_+ \\ i_+ \end{bmatrix} = \begin{bmatrix} \text{stuff}_1 f_b + \text{stuff}_2 \cdot s_+ \\ \text{stuff}_3 f_a + \text{stuff}_4 \cdot s_+ \\ 0 \quad (\text{or } \text{stuff}_5 \cdot s_1 = [0 \ 1 \ 0] s_+) \end{bmatrix}$$

if you wanna include a manifold  
work

$$z_+ = \mathcal{G}^{-1} \cdot \underbrace{\begin{bmatrix} \pi_+ \\ x_+ \\ i_+ \end{bmatrix}}_{=: \mathcal{L} \text{ ("gimed")}} \quad \begin{array}{l} \text{matlab:} \\ \Rightarrow z_+ = \text{aleph-gimed.m} \end{array}$$

Ok, btw now I've corrected  $\mathcal{G}$  on materials2S.nb and now it equals  $\mathcal{G}^{-1}$  from evaluating the inverse numerically. (see red corrections)

Simulation using aleph-gmrd.m yields EXACTLY the same thing when TR-5.m or exog-sim. But, both are not E-stable.

In fact, w/ the correction, so does simulation using matrices - A-25-time-baseline.m

↳ Ok, so this suggests that it's not just the earlier A-matrices are wrong per se, it's just that they embody a particular info-set.

I'm looking at materials13-time-baseline.nb

Now, w/ corrections, compare  $(A_a, A_b, A_s)$  w/  $(A_a, A_b, A_s)^{\text{apdl}}$

$$A_a = A_a^{\text{apdl}} \checkmark$$

$$A_b - A_b^{\text{apdl}} = \begin{bmatrix} * & 0 & * \\ * & 0 & * \\ * & 0 & * \end{bmatrix} \rightarrow \text{loadings of } z_4 \text{ on } f_b(1) \text{ and } f_b(3) \text{ are different}$$

$$A_s - A_s^{\text{apdl}} = \begin{bmatrix} 0 & * & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix} \rightarrow \text{loadings of } z_3 \text{ on } \hat{f}_s \text{ are diff.}$$

Let's deal w/  $A_S$  first, as I think this just comes from whether I include a mon. pd slant:

$$A_S - A_S^{\text{April}} = \begin{bmatrix} 0 & * & 0 \\ 0 & * & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \text{loadings of } z_4 \text{ on } \hat{z}_1 \text{ are diff.}$$

$\rightarrow$  yes, w/ the blue corrections

$$A_S = A_S^{\text{April}}$$

Ok so now to  $A_B$ .

$$A_B - A_B^{\text{April}} = \begin{bmatrix} * & 0 & * \\ * & 0 & * \\ * & 0 & * \end{bmatrix} \rightarrow \text{loadings of } z_4 \text{ on } f_B(1) \text{ and } f_B(3)$$

are different

$\hookrightarrow$  this should change if  $\gamma_x \neq 0$ .  $\rightarrow$  Exactly! It does!

$$\text{When } \gamma_x = 0.5, \quad A_B - A_B^{\text{April}} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

Ok, so now I know that the reason there's a diff between  $A_B - A_B^{\text{April}}$  is b/c of how  $E(\cdot)$  we treated.

$$\text{Stuff}_1 = [1 \ 0.01 \ -0.99] \quad (\text{w/ current params, } \gamma_x=0.5)$$

$$A_B^{\text{april}} = \begin{bmatrix} k/w \\ 1/w \\ \frac{\gamma_x + k\gamma_n}{w} \end{bmatrix} \text{stuff}_1$$

Ok. My conjecture is that for the old  $A_B$ ,  $\text{stuff}_1^{\text{old}}$  is diff:

$$A_B^{\text{old}} = \begin{bmatrix} k/w \\ 1/w \\ \frac{\gamma_x + k\gamma_n}{w} \end{bmatrix} \text{stuff}_1^{\text{old}}$$

$\text{stuff}_1 = [3 \ 1-\beta, -\beta\beta]$ . If we impose (\*) from my wifz-assumptions work in January, i.e. we impose

$$f_\beta(3) = \gamma_n f_\beta(1) + \gamma_x f_\beta(2), \text{ stuff}_1^{\text{old}} \text{ becomes } [3 - 3\beta\gamma_n, 1 - \beta - 3\beta\gamma_x, 0] = \text{stuff}_1^{\text{old}}$$

↪ and it's correct!

$\Rightarrow$  So now:

matrices\_A-25-one-baseline.m gives you the same as aleph-gamel.m. Whether they give you the same A-matrices as  
matrices\_A-13-one-baseline.m

depends on whether you impose that  $\text{stuff}_1 = \text{stuff}_1^{\text{old}}$ .  
w/  $\text{stuff}_1^{\text{old}}$ , you impose condition (\*), i.e. that  
agents know the Taylor-rule. In that case,  
the model is E-stable, and all equations are  
fulfilled. If you impose the new  $\text{stuff}_1$ , the  
model is NOT E-stable, but all equations are  
fulfilled.

Now the challenge is :

- ① it doesn't make sense to simulate a target TR simulation in which agents don't internalize the TR  $\rightarrow$  need to impose  $(t)$  for TR.
- ② but if for the exog-seq, I do not impose  $(t)$ , the equations won't be fulfilled.
- ③ But it doesn't really make sense to impose  $(t)$  when the CB isn't using a TR  $\rightarrow$  that amounts to assuming that agents believe the CB to follow a TR, while it's really just choosing the  $i$ -sequences.

$\Rightarrow$  Yet I feel ③ is my only option b/c in the target TR-sim, agents should know the TR and the model needs to be E-stable, yet for the exog-seq thing to work, the egs. need to match. So I am say: suppose the CB thinks agents believe in a TR-type regime, then what sequences to choose?

③ works when initialized at TR, and when initialized at random seq., except if you input TR as well. (then "stopped prematurely")

Let's just try diff. seeds for  $\text{ng}(0)$  to confirm that this is independent of initialization.  $\rightarrow$  Yup!

$\Rightarrow$  OK, so I know the fsolve machinery works now. But several challenges arise. Supp. I want to implement the RE-TC. Now the TR is NOT among the model equations. So I need to simulate the model w/o recourse to a Taylor-rule. But this is not expectationally stable unless I impose condition (\*) ( $\text{trig}, \text{ord}$ ). Right now I feel it's the only choice I have. But it seems like this problem has an economic interpretation which is interesting. It highlights the communication role of the Taylor rule (or potentially of any int rate rule);

I not only does it itself stabilize, but the majority of its effectiveness comes from its effect on  $U_1$  expectations, so that they become/remain anchored, and the model thus becomes E-stable.

→ ideally, my choice of optimal reaction function would need to reflect this as well.

I just implemented the numerical Ramsey. It seems like it finds sols (although if  $\lambda_x = 0$ , it doesn't seem to.)

It finds min at  $\lambda_x = 0.01$ , but it seems to take longer the smaller  $\lambda_x$ . It also seems to find the same seq every time, but incurs higher loss when  $\lambda_x$  is higher.

Let's initialize somewhere else. Even if we initialize at TR, we get the same thing.

And if we input  $\{x_q, i_1\}$ ? But that's not conceptually correct!  
No you get sthg diff.

Honestly, I don't know why bother w/ implementing the TC numerically when you can simply solve the Ramsey problem numerically.

If this and the "implement-and-TC" exercise are both correct, and my understanding is also correct, then I should obtain the same sequences w/ Ramsey & the "implement-and-TC" approach.

What I don't like:  $\pi$  &  $x$  don't move, even if  $i$  changes  $\rightarrow$  suggests to me that AGA10.m still isn't quite right. (Even if you input  $i = \text{ones}(\cdot)$ . But there is nothing in AGA10.m that "could" be wrong - I mean, there's barely anything in it!  $\Rightarrow$  things do move! Inputting  $\text{ones}(\cdot)$  vs. some sequence and computing  $y^{\text{ones}} - y^{\text{seq}} \neq 0$ . But the point is that  $\pi$  doesn't move a lot b/c  $\pi_t = kx_t = i_t + \epsilon(\cdot) \rightarrow$  whether you set  $\uparrow_{\text{small}}$

higher or smaller  $i_t$  is only a function of how much there's a concern for  $i$ -stabilization ( $\gamma_i$ ), since it doesn't move  $\pi_t$  by a lot, and moving  $\pi_t$  is the same as moving  $\theta_t$ , just by more.

→ but it seems then that the TR is optimal, or its optimality seems to stem from stabilizing  $E(\cdot)$ .

↳ I wonder if this obtains in vector learning.

I also worry about the optimal loss being huge.

Ok: tomorrow: try to implement RE-TC w/ the assumption of "TR-Expectations".

And think more this: why do  $E(\cdot)$  matter in a scalar learning environment? Is it that I change (how agents react to)  $\pi - E(\cdot)$  that makes the diff?

→ I think so!

A9 & A10:

12 April 2020

$$\text{A10: } x_+ = -\beta v_t + \underbrace{[3, 1-\beta, -\beta\beta]}_{\text{stuff}_1} f_b + \text{shocks}$$
$$+ \underbrace{[3 - 3\beta\gamma_\pi, 1 - \beta - 3\beta\gamma_x, 0]}_{\text{stuff}_1^{\text{all}}} f_b$$

$$\text{A10: } \pi_t = NK_+ + [(\alpha-\beta), \alpha\beta, 0] f_a + \text{shocks}$$

↑  $E(i)$  don't play a role

here, so  $i$ 's effect on  $E(\cdot)$  goes all thru the NKIS.

Suppose  $\gamma_x < 0$ . Then the difference between  $\text{stuff}_1$ ,  $\text{stuff}_1^{\text{all}}$  is that  $x_+$  depends on  $E(\pi)$  w/ 3 instead of

w/  $3 - 3\beta\gamma_\pi$ , and on  $E(i)$  w/  $-\beta\beta$  instead of 0.

