Materials 3 - Special cases

Laura Gáti

September 19, 2019

Overview

1 Special cases towards general case: agenda

1. Simulate RE model ✓
2. Simulate EE model where agents learn both slope and constant ✓
   - Simulate using the "implicit ALM": rearranging the expectational matrix equation that underlies the solution to the model, you obtain the simulated observables z_t without explicitly writing out the ALM ✓
   - Simulate using the "explicit ALM", equation (6), plugging in expectations evaluated separately. ✓

   The cool thing is: when I do the above two steps, I obtain the same simulated observables, so I know I'm doing it correctly.
3. Simulate LR model where agents learn both slope and constant, extend horizons from 1 to ∞ ✓
4. Simulate EE model where agents learn only the constant ✓
5. Simulate LR model where agents learn only the constant, extend horizons from 1 to infinity ✓
6. Simulate LR model with anchoring ✓

   So far, they all converge to RE!
7. Simulate LR model with anchoring, when agents only learn LR inflation

2 Timeline in the learning models

<u>$t = 0$</u>: Initialize learning coefficients $\phi_{t-1} = \phi_0$ at the RE solution.

For each t:

1. Evaluate expectations $t + s$ (the one-period ahead, ($s = 1$) or the full 1 to ∞-period ahead ($s = 1, \dots, \infty$)) given ϕ_t and states dated t
2. Evaluate ALM given expectations: "today's observables are a function of expectations and today's state"
3. Update learning: ϕ_{t+1} = RLS of ϕ_t and fcst error between today's data and yesterday's forecast

3 The models to be simulated

1. Rational expectations NK model (RE)
2. Euler equation approach learning NK model à la Bullard & Mitra (2002) (EE)
3. LR expectations learning NK model à la Preston (2005) (LR)
4. LR expectations learning NK model à la Preston with anchoring à la CEMP
5. Same, but with agents only learning about the drift in inflation

The difference between these models is 1) in the expectations (rational or not), 2) in the number of horizons of expectations that need to be summed (1 vs. infinite) and 3) the gain dynamics. So what I'm going to do consists of 3 steps:

1. Write a learning rule that takes the form of Preston's, but that nests CEMP, and has a decreasing gain.
2. Write out f_a and f_b as truncated sums of h-period ahead forecasts. When $h = 1$, EE and LR (models (6) and (7)) should coincide. (Actually - maybe they shouldn't. See later.)
3. Add variations: anchoring, and scalar anchoring

3.1 Rational expectations NK model (RE)

$$x_t = \mathbb{E}_t\, x_{t+1} - \sigma(i_t - \mathbb{E}_t\, \pi_{t+1}) + \sigma r_t^n \tag{1}$$

$$\pi_t = \kappa x_t + \beta\, \mathbb{E}_t\, \pi_{t+1} + u_t \tag{2}$$

$$i_t = \bar{i}_t + \psi_\pi \pi_t + \psi_x x_t \tag{3}$$

3.2 NK model with Euler equation learning (EE)

$$x_t = \hat{\mathbb{E}}_t x_{t+1} - \sigma(i_t - \hat{\mathbb{E}}_t \pi_{t+1}) + \sigma r_t^n \qquad \text{(Preston, eq. (13))}$$

$$\pi_t = \kappa x_t + \beta \hat{\mathbb{E}}_t \pi_{t+1} + u_t \qquad \text{(Preston, eq. (14))}$$

$$i_t = \bar{i}_t + \psi_\pi \pi_t + \psi_x x_t \qquad \text{(Preston, eq. (27))}$$

3.3 NK model with ∞-horizon forecasts (LR)

$$x_t = -\sigma i_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \beta^{T-t}\big((1-\beta)x_{T+1} - \sigma(\beta i_{T+1} - \pi_{T+1}) + \sigma r_T^n\big) \tag{Preston, eq. (18)}$$

$$\pi_t = \kappa x_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t}\big(\kappa\alpha\beta x_{T+1} + (1-\alpha)\beta\pi_{T+1} + u_T\big) \tag{Preston, eq. (19)}$$

$$i_t = \psi_\pi \pi_t + \psi_x x_t + \bar{i}_t \tag{Preston, eq. (27)}$$

One issue is that if I set $T = t$, I don't think Preston, eq. (18) reduces to Preston, eq. (13), nor does Preston, eq. (19) reduce to Preston, eq. (14). But actually, maybe I shouldn't expect them to reduce to the EE learning equations. Why not? Because, as Preston stresses, the EE approach not only neglects future state variables that individuals find relevant to their decision (future wealth) but also imposes equilibrium conditions that agents wouldn't know (market clearing). Therefore if I subtract the ∞-future element, the element of equilibrium conditions remains as a difference.