→ I think this is the part that matters in scalar learning, b/c  $E(i) = b \cdot s_i$  anyway (so "fixed")

$$\text{So } x_+ = 3 f_\beta(\pi) \text{ vs. } x_+ = 3(1 - \beta\gamma_\pi) f_\beta(\pi)$$

The creepy thing is, this is the same thing that causes the oscillations b/c  $|3| > |3(1 - \beta\gamma_\pi)| > 0 > 3(1 - \beta\gamma_\pi)$

In vector learning, this wouldn't make a diff b/c  $E(\cdot)$  would converge to  $\gamma_\pi E(\pi) + \gamma_x E(x)$ .

Since  $\pi_t$  is just a scaled-down version of  $x_t$ ,  
the CB can only affect  $\pi_t$  by affecting  $x_t$  if more.  
And if  $x_t$  barely moves b/c it's based on  $E(\pi)$   
by a small (negative) number,  $\alpha(1-\beta\pi_t)$ , if  
 $E(\cdot)$  anchored, so gain is small, oscillations become  
very small,  $f_k(\pi)$  is small,  $E(\cdot)$  anchored, and  
moving  $i$  a lot doesn't have a huge effect,  
can't get  $E(\cdot)$  unanchored b/c  $\pi$  is moving so little  
that  $f_k$  is too small ( $\rightarrow$  need a better  $g$ -function).  
 $\hookrightarrow$  The point is that Taylor rule beliefs (together w/  
this form of the anchoring function) stabilize so much  
that the chosen  $i$ -sequence doesn't stir much  
water.  $\rightarrow$  That's why vector learning is more  
interesting b/c then a given  $i$ -sequence would  
allow (or not) Taylor rule beliefs to emerge.

$\rightarrow$  check work given vector learning.

The problem is, I haven't implemented  $g$  for vector learning. Also, the entire anchoring TC is derived for scalar learning; it doesn't even make sense for vector learning. But I could get an optimal  $t$ -sequence for vector learning if I generalize  $g$ .

Does it even make sense to continue w/ implementing the scalar anchTC?  $\rightarrow$  I don't think so b/c in that scalar case, if beliefs are not tied to  $t$ , they somehow, perhaps explode. But in a sense, this also has implications for the analytical TC. It should really just be seen as a characterization of what the sol. would look like in such a simple setting that otherwise doesn't make sense.

Ramsey i-sequence for vector learning and general g

→ sim-learn L1-clean-g.m (from sim-learn L1-clean.m)

This is as a first step just a g-function where I work w) k<sup>-1</sup> to start w/. Let me refer to it in the code as k1. (Scalar)

then implement vector

fk-smooth.m

this stuff still blows up in my face. It's

anonymous fct. not the right time to

$$k_t = k_{t-1} + g(\cdot)$$

$$k_t^{-1} = k_{t-1}^{-1} + g(\cdot)$$

Smooth to k<sup>-1</sup>.

where g() should be something like d · fk<sup>2</sup> + c

fk fk' → ny × ny   vs. fk' fk ⇒ 1 × 1  
ny × 1   1 × ny

$$\begin{bmatrix} fk_n \\ fk_x \\ fk_i \end{bmatrix} \begin{bmatrix} fk_n fk_x fk_i \end{bmatrix} = fk_n^2 \underline{fk_n fk_x} \underline{fk_n fk_i} \\ \underline{fk_n fk_x} fk_x^2 \underline{\underline{fk_x fk_i}} \\ \underline{\underline{fk_n fk_i}} \underline{\underline{fk_x fk_i}} k^2$$

$$\begin{bmatrix} fk_n fk_x fk_i \end{bmatrix} \begin{bmatrix} fk_n \\ fk_x \\ fk_i \end{bmatrix} = [fk_n^2 + fk_x^2 + fk_i^2] \quad \begin{matrix} \text{Not's fine} \\ \text{for now.} \end{matrix}$$

In fact, even the vector version isn't cool. Not the right time

→ But the good news is: for vector learning, we don't necessarily need a "vector-criterion" for  $\ell$ . It's fine if agents only evaluate  $f_c(\pi)$ , not  $f_c(z)$ .

Ok, right now sim team H - clean g. m

= sim/earnLH clean.m

and I've set it (-clear\_g) to use  $f(\pi)$  only to  
set  $k_+$ , but it's doing vector learning. Now  
try to input a sequence of  $i$  in first.

## Vector vs scalar learning

DEF

$$e_L = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \rightarrow e_L \neq \phi$$

Vector B. similar anchoring.

$$f_e = f_e(z_t) \quad vs \quad f_e = f_e(\bar{z}_t)$$

Now I need simLearnH-clean-given-seq  
→ ha! listen:

- simLearnH-clean-given-seq.m;
- set up w/ vector learning
- initialized at TR seq
- NIT setting stuff<sub>1</sub> = stuff<sub>1</sub><sup>old</sup>

\* drumroll \*

... doesn't explode!!

---

If you initialize it elsewhere, it explodes, c.p.  
What is worrisome though is that if you input  
anything else than the TR-i, it explodes. Even  
if you input  $i=1.6 \cdot Tt_t^{TR}$ . Oh I think I see why.  
B/c also that is an exog i-sequence but  
doesn't correspond to what  $\pi_i$  turns out to be in  
the simulation. But this means that inputting an  
exog. sequence is never gonna work unless it

coincides exactly w/ a TR that's used to generate that very sequence. The reason is that if you do the following (which I called "thingy" in `command-numerical-ramsey.m`):

- generate  $y^{TR}$  using the TR
- input in `sim-command-clean-given-seq.m`

(w/ vector learning, scalar anchoring)

anything else than  $\gamma_{\pi} \pi_t^{TR}$  ( $\gamma_x = 0$ )

→ then in the model, agents won't learn the TR blk in their world, other  $\pi_t \neq \pi_t^{TR}$  will be realized, so the relationship  $\gamma_{\pi} \pi_t$  won't hold!

Ok so what I've learnt in all this is: inputting exog i-sequence doesn't make sense if agents aren't learning the Taylor-rule, AND, if they're learning the TR, they won't learn the rule you're following if it's not conditioned on the  $\pi_t$  they're facing.

→ The point is → what makes sense to do is to find the optimal reaction-function for  $i$ , not an exoy sequence.

### Time-out

- 1) Why not simply min CB-loss over  $\text{exoy}\{i\}$ ?  
→ b/c  $\text{exoy}\{i\}$  will always explode, even w/ vector learning.
- 2) So why not simply min CB-loss over  $\gamma_{\pi}, \gamma_x, \gamma_{\pi^2}, \gamma_{x^2}$ ?  
→ I think I've already done that (at least over  $\gamma_{\pi}$ ) and it answers the wrong question: it answers: "given adhering to a TR-rule (or a higher-order thing), what are optimal params?"
- 3.) The right question is: "what's the optimal reaction-fct?"  
→ optimal means that it implements the target criterion. But: the TC is only valid for my scalar-learning and highly simplified anchoring fct setting. That's no good. So what I can do is a mix of 2) & 3)

- i) min CB loss over  $\gamma_{\pi_t}, \gamma_x$  in a vector learning world  
 (the latter to avoid the "They don't know the TR, so things explode or they do know the TR and so nothing matters" issue) (grid-search.m did this step in a constant-only, vector learning world with the USUM-crit)
- compute residual
- ii) min residual over  $\gamma_{\pi_t}, \gamma_x, \gamma_{\pi_t}^2, \gamma_x^2, \rho_{\pi x}$   
 keep going until residual is small enough.  
 Do this for other input candidates

$$\pi_{t+1}, \hat{E}_t(\pi_{t+1})$$

For simplicity, set  $\gamma_x = 0$ .

Some confusion is caused by the recursive nature of this process. Why can't I start w/ the most generous specification? B/c it's " $\infty$ "!

What can I do?

- I can evaluate the archTC for the TR to see how optimal the TR is
- I can min overall loss wrt. a very general TR and see how that does for the archTC.  
     $\hookrightarrow$  I wonder if the  $C_{loss} = \text{res}(\text{archTC})$ .

Skimble for today I'll prepare the presentation slides, maybe cont w/ Matlab in the evening.

I'm looking to see if one or two iterations on analytical VFI can't help me guess what form the optimal \*policy function\* for  $i$  should be.

$$\begin{aligned}& (-k\beta i_t + \Omega_4 \bar{\pi}_{t+1} + \Omega_5 s_t)^2 \\&= (\cancel{k\beta})^2 i_t^2 + \cancel{\Omega_4^2 \bar{\pi}_{t+1}^2} + \cancel{\Omega_5^2 s_t^2} \\&\quad - \cancel{k\beta \Omega_4 i_t \bar{\pi}_{t+1}} - \cancel{k\beta \Omega_5 i_t s_t} + \cancel{\Omega_4 \Omega_5 \bar{\pi}_{t+1} s_t}\end{aligned}$$

$\rightarrow$  Eqs (32) - (34) in materials25 describe the VFI problem.

13 April 2020

Suppose  $s_t = \begin{bmatrix} r_t \\ u_t \end{bmatrix} \sim \text{iid}$  so  $E_t(s_t) = 0$ .

Given:  $V^0 = 0$

Step 0.

$$V^1 = \max_{i_t} - \left\{ R_6 i_t^2 + R_7 \bar{\pi}_{t+1}^2 + R_8 i_t \bar{\pi}_{t+1} + R_9 i_t s_t + R_{10} \bar{\pi}_{t+1} s_t + R_{11} s_t^2 \right\}$$

s.t.

$$\bar{\pi}_t = \bar{\pi}_{t+1} + k_t^{-1} (-R_6 i_t + R_{12} \bar{\pi}_{t+1} + R_5 s_t - b_1 s_{t+1}) \quad (33)$$

$$k_t^{-1} = k_{t+1}^{-1} + g(i_t, \bar{\pi}_{t+1}, s_t, s_{t+1}) \quad (34)$$

$$\text{FDC} - 2R_6 i_t + R_8 \bar{\pi}_{t+1} + R_9 s_t = 0$$

$$i_t = \frac{R_8}{2R_6} \bar{\pi}_{t+1} + \frac{R_9}{2R_6} s_t \stackrel{\text{=: } P^1}{=} \text{(policy fct)}$$

I'm not sure if the constraints should be included here.

I think maybe only in the next step to sub out  $t+1$  states. Plug policy in  $V^1$ :

$$V^1 = - \left\{ R_6 \left( \frac{R_8}{2R_6} \bar{\pi}_{t+1} + \frac{R_9}{2R_6} s_t \right)^2 + R_7 \bar{\pi}_{t+1} + R_8 i_t \bar{\pi}_{t+1} + R_9 i_t s_t + R_{10} \bar{\pi}_{t+1} s_t \right\}$$

$$V_t^1 = - \left\{ \mathcal{L}_6 \left( \frac{\mathcal{L}_8}{2R_6} \bar{a}_{t-1} + \frac{\mathcal{L}_9}{2R_6} s_t \right)^2 + \mathcal{L}_7 \bar{\pi}_{t-1} + \mathcal{L}_8 i_t \bar{\pi}_{t-1} + \mathcal{L}_9 i_t s_t + \mathcal{L}_{10} \bar{\pi}_{t-1} s_t \right\}$$

$$= - \left\{ \mathcal{L}_6 \left( \frac{\mathcal{L}_8}{2R_6} \right)^2 \bar{\pi}_{t-1}^2 + \mathcal{L}_6 \left( \frac{\mathcal{L}_9}{2R_6} \right)^2 s_t^2 + \mathcal{L}_6 \frac{\mathcal{L}_8 \mathcal{L}_9}{(2R_6)^2} \bar{\pi}_{t-1} s_t \right.$$

$$\left. + \mathcal{L}_7 \bar{\pi}_{t-1} + \mathcal{L}_8 i_t \bar{\pi}_{t-1} + \mathcal{L}_9 i_t s_t + \mathcal{L}_{10} \bar{\pi}_{t-1} s_t \right\}$$

$$= - \left\{ \frac{\mathcal{L}_8^2}{4R_6} \bar{\pi}_{t-1}^2 + \frac{\mathcal{L}_9^2}{4R_6} s_t^2 + \mathcal{L}_7 \bar{\pi}_{t-1} + \mathcal{L}_8 i_t \bar{\pi}_{t-1} + \mathcal{L}_9 i_t s_t \right.$$

$$\left. + \left( \frac{\mathcal{L}_8 \mathcal{L}_9}{4R_6} + \mathcal{L}_{10} \right) \bar{\pi}_{t-1} s_t \right\}$$



Forgot to sub in policy for it here

$$= - \left\{ \frac{\mathcal{L}_8^2}{4R_6} \bar{\pi}_{t-1}^2 + \frac{\mathcal{L}_9^2}{4R_6} s_t^2 + \mathcal{L}_7 \bar{\pi}_{t-1} \right.$$

$$\left. + \mathcal{L}_8 \bar{\pi}_{t-1} \left( \frac{\mathcal{L}_8}{2R_6} \bar{a}_{t-1} + \frac{\mathcal{L}_9}{2R_6} s_t \right) \right.$$

$$\left. + \mathcal{L}_9 s_t \left( \frac{\mathcal{L}_8}{2R_6} \bar{a}_{t-1} + \frac{\mathcal{L}_9}{2R_6} s_t \right) \right.$$

$$\left. + \left( \frac{\mathcal{L}_8 \mathcal{L}_9}{4R_6} + \mathcal{L}_{10} \right) \bar{\pi}_{t-1} s_t \right\}$$

$$V^1(x_t) = - \left\{ \frac{\Omega_8^2}{4\Omega_6} \bar{\pi}_{t-1}^2 + \frac{\Omega_9^2}{4\Omega_6} s_t^2 + \Omega_7 \bar{a}_{t-1} \right.$$

$$+ \Omega_8 \bar{\pi}_{t-1} \left( \frac{\Omega_8}{2\Omega_6} \bar{a}_{t-1} + \frac{\Omega_9}{2\Omega_6} s_t \right)$$

$$+ \Omega_9 s_t \left( \frac{\Omega_8}{2\Omega_6} \bar{a}_{t-1} + \frac{\Omega_9}{2\Omega_6} s_t \right)$$

$$\left. + \left( \frac{\Omega_8 \Omega_9}{4\Omega_6} + \Omega_{10} \right) \bar{\pi}_{t-1} s_t \right\}$$

$$= - \left\{ \frac{\Omega_8^2}{4\Omega_6} \bar{\pi}_{t-1}^2 + \frac{\Omega_9^2}{4\Omega_6} s_t^2 + \Omega_7 \bar{a}_{t-1} \right.$$

$$+ \left( \frac{\Omega_8^2}{2\Omega_6} \bar{a}_{t-1}^2 + \Omega_8 \frac{\Omega_9}{2\Omega_6} \bar{\pi}_{t-1} s_t \right)$$

$$+ \left( \frac{\Omega_8 \Omega_9}{2\Omega_6} \bar{a}_{t-1} s_t + \frac{\Omega_9^2}{2\Omega_6} s_t^2 \right)$$

$$\left. + \left( \frac{\Omega_8 \Omega_9}{4\Omega_6} + \Omega_{10} \right) \bar{\pi}_{t-1} s_t \right\}$$

$$V^1(x_t) = - \left\{ \frac{3\Omega_8^2}{4\Omega_6} \bar{\pi}_{t-1}^2 + \frac{3\Omega_9^2}{4\Omega_6} s_t^2 + \Omega_7 \bar{\pi}_{t-1} + \left( \frac{5\Omega_8 \Omega_9}{4\Omega_6} + \Omega_{10} \right) \bar{\pi}_{t-1} s_t \right\}$$

Here the iid assumption comes in handy b/c  $E[s_{t+1}] = 0$   $E[s_{t+1}^2] = \Sigma$

$$E[V^1(x_{t+1})] = - \left\{ \frac{3\Omega_8^2}{4\Omega_6} \bar{\pi}_t^2 + \frac{3\Omega_9^2}{4\Omega_6} \Sigma + \Omega_7 \bar{a}_t \right\}$$

Now need to use constraints to sub out  $\bar{\pi}_t$

Step 1.

$$V^2(x_t) = \max_{i_t} - \left\{ R_6 i_t^2 + R_7 \bar{\pi}_{t-1}^2 + R_8 i_t \bar{\pi}_{t-1} + R_9 i_t s_t + R_{10} \bar{\pi}_{t-1} s_t + R_{11} s_t + \beta E_t V^1(x_{t+1}) \right\}$$

s.t. constraints

which is

$$V^2(x_t) = \max_{i_t} - \left\{ R_6 i_t^2 + R_7 \bar{\pi}_{t-1}^2 + R_8 i_t \bar{\pi}_{t-1} + R_9 i_t s_t + R_{10} \bar{\pi}_{t-1} s_t + R_{11} s_t \right. \\ \left. + \beta \left[ - \left\{ \frac{3R_8^2}{4R_6} \bar{\pi}_t^2 + \frac{3R_9^2}{4R_6} \sum + R_7 \bar{\pi}_t \right\} \right] \right\}$$

$$s.t. \quad \bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1} (-k b i_t + R_{12} \bar{\pi}_{t-1} + R_5 s_t - b_1 s_{t-1}) \\ + \beta \left[ - \left\{ \frac{3R_8^2}{4R_6} \bar{\pi}_t^2 + \frac{3R_9^2}{4R_6} \sum + R_7 \bar{\pi}_t \right\} \right]$$

$$s.t. \quad \bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1} (-k b i_t + R_{12} \bar{\pi}_{t-1} + R_5 s_t - b_1 s_{t-1})$$

$$k_t^{-1} = k_{t-1}^{-1} + g(i_t, \bar{\pi}_{t-1}, s_t, s_{t-1})$$

$$V^2(x_t) = \max_{i_t} - \left\{ R_6 i_t^2 + R_7 \bar{\pi}_{t-1}^2 + R_8 i_t \bar{\pi}_{t-1} + R_9 i_t s_t + R_{10} \bar{\pi}_{t-1} s_t + R_{11} s_t \right. \\ - \beta \left[ \frac{3R_8^2}{4R_6} \left( \bar{\pi}_{t-1} + (k_{t-1}^{-1} + g(\cdot))(-k b i_t + R_{12} \bar{\pi}_{t-1} + R_5 s_t - b_1 s_{t-1}) \right)^2 \right. \\ \left. + \frac{3R_9^2}{4R_6} \sum + R_7 \left( \bar{\pi}_{t-1} + (k_{t-1}^{-1} + g(\cdot))(-k b i_t + R_{12} \bar{\pi}_{t-1} + R_5 s_t - b_1 s_{t-1}) \right) \right]$$

$$V(x_t) = \max_{i_t} - \left\{ \begin{aligned} & \mathcal{R}_6 i_t^2 + \mathcal{R}_7 \bar{\pi}_{t+1}^2 + \mathcal{R}_8 i_t \bar{\pi}_{t+1} + \mathcal{R}_9 i_t s_t + \mathcal{R}_{10} \bar{\pi}_{t+1} s_t + \mathcal{R}_{11} s_t \\ & - \beta \left[ \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} \left( \bar{\pi}_{t+1} + (k_{t+1}^{-1} + g(\cdot))(-k_3 i_t + \mathcal{R}_{12} \bar{\pi}_{t+1} + \mathcal{R}_5 s_t - b_1 s_{t+1}) \right)^2 \right. \\ & \left. + \frac{3\mathcal{R}_9^2}{4\mathcal{R}_6} \sum_i + \mathcal{R}_7 \left( \bar{\pi}_{t+1} + (k_{t+1}^{-1} + g(\cdot))(-k_3 i_t + \mathcal{R}_{12} \bar{\pi}_{t+1} + \mathcal{R}_5 s_t - b_1 s_{t+1}) \right)^2 \right] \end{aligned} \right\}$$

FOC

i

$$-2\mathcal{R}_6 i_t - \mathcal{R}_8 \bar{\pi}_{t+1} + \mathcal{R}_9 s_t$$

$$\begin{aligned} & -\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} 2 \left( \bar{\pi}_{t+1} + (k_{t+1}^{-1} + g(\cdot))(-k_3 i_t + \mathcal{R}_{12} \bar{\pi}_{t+1} + \mathcal{R}_5 s_t - b_1 s_{t+1}) \right) \\ & \cdot \left( -k_3 k_{t+1}^{-1} - k_3 g(\cdot) - k_3 i_t \cdot g_i \right) \\ & - \beta \mathcal{R}_7 \left( -k_3 k_{t+1}^{-1} - k_3 g(\cdot) - k_3 i_t \cdot g_i \right) = 0 \end{aligned}$$

$$\Leftrightarrow -2\mathcal{R}_6 i_t - \mathcal{R}_8 \bar{\pi}_{t+1} + \mathcal{R}_9 s_t$$

$$\begin{aligned} & -2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} \left( k_{t+1}^{-1} + g(\cdot) \right) \left( -k_3 i_t \right) \left( -k_3 k_{t+1}^{-1} - k_3 g(\cdot) - k_3 i_t \cdot g_i \right) \\ & - 2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} \left( \bar{\pi}_{t+1} + (k_{t+1}^{-1} + g(\cdot)) \left( \mathcal{R}_{12} \bar{\pi}_{t+1} + \mathcal{R}_5 s_t - b_1 s_{t+1} \right) \right) \\ & \left( -k_3 k_{t+1}^{-1} - k_3 g(\cdot) - k_3 i_t \cdot g_i \right) \\ & + \beta \mathcal{R}_7 k_3 g_i i_t - \beta \mathcal{R}_7 \left( -k_3 k_{t+1}^{-1} - k_3 g(\cdot) \right) = 0 \end{aligned}$$

$$-2\mathcal{R}_6 i_t - \mathcal{R}_8 \bar{\pi}_{t-1} + \mathcal{R}_g s_t$$

$$-2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (k_{t-1}^{-1} + g(\cdot)) (-k_3 i_t) (-k_3 k_{t-1}^{-1} - k_3 g(\cdot) - k_3 i_t \cdot g_i)$$

$$-2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (\bar{\pi}_{t-1} - (k_{t-1}^{-1} + g(\cdot))) (\mathcal{R}_{12} \bar{\pi}_{t-1} + \mathcal{R}_5 s_t - b_1 s_{t-1}) \\ (-k_3 k_{t-1}^{-1} - k_3 g(\cdot) - k_3 i_t \cdot g_i)$$

$$+ \beta \mathcal{R}_7 k_3 g_i i_t - \beta \mathcal{R}_7 (-k_3 k_{t-1}^{-1} - k_3 g(\cdot)) = 0$$

$$\Leftrightarrow \underline{-2\mathcal{R}_6 i_t - \mathcal{R}_8 \bar{\pi}_{t-1} + \mathcal{R}_g s_t}$$

$$-2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (k_{t-1}^{-1} + g(\cdot)) (k_3)^2 g_i i_t^2$$

$$\underline{-2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (k_{t-1}^{-1} + g(\cdot)) (-k_3 k_{t-1}^{-1} - k_3 g(\cdot)) (-k_3 i_t)}$$

$$\underline{-2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (\bar{\pi}_{t-1} - (k_{t-1}^{-1} + g(\cdot))) (\mathcal{R}_{12} \bar{\pi}_{t-1} + \mathcal{R}_5 s_t - b_1 s_{t-1}) (-k_3 g_i) i_t}$$

$$\underline{-2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (\bar{\pi}_{t-1} - (k_{t-1}^{-1} + g(\cdot))) (\mathcal{R}_{12} \bar{\pi}_{t-1} + \mathcal{R}_5 s_t - b_1 s_{t-1}) (-k_3 k_{t-1}^{-1} - k_3 g(\cdot))}$$

$$\underline{+ \beta \mathcal{R}_7 k_3 g_i i_t - \beta \mathcal{R}_7 (-k_3 k_{t-1}^{-1} - k_3 g(\cdot)) = 0}$$

$$\Rightarrow 2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (k_{t-1}^{-1} + g(\cdot)) (k_3)^2 g_i i_t^2$$

$$+ \left[ 2\mathcal{R}_6 - \beta \mathcal{R}_7 k_3 g_i + 2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (k_{t-1}^{-1} + g(\cdot))^2 (k_3)^2 \right.$$

$$\left. - 2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (\bar{\pi}_{t-1} - (k_{t-1}^{-1} + g(\cdot))) (\mathcal{R}_{12} \bar{\pi}_{t-1} + \mathcal{R}_5 s_t - b_1 s_{t-1}) (k_3 g_i) \right] i_t$$

$$+ \left[ -\mathcal{R}_8 \bar{\pi}_t + \mathcal{R}_g s_t + \beta \mathcal{R}_7 (-k_3 k_{t-1}^{-1} - k_3 g(\cdot)) \right]$$

$$+ \left[ 2\beta \frac{3\mathcal{R}_8^2}{4\mathcal{R}_6} (\bar{\pi}_{t-1} - (k_{t-1}^{-1} + g(\cdot))) (\mathcal{R}_{12} \bar{\pi}_{t-1} + \mathcal{R}_5 s_t - b_1 s_{t-1}) (-k_3 k_{t-1}^{-1} - k_3 g(\cdot)) \right]$$

$= 0$       the policy function becomes a quadratic.

Before we go on the PL, let's interpret this policy:

$$\begin{aligned}
 & 2\beta \frac{3R_8^2}{4R_6} (k_{t-1}^{-1} + g(\cdot)) (k_0)^2 g_i i_t^2 \\
 & + \left[ 2R_6 - \beta R_7 k_0^2 g_i + 2\beta \frac{3R_8^2}{4R_6} (k_{t-1}^{-1} + g(\cdot))^2 (k_0)^2 \right. \\
 & \quad \left. - 2\beta \frac{3R_8^2}{4R_6} (\bar{\pi}_{t-1} - (k_{t-1}^{-1} + g(\cdot)) (R_{12}\bar{\pi}_{t-1} + R_5 S_t - b_1 S_{t-1})) (k_0 g_i) \right] i_t \\
 & + \left[ -R_8 \bar{\pi}_t + R_9 S_t + \beta R_7 (-k_0 k_{t-1}^{-1} - k_0 g(\cdot)) \right. \\
 & \quad \left. + 2\beta \frac{3R_8^2}{4R_6} (\bar{\pi}_{t-1} - (k_{t-1}^{-1} + g(\cdot)) (R_{12}\bar{\pi}_{t-1} + R_5 S_t - b_1 S_{t-1})) (-k_0 k_{t-1}^{-1} - k_0 g(\cdot)) \right] \\
 & = 0
 \end{aligned}$$

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad \text{Rightmost.}$$

But it is a function of

$$i_t = p(k_{t-1}^{-1}, g(\cdot), g_i, \bar{\pi}_{t-1}, S_t, S_{t-1})$$

which makes me think that the optimal reaction fit should be something like this:

$$i_t = r(\bar{\pi}_t, k_{t-1}^{-1}, \bar{\pi}_{t-1}, S_t, S_{t-1})$$

↑  
blk of FE.

$$r(\cdot)$$

interesting that both show up

↳ So one thing one could do next to value fit iter is

to consider a rule of the form

$$l_t = \gamma_{\pi} \pi_t + \gamma_k k_{t-1}^{-1} + \gamma_{\bar{\pi}} \bar{\pi}_{t-1} + \gamma_s s_t + \gamma_{sl} s_{t-1} \quad (1.7)$$

and maximize (-CB loss) over  $(\gamma_{\pi}, \gamma_k, \gamma_{\bar{\pi}}, \gamma_s, \gamma_{sl})$   
as well as higher orders or interactions.

The previous FOC (policy fit) suggests the following  
interactions :  $k_{t-1}^{-1} \pi_t, k_{t-1}^{-1} \bar{\pi}_{t-1}, k_{t-1}^{-1} s_t, k_{t-1}^{-1} s_{t-1}$   
 $\pi_t \bar{\pi}_{t-1}, \pi_t s_t, \pi_t s_{t-1}$ .

Ok, but let's turn to Lillard's VFI notes to see  
how I should discretize i) the iid shocks ii) the  
partly nonlinear LOMs : S

For the iid shocks, let's follow Peter in ass. that they  
take on only 'high' & 'low' values & are iid.

Then  $s_t = \begin{bmatrix} r_{high, Uhigh} \\ r_{high, Ulow} \\ r_{low, Uhigh} \\ r_{low, Ulow} \end{bmatrix}$

Supp for simplicity  
these values are  $\{-1, 1\}$ .

Then the state-space of the exog states is

$$\begin{Bmatrix} [1 \ 1] \\ [1 \ -1] \\ [-1 \ 1] \\ [-1 \ -1] \end{Bmatrix}$$

$$\rightarrow \text{sgrid} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$\rightarrow$  and a state  $i$  is  $\text{sgrid}(:, i)$ .

Let's also ass. that each of these states is equally likely, so  $\text{Prob}(\text{sgrid}(:, i)) = \frac{1}{4}$ . I'm not sure but I think that  $E(s_{t+1} | \mathcal{I}_{t-1}) = 0$  still.

Now to the endog. states.

Collard's summary of Tauden & Nursey seems to imply the following logic. For an AR(1) of the form

$$s_{t+1} = \rho s_t + (1-\rho)\bar{s} + \epsilon_{t+1} \quad \epsilon_{t+1} \sim N(0, \sigma^2)$$

Step 1. the likelihood is given by the Normal pdf/cdf  
( $\theta, s$ )

Step 2. Tanchen & Hussey replace the integral in the normal pdf/cdf by the quadrature

$$\frac{1}{\sqrt{\pi}} \sum_{j=1}^n w_j \bar{\Phi}(z_j, z_i, \bar{x})$$

$\nwarrow$  weights  $\uparrow$  quadrature nodes

where  $\bar{\Phi}(s_{t+1}, s_t, \bar{s}) = \exp \left\{ -\frac{1}{2} \left[ \left( \frac{s_{t+1} - ps_t - (1-p)\bar{s}}{\bar{s}} \right)^2 - \left( \frac{s_{t+1} - \bar{s}}{\bar{s}} \right)^2 \right] \right\}$

so that  $w_j \bar{\Phi}$  is interpreted as an estimate of the transition prob  $\text{Prob}(s_{t+1} = s_j | s_t = s_i)$ .

I think you get the weights using Gauss-Hermite quadrature  $\rightarrow$  and here's a function GHquadrature.m in Ragan PS6 code.

$\hookrightarrow$  he seems to be doing pdby fit iter of fodor ... humm ...

There's also tanchenhussey.m (Martin Foden)  
in "International" which does all of it for you.

But my concern is that I can't simply use AR(1)-discretization b/c even the LOM for  $b_t^{-1}$  isn't

a standard AR(1). If  $g(\cdot)$  were the simplest possible form, which is  $g(\mu) = d \cdot \mu^2 - c$

$$\text{then } k_t^{-1} = k_{t-1}^{-1} + d(-k_3 i_t + \Omega_{12} \bar{\pi}_{t-1} + \Omega_5 s_t - b_1 s_{t-1})^2 - c.$$

→ this may not even be Normal anymore.

As for  $\bar{\pi}_t$ , even if I don't plug  $k_t^{-1}$ ,

$$\bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1} \Omega_{12} \bar{\pi}_{t-1} + k_t^{-1} (-k_3 i_t + \Omega_5 s_t - b_1 s_{t-1})$$

$$\bar{\pi}_t = (1 + k_t^{-1} \Omega_{12}) \bar{\pi}_{t-1} + \underbrace{k_t^{-1} (-k_3 i_t + \Omega_5 s_t - b_1 s_{t-1})}_{\begin{array}{l} \uparrow \text{time-varying AR coeff} \\ \text{nonlinear MA terms} \end{array}}$$

→ may also not be Normal, and I haven't even plugged  $k_{t-1}^{-1}$ !

Judd 12.5 (Mac 113) deals w/ discretization.

But I want to leave that aside for now and see if a rule of the form of  $r(\cdot)$  can do better than a TR.

Solving the model for a rule of the form: let's shut these off

$$i_t = \gamma_\pi \pi_t + \gamma_k k_{t-1}^{-1} + \gamma_{\bar{\pi}} \bar{\pi}_{t-1} + \gamma_s s_t + \cancel{\gamma_{sl} s_{t-1}}$$

First of all, I need to i) either impose vector learning but then change the rule ii) have scalar learning but rewrite the model  $\rightarrow$  rewrite stuff.

I opt for i) b/c it's more flexible.

So let the rule look like this

$$i_t = \gamma_\pi \pi_t + \gamma_x x_t + \gamma_k k_{t-1}^{-1} + \gamma_{\bar{\pi}} \bar{\pi}_{t-1} + \gamma_{\bar{x}} \bar{x}_{t-1} \quad [r(1)]$$

$$\mathbb{C}_{a(1)} \quad \mathbb{C}_{a(2)}$$

Still use constant-only learning. ( $b = g x \cdot h x$ )

(as short of  $\gamma_x = \gamma_{\bar{x}} = 0$ , and it might be sufficient actually b/c  $\bar{\pi}_{t-1}$  might be a sufficient stat.)

$\rightarrow$  Model

$$\begin{bmatrix} 0 & 1 & 3 \\ 1 - \alpha & 0 & \\ -\gamma_\pi - \gamma_x & 1 & \end{bmatrix} \begin{bmatrix} \pi_t \\ x_t \\ i_t \end{bmatrix} = \begin{bmatrix} \text{stuff}_1 f_b + \text{stuff}_2 s_t \\ \text{stuff}_3 f_a + \text{stuff}_4 s_t \\ \gamma_k k_{t-1}^{-1} + \gamma_{\bar{\pi}} \bar{\pi}_{t-1} + \gamma_{\bar{x}} \bar{x}_{t-1} \end{bmatrix}$$

$\hookrightarrow \text{aleph-gind2.m}$

! only new thing

It seems to me that in materials 25.m, when I'm doing this  $(\gamma_1, \dots, \gamma_{\text{bar}}) = \arg \min (\text{B loss})$ , what I'm finding is that adding additional  $\gamma_s$  does not improve on the TR. On the contrary, the new  $\gamma_s$  seem to be substitutes to  $\gamma_0$  in a sense that the model just doesn't want high negative feedback, so it distributes more available  $\gamma_s$ , but  $\gamma_0$  can bear all the burden alone.

Need to get smooth anchoring fit to work 14 April 2020

- Side on multipliers, no commitment,  $\alpha_{pp}$
- Put TR on slide w/ SR costs vs long-run benefits
- Need to explain SR costs vs long-run benefits better
- Need to think of episodes the model can explain

## Peter meeting

19 April 2020

→ Macro lesson

- Need for a feedback rule: RE too?
- Implement  $T_C = \min CB \log_3$
- Covid-19?

→ RE: Taylor-principle needs to hold

↳ indeterministic w/ opt mon. pol when

RE is fit of easy shocks at + & -1.

→ One trick: Lautenbacher describes mon. pol on the cgb  
but off the cgb path I assume a policy

↳ learning based on RE to say that agents can  
distinguish between yours (CB's) behavior on &  
off the cgb path.

→ check these references to see if they have some  
trick

- TR preferred is an interesting result  
 $\rightarrow$  if it's robust, make that another section of the paper.

the same kind of bounded use rational EV  
 that gives  $\pi$  nice to anchored E also makes it hard for the CB to articulate a policy  
 Even this in theory it might be optimal to adjust to  $\bar{\pi}$  &  $k$ , if you in the TR include those in the TR, actually those coeffs are close to zero.

$\hookrightarrow$  Approximate the decision rules that depend on the nonlinear states as a finite dimensional grid.  
 e.g. guess that Vardle for is a linear grid  
 $\rightarrow$  iterate until close (but it won't equal) 0.

Fixed point which you won't.

Tschbeyscher-polynomials are the way to approx

nonlinear value pts, says Ken Judd.

### Work now

- Prezi mit Kühn + M
- ✗ Bring back old figure

Need to cut by 10 min.

Slide 2 2010-2020 instead of 2015

✓ → Puzzling interest rate setting 2015.  
Fed behavior

✗ "Taylor rule is not followed" vs.  
am Ende "Taylor rule ist doch besser"

Slide 3 ✓ endog. gain "fairly unusual" kann  
dann nicht einfach dekomponieren.

Slide 4 Nicht betonen den "special case" - sollte  
nicht wie ein failure darüber kommen

Slide 7 buttons lenken ab.

mindestens unterschreiten

✗ "Appendix: cons, price level"

Doppelprämie anstelle von "-"

Slide 10 ✓ misspecification is not irrational

Slide 13 ✓ cut!

Slide 19 ✓ 2 layers of novel intertemporal  
tradeoffs

✓ - emphasize the  $E(\cdot)$  of agents' tr (Baffets)

✓ - Take out  $\oplus$

- get in touch w/ Mdawar

Slide 21 ✓ "already true for every year" so why endog?

- ✓ Weswegen ist es gut den endog ganz hier explizit zu schen? Weil wir deswegen das Ted besser erklren knnen.
- ✓ "2nd interesting tradeoff" beschreibt es.

Slide 23. Betonen was ist SR  
X Vielleicht LCB schreiben?

Slide 24. / "TR disciplines expectations:"  
equation

Slide 25 Cent Taylor - rule prepared

- ✓ Next steps: form of reaction function  
Puzzling: problem results preps simple TR.  
Rephrase "TR better approx" and put to a different slide.

Slide 26: "Optimal policy is conditioned on /  
✓ conditional on"  
"Explains departures..."

### Mittel:

Strukturen:  
• Modell geht extrem lang  
✓ Implikationen      • Ramsey ist Algo dann fertig, &  
theoretisch      ← Übergang ist dann flüssig.  
partiell      10 min zu lang.

↓  
• Das Skriptus leidet am meisten:  
How to implement

Kapitel angeben → TOC / "Solving the Ramsey problem"

✓ · Bälle grünig beim Churneration

✗ · Größere margins

✓ · Buttons

✓ · Bindeshirt/H

## Notes for me

$$TC(\text{commitment RF}) \Rightarrow \pi_t = -\frac{\gamma_x}{k} x_t + \frac{\gamma_x}{k} x_{t-1}$$

- get in touch w/ Mawas

- slide 18: exoy gain? } the TC doesn't say

- SR vs LR tradeoffs } these LR-SR gains

- examples of other deviating from TR episodes

✓ Not today slide 4.

✓ PS doesn't know of slide 9.  $\rightarrow$  doesn't know  
they're identical

✓ slide 10, space

✓ Svensson target inflation (2003)

✓ Smith corollary w/ No commitment slide 15.

• Implementation: feedback rule

## Notes poszi

- Ryan: what if in RLS replace  $s_t$  for  $s_{t-1}$ ?
- Vito: normative implications of fed guidance  
outcome-based vs. instrument-based?

## Lynn meeting

Biggest challenge: RLS & poszzi's question

→ Lucas critique

→ fair critique of EC( ) formation

3 response points:

1. Are these expectations good positive theory?

→ Yes, I many papers in the lit. incl. Comp that show that this learning fit beliefs citing some facts that may match

e.g. inestin

2. Good for normative analysis?

RE isn't good positive, so an alternative is

needed.

This may not be the best tool, but it's what economists do! So it may be as good as any we have in terms of being robust to policy.

It may be better than models that assume over- or underreaction directly

b/c o- or u-reaction depends on policy and so regressions can capture those

3. Gain:  $g$  isn't a result of an opti-problem but it does give agents some ability to have E-formation respond to changes in environment.

- "I tried putting  $s_t$  there and it didn't matter"

• Ramsey problem min  $L^{\text{CB}}$

s.t. model equations  
and expectation formation

- Hope to get from me  
give a description of the signs of TC  
→ do I respond more or less to a cost-push  
⇒ ! ~

- The safe way is to say not that opt. policy  
is time-inconsistent  
↳ check time-inconsistency def

talk only about tech!

What Ryan wants to see: 48 hours before

objective function (history of  $x, i, \pi$ ;  
(exog states))

↳ complete history of beliefs that go w/ it  
and residuals to IS, PC & TR.

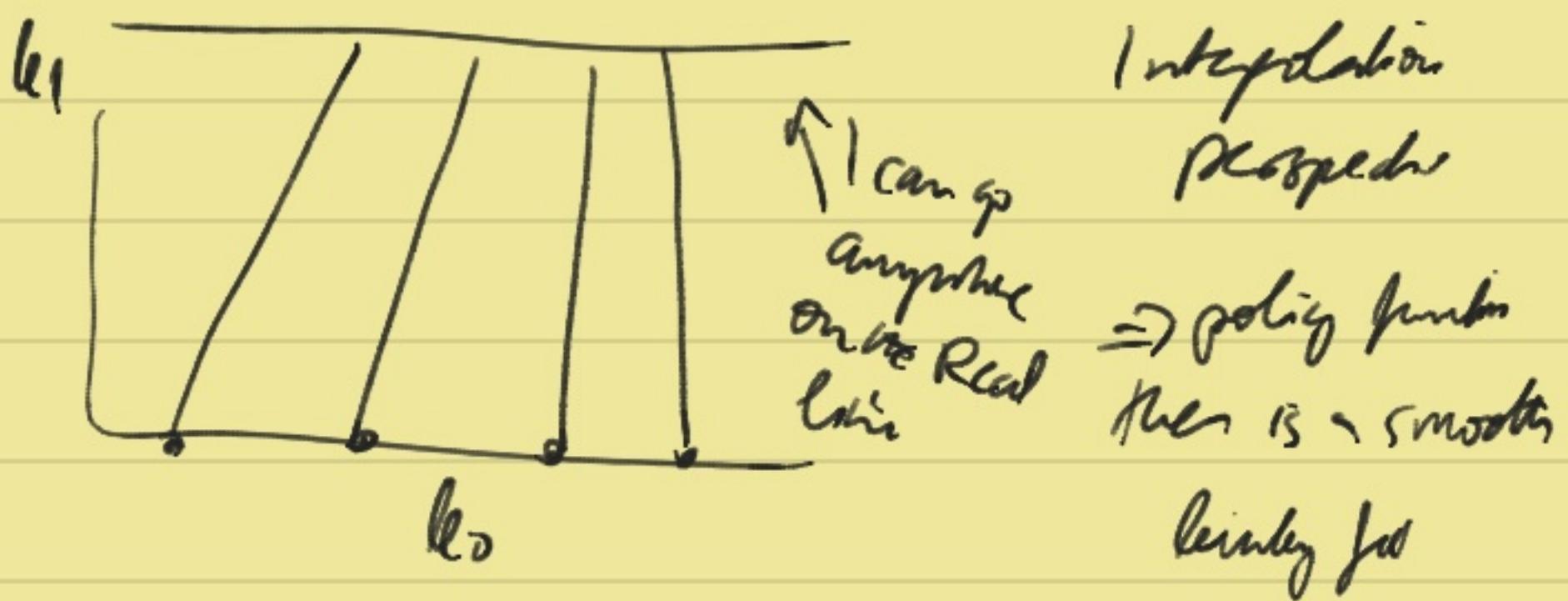
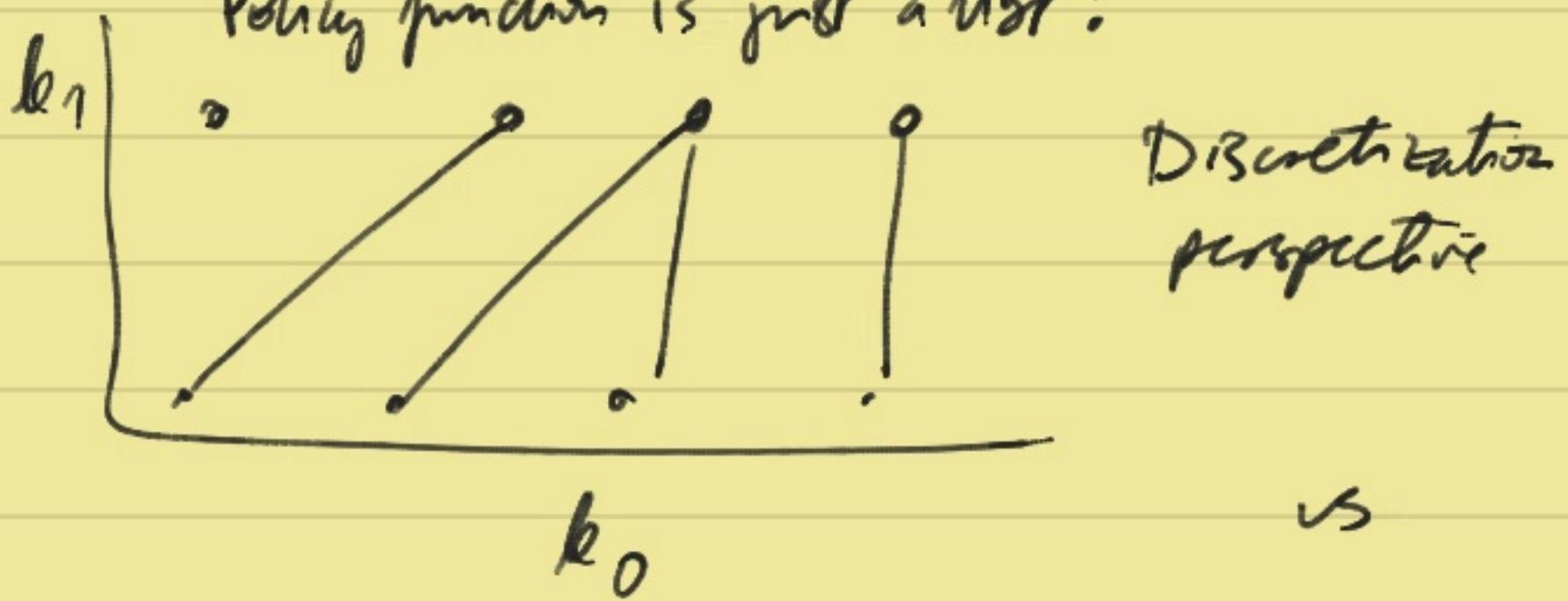
A starting point from a seq generated by  
the Taylor-rule.

Eval obj. pt once to show that it's working.

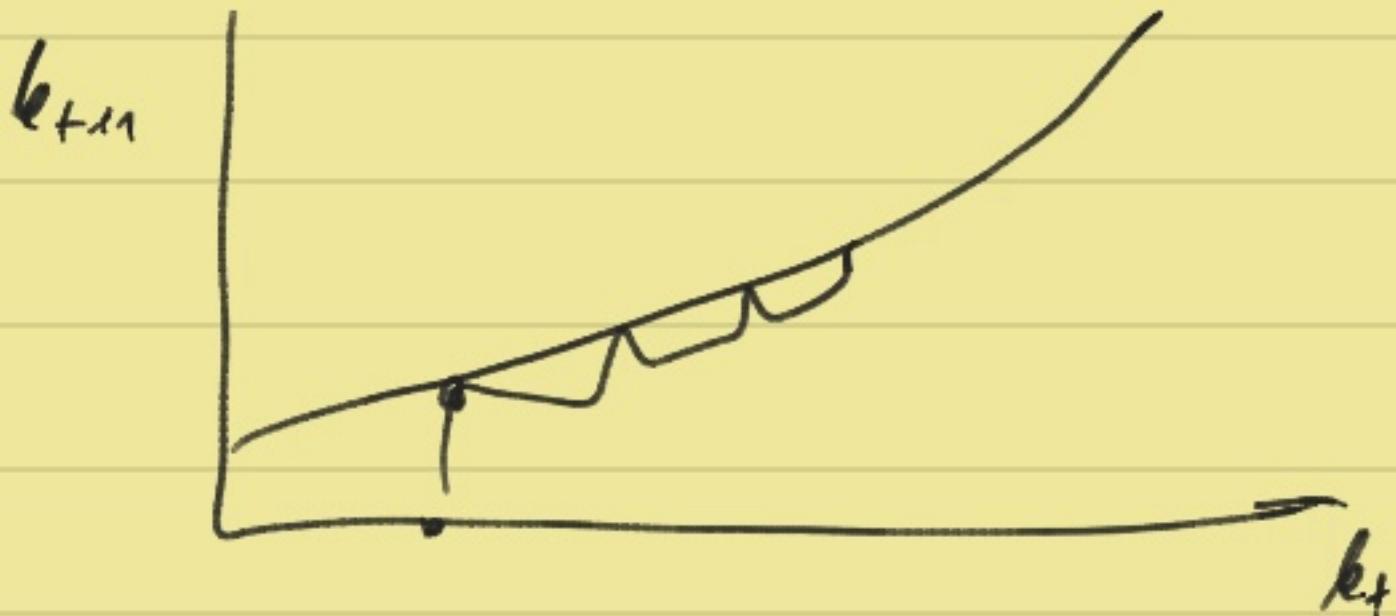
- Good bottom line that opt policy would deviate from the TR in which decisions → in the dir suggested by the TC!

Might claim string by putting in a less aggressive TR and how it violates the TC.

Policy function is just a list:



$\Rightarrow$  policy fit:



$\rightarrow$  interpolation allows me to have a smooth  $k_t, k_{t+1}$

$\rightarrow$  discretization is easier, interp is more efficient  
Ryan's approach:

- discretize exog states
- interpolate endog states

Work after materials26

18 April 2020

I want to clean up the laundry code once more so that before I submit it to Ryan, it's neater. Therefore:

- sim-learn-clean-smooth is RETIRED
  - sim-learnH1-clean-g is RETIRED
- focus on

sim-learnH1-clean.m, and create a counterpart, sim-learnH1-clean-optimal-seq.m

Also, Aa, Ab, As (A-matrices) RETIRED.

Stuff.m is replaced by smat.m  $\xrightarrow{\quad}$  stuff.m

depth-optimal.m replaced by ALM.m  $\xrightarrow{\quad}$  depth-optimal.m

there's a distance between A's and depth-optimal  
but I think it's only either very small (stuff.old)  
or large b/c things explode.

All of this & the procedure is summarized in  
materials26.tex!

The objective function will be called 20 April 2020

objchic-seq-clean.m  $\rightarrow$  replaces objchic-seq-future.m

So for Ryan I'm gonna do a folder  
called input-sequences-for-Ryan

Now that I've done that, I want to address the  
conceptual points Ryan mentioned about my prezi:

- ① Empirical facts the adaptive learning can match
- ② commitment, discretion, time-inconsistency, history-dep.
- ③ Does TL make me respond more or less to  
a cost-push shock?  $\rightarrow$  later. (I think len.)
- ④ Preston & int-rate-expectations

- ① Is this adaptive E-pomotion good positive theory?  
Yes b/c there are many papers that have shown  
that they can fit the following facts:

- backward-looking Phillips-curve / high persistence of  $\pi$   
Kilani (2007), Europ. Preston 2011 AER Inertia

- CEMP: anchoring can match long-term  $\pi - E(\cdot)$   
even tho it wasn't targeted

One thing I've just understood: CEMP can explain Missing Deflation/Inflation, but I can additionally explain why the Fed responds to it off the TR.

- Milani (2014): estimates an endog gain model  
and shows how the model generates time-varying volatility in macro variables  
→ not the shocks have time-varying volatility,  
but the endog gain feeds them to variables.  
→ matches higher volatility in Great Inflation &  
decreasing in Great Moderation

## ② Commitment & discretion

In particular, I want to clarify time-inconsistency  
Fed posts & Mankiw's blog seem to treat discretion  
and time-inconsistency as analogous.

Woodford p. 14

Commitment := CB chooses to commit itself to a policy rule / systematic procedure for setting  $i$

Discretion := if the bank acts at each date under the assumption that it cannot commit itself to any future behavior (and is not bound by any past commitments)

↳ the point is that if I take a simple analogy to my TC, then it says

$$\pi_t = -\frac{\lambda}{K} x_t + \frac{\lambda}{K} \text{stuff} \cdot x_{t+1}$$

i.e. "I'll raise  $x$  today and I promise to lower it tomorrow" but when tomorrow comes, I face the same TC and b/c no lagged  $x$  shows up, I can't honor my commitment

$\Rightarrow$  time-inconsistent b/c it doesn't follow thru w/ its promise

Note that the contribution of Kydland & Prescott 1977

is to show that optimal control theory leads to time-inconsistent results (opti commitment).

"Only if expectations were invariant to the future policy plan would optimal control theory be appropriate."  
(Kl, p. 474)

Discretion is time-inconsistent b/c at each point in time it picks what is optimal ("best") given the current stance of things & expected stuff gives today but ignores the effects of this choice on expectations and thus on endog vars.

→ history-dependence is a way to implement commitment.

p. 21. Woodford "distinguishes the approach advocated here from the usual understanding of what an optimal commitment involves"

→ the usual understanding (time-consistent optimal commitment) is not time-consistent, by which

Woodford means that "solving the same opt problem at a later date  $t_1$  to determine the opt commitment from that date onward does not result in a state-contingent plan from  $t_1$  onward that contains the plan judged to be optimal at  $t_0$ ." p. 22.

→ The reason why this time-inconsistent is the same as why discretion yields a diff outcome b/c  $E(\cdot)$  are taken as given at  $t_1$ .

And indeed the timeless commitment policy is time-consistent.

Now to close this question, let's 21 April 2020 read the part about  $\pi$ - vs. price-level targeting.

Recall: cost of discretion:

inflation bias (st. st.  $\pi$  too high)

suboptimal response to shocks

Subtle:  $t_0$ -optimal commitment also has an initial inflation bias b/c in the first few periods you're free

of commitment: can take  $E(\cdot)$  as given initially.  
→ It seems to me to be only the initial term,  $\pi_0$ ,  
which is  $> 0$ . (see p. 473)

bc initial  $\pi$  cannot affect prior expectations

[the key figure is Fig. 7.1. p. 476]

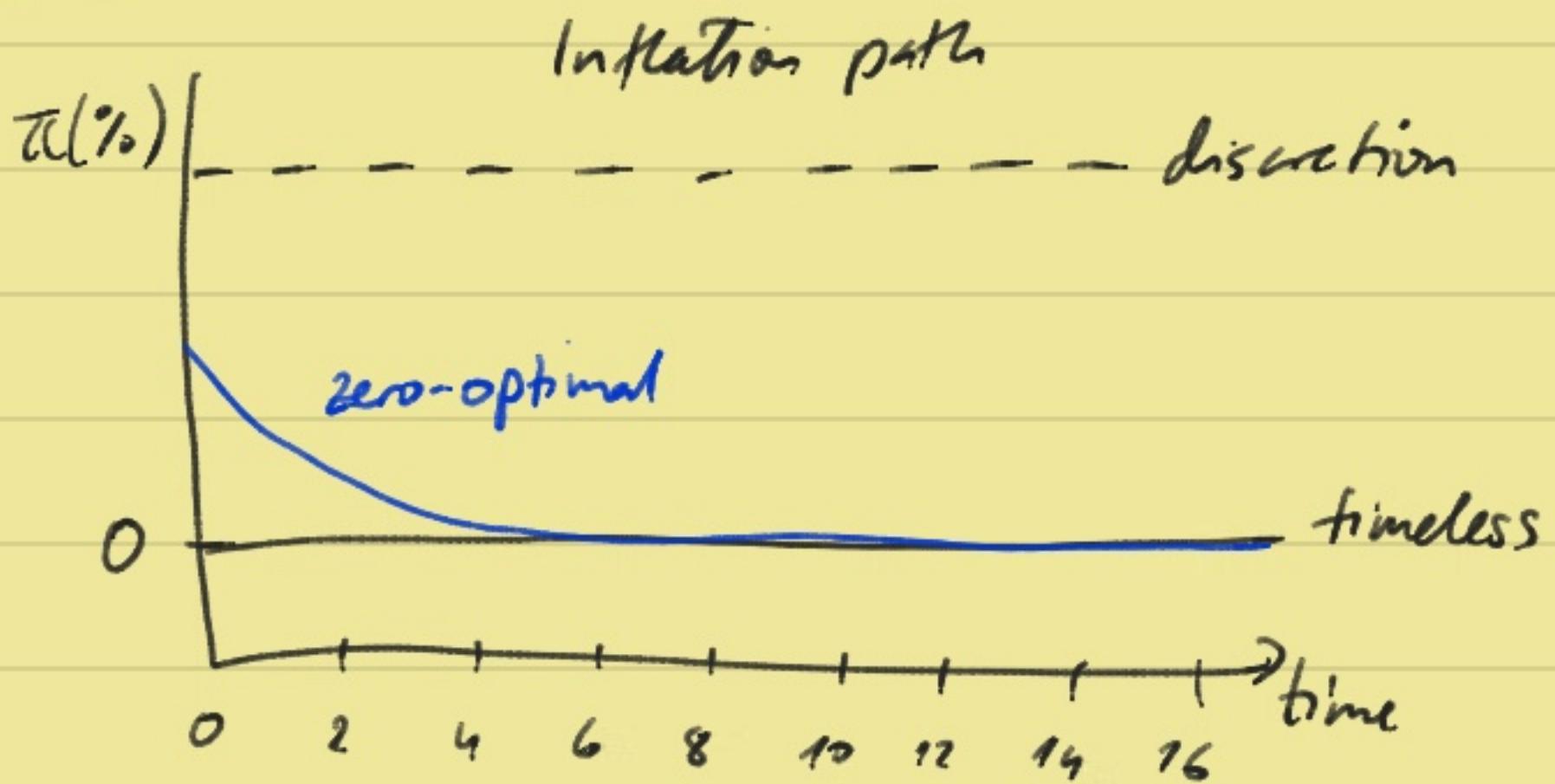


Fig 7.1. Inflation bias under discretion & zero-optimal committed

The key fig - showing suboptimal responses to shocks  
is Fig. 7.3 [ p. 494 ]

--- disruption  
— optimal commitment

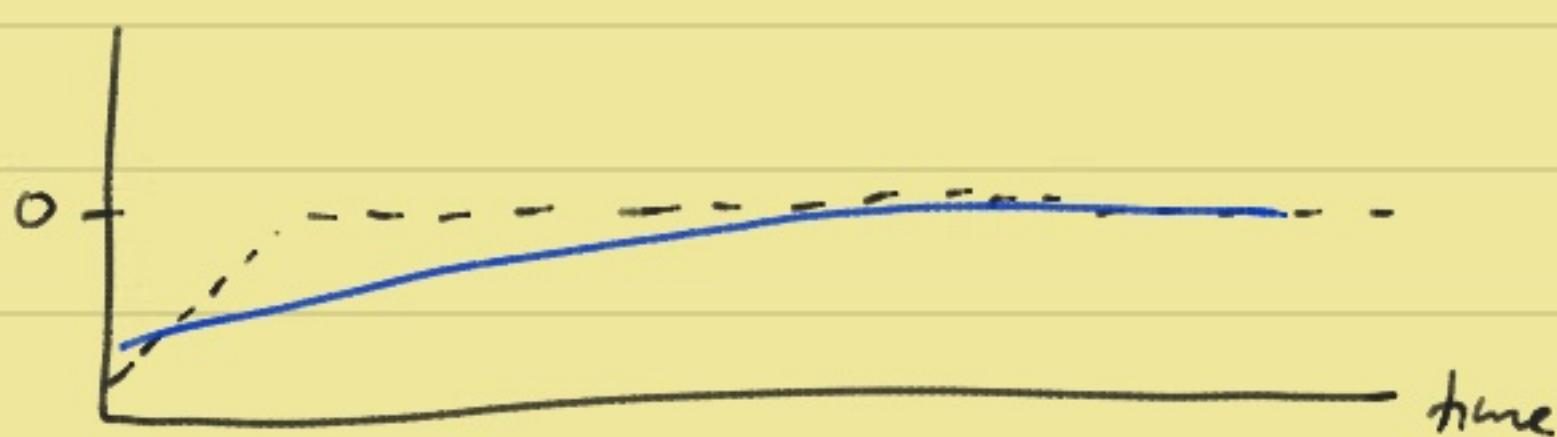


Fig. 7.3. Stabilization bias of disruption.

Optimal responses to a transitory cost-push shock

$t_0$  = timeless optimal commitment in terms of responses to shocks.

From this I know now why Bauger et al said that w/ the again TL, responses to  $u_f$  resemble commitment more  $\rightarrow$  it's b/c the main feature of discretion is that it takes the shock full on at time  $t$ , while commitment spreads it out  $\rightarrow$  which is exactly what learning allows you to do. So on impact, you respond <sup>more  $\rightarrow$  to TL moves</sup> less to an  $u_f$ -shock than you would under RE-discretion b/c you can postpone the tradeoff to the future. But the reason is very different from commitment. In commitment you can respond less b/c  $E(\cdot)$  anticipating your future responses makes a smaller response necessary. W/ learning, instead, you don't have to face the full discretion-tradeoff b/c  $E(\cdot)$  isn't responding fast enough.

I'm not sure about the size of the response in  $u_f$  necessary. All I know is that  $\pi_8 \times$  more less under commitment & in learning.

Let me note that there are two distinct questions in  
discretion vs. commitment

1) st. st. or long-run target or time-path of variables

→ this is where time-consistency matters

discretion  $\rightarrow$  zero-optimal commitment  $\rightarrow$  timeless

"inflation bias" "inflation bias initially" "no inflation bias"

time-inconsistent

time-inconsistent

time-consistent

Kydland & Prescott 1977

2) stabilizer / responses to shocks

→ I think that here there is no distinction between

zero- & timeless optimal commitment b/c they

respond to shocks the same way, namely by

spreading out the shock. I'm not sure it makes sense

to even talk about time-consistency wrt to responding to

shocks, b/c you're not answering the question about the

optimal time path!

As for the response to a cost-push shock  
 RE discretion  $\pi_t = -\frac{\alpha}{K} x_t$

RE commitment  $\pi_t = -\frac{\alpha}{K} x_t + \frac{\alpha}{K} x_{t-1}$

Learning TL:  $\pi_t = -\frac{\alpha}{K} x_t + \frac{\alpha}{K} E_t x_{t+1}$

Suppose we start w/  $\pi_{t+1}, x_t, x_{t-1} = 0$ . A cost-push  
 shock hits,  $u_t = 1 \rightarrow \pi_t \uparrow$  by 1.

- On impact: discretion & commitment face the same situation

$$1 = -\frac{\alpha}{K} x_t$$

while learning faces an attenuated tradeoff:

$$1 = -\frac{\alpha}{K} x_t + \frac{\alpha}{K} E_t x_{t+1}$$

which depends on the stance of learning.

But commitment needs to respond less today b/c it  
 will respond tomorrow (and people know this), discretion  
 responds 1:1 and learning postpones responding.

But I don't think you can tell whether the CB is more aggressive on inflation! Supp CB only cares about  $\pi_t$ . Then it responds most aggressively under RE discretion, not b/c it cares more, but b/c under commitment, expectations adjust to do part of the job, and learning would be indiff b/t moving  $y_t$  or  $x_{t+1}$ . So I'm unsure about Gaspar et al's 2011 conclusions.

Learning also cannot be time-consistent b/c the choice of the optimal time-path depends on the stance of learning.

Price-level targeting vs. inflation-targeting (PT vs IT)

PLT > IT b/c it leads price-setters to expect that  $\Delta P_t$  will be undone, thus adapting their behavior. Note that inflation inertia undoes the result

that the same price level  $p_t$  is always fully restored, as well as it involves the result that timely optimal commitment has no initial ↑ in inflation on the optimal time-path.

### Peter meeting

29 April 2020

- Janneir: why this model? → B/c it has been developed in lit as the leading model of anchored  $E(\cdot)$ 
  - + it has the ability to deal w/ the PC in a fwd-looking way w/o RE.
- Liked the highlight of the TC.
- Next general comment: pull set of results & figure out how to present. This color-coded results is helpful to foreshadow the results but w/o results you don't know yet what they're foreshadowing.

Time-inconsistent: begin at time 0, allow CB to choose a state-contingent path, and then at a later date, allow you to rechoose, then CB won't choose to continue the same plans.

But Peter interprets the same problem to be applicable in response to shocks

Kydland & Prescott 1977: time-inconsistency only arose once  $E(\cdot)$  was modeled as RE vs adaptive

Some classes of models e.g. adaptive learning where discretion & commitment coincide.

Say instead of GQR & Lemma: the discretion & commitment sols to Ramsey problem coincide and are more akin to the discretion sol.  
→ You don't wanna argue about terminology!

## Work after

Ok, clearly I need to double-check whether  
discretion is time-consistent or not.

I'm taking a time-out to consider the role of int rate beliefs. Ch. 4 in Woodford doesn't actually mention them, so let's go to Preston:

- Eusepi & Preston 2018 (Science of Mon Pol)

Result 6: the IS curve is a constraint in the Ramsey problem b/c i-beliefs are a relevant state variable. (p. 30 Mac)

(ii) close to REE, opt policy is more aggressive toward inflation, far from REE, it is less.

- Eusepi, Giannoni, Preston 2018 (Limits)

→ emphasize the importance of (lt H-) int rate expectations for the stabilization effectiveness of mon. pol. The main result is that int rate

beliefs are a constraint to mon. pol.

- p. 16 top: "b/c beliefs are a state variable there is no distinction b/w optimal commitment & discretion"
- and their interpretation of this is that the AD eq being a constraint is a manifestation of the effect of it on  $E(i)$ .

Corollary 2: (p. 21)

If the int-rate beliefs are anchored to RE,  
the AD curve is no longer a constraint &  
divine coincidence is restored.

→ I disagree b/c they still need a TR to be anchored to.

• Preston 2002

only really talks about int-rate beliefs on p. 33

(→ I'm not sure how to interpret his interpretation<sup>(Mac)</sup>)

though! He doesn't go into it in the App.

Back to discretion: is discretion time-consistent?

Taylor Lecture notes: discretion  $\rightarrow$  consistent  
rule  $\rightarrow$  optimal  
cheating  $\rightarrow$  inconsistent

Is time-inconsistency tied to the idea that  
preferences change over time?

- Odysseus & the sirens
- Procrastination

What I don't get is that Kydland Prescott's 1977  
paper kinda says rules are better than discretion  
b/c we overcome the time-inconsistency problem.

But maybe "Rules vs discretion" is a misnomer?

Maybe Kydland & Prescott are trying to say that  
the optimal Ramsey policy under commitment is  
time-consistent, so that absent rules, it can't be  
implemented and we get discretion?

Kydland & Prescott seem to be saying that discretionary policy is time-consistent but not optimal in the sense that commitment yields a superior outcome. But commitment is inconsistent, so that's why optimal plans are inconsistent.

Their def of time-consistency: a policy is consistent if it maximizes the objective fn, taking previous policy decisions as given and that future policy decisions are similarly selected.

"What is best given the current sitn"

I think I know why discretion is consistent: b/c it doesn't mean that you reoptimize at a later date - it just means that when you do optimize, you don't take the effects of your policy on  $E(\cdot)$  into account. So that if you would reoptimize later, you still don't take them into account! Consistent!

Ok - here's the (final?) word on

22 April 2020

time-consistency:

Time consistency means that if you could reoptimise at a later date, you would choose the same plan (but not necessarily the same action!). Therefore commitment is inconsistent b/c a planner wouldn't choose the commitment plan once the plains are flooded. Discretion (the plan involving \*inconsistent\* actions (forbid building houses in floodplain, rescue flooded areas)) is time consistent b/c even once the plains are flooded, the planner would choose to adhere to this plan of action. Discretion is by design time consistent b/c the plan it selects is to always choose the action that's best in that moment.

Timeless commitment is time consistent b/c the initial multiplier forces you to think the same way once the plains are flooded.

And time-consistency is an issue relevant for the responses to shocks. What isn't relevant is to- vs. timeless commitment b/c they ascribe the same responses to shocks.

S. Fischer, 1980, Section 9, last paragraph p.104 (12 Mar)  
time inconsistency arises as a problem only if expectations are partly rational in the sense that they depend in some way on announcements of future policy. If  $E_t(\cdot)$  are purely adaptive (only depend on past behavior) then the time inconsistency problem does not arise unless the controller's tasks change.

↳ This is why procrastination is actually time inconsistent b/c w/ hyperbolic discounting, your preferences change.

Chen, Kehoe & Prescott 1988

p. 11, Prop 1. (13 Mac)

For there to be a time inconsistency problem, there needs to be a conflict of interest, either between society or agents, or between agents themselves.

Sounds counterintuitive but actually makes sense:

- discretion is consistent b/c you always max U given current circumstances → agents would agree w/ you
- commitment is inconsistent b/c when you announce that no one may build houses in the floodplain, agents disagree w/ you.

Let me close the discussion here: so in my model, there is no inconsistency problem.