4 Compact notation

Exogenous states are summarized as:

$$s_t = P s_{t-1} + \epsilon_t \quad \text{where} \quad s_t \equiv \begin{pmatrix} r_t^n \\ \bar{i}_t \\ u_t \end{pmatrix} \quad P \equiv \begin{pmatrix} \rho_r & 0 & 0 \\ 0 & \rho_i & 0 \\ 0 & 0 & \rho_u \end{pmatrix} \quad \epsilon_t \equiv \begin{pmatrix} \varepsilon_t^r \\ \varepsilon_t^i \\ \varepsilon_t^u \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_r & 0 & 0 \\ 0 & \sigma_i & 0 \\ 0 & 0 & \sigma_u \end{pmatrix}$$

Let z_t summarize the endogenous variables as

$$z_t \equiv \begin{pmatrix} \pi_t \\ x_t \\ i_t \end{pmatrix} \tag{4}$$

Then I can write the models compactly as

$$z_t = A_p^{RE}\, \mathbb{E}_t\, z_{t+1} + A_s^{RE} s_t \tag{5}$$

$$z_t = A_p^{RE} \hat{\mathbb{E}}_t z_{t+1} + A_s^{RE} s_t \tag{6}$$

$$z_t = A_a^{LR} f_a(t) + A_b^{LR} f_b(t) + A_s^{LR} s_t \tag{7}$$

$$s_t = P s_{t-1} + \epsilon_t \tag{8}$$

where f_a and f_b capture discounted sums of expectations at all horizons of the endogenous states z (following Preston, I refer to these objects as "long-run expectations"):

$$f_a(t) \equiv \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} z_{T+1} \qquad f_b(t) \equiv \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\beta)^{T-t} z_{T+1} \tag{9}$$

and the coefficient matrices are given by:

$$A_p^{RE} = \begin{pmatrix} \beta + \frac{\kappa\sigma}{w}(1-\psi_\pi\beta) & \frac{\kappa}{w} & 0 \\ \frac{\sigma}{w}(1-\psi_\pi\beta) & \frac{1}{w} & 0 \\ \psi_\pi\left(\beta + \frac{\kappa\sigma}{w}(1-\psi_\pi\beta)\right) + \psi_x \frac{\sigma}{w}(1-\psi_\pi\beta) & \psi_x(\frac{1}{w}) + \psi_\pi(\frac{\kappa}{w}) & 0 \end{pmatrix} \tag{10}$$

$$A_s^{RE} = \begin{pmatrix} \frac{\kappa\sigma}{w} & -\frac{\kappa\sigma}{w} & 1 - \frac{\kappa\sigma\psi_\pi}{w} \\ \frac{\sigma}{w} & -\frac{\sigma}{w} & -\frac{\sigma\psi_\pi}{w} \\ \psi_x(\frac{\sigma}{w}) + \psi_\pi(\frac{\kappa\sigma}{w}) & \psi_x(-\frac{\sigma}{w}) + \psi_\pi(-\frac{\kappa\sigma}{w}) + 1 & \psi_x(-\frac{\sigma\psi_\pi}{w}) + \psi_\pi(1 - \frac{\kappa\sigma\psi_\pi}{w}) \end{pmatrix} \tag{11}$$

$$A_a^{LR} = \begin{pmatrix} g_{\pi a} \\ g_{xa} \\ \psi_\pi g_{\pi a} + \psi_x g_{xa} \end{pmatrix} \quad A_b^{LR} = \begin{pmatrix} g_{\pi b} \\ g_{xb} \\ \psi_\pi g_{\pi b} + \psi_x g_{xb} \end{pmatrix} \quad A_s^{LR} = \begin{pmatrix} g_{\pi s} \\ g_{xs} \\ \psi_\pi g_{\pi s} + \psi_x g_{xs} + \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \end{pmatrix} \tag{12}$$

$$g_{\pi a} = (1 - \frac{\kappa\sigma\psi_\pi}{w}) \left[(1-\alpha)\beta, \kappa\alpha\beta, 0\right] \tag{13}$$

$$g_{xa} = \frac{-\sigma\psi_\pi}{w} \left[(1-\alpha)\beta, \kappa\alpha\beta, 0\right] \tag{14}$$

$$g_{\pi b} = \frac{\kappa}{w} \left[\sigma(1-\beta\psi_\pi), (1-\beta-\beta\sigma\psi_x, 0\right] \tag{15}$$

$$g_{xb} = \frac{1}{w} \left[\sigma(1-\beta\psi_\pi), (1-\beta-\beta\sigma\psi_x, 0\right] \tag{16}$$

$$g_{\pi s} = (1 - \frac{\kappa\sigma\psi_\pi}{w}) \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} (I_3 - \alpha\beta P)^{-1} - \frac{\kappa\sigma}{w} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} (I_3 - \beta P)^{-1} \tag{17}$$

$$g_{xs} = \frac{-\sigma\psi_\pi}{w} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} (I_3 - \alpha\beta P)^{-1} - \frac{\sigma}{w} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} (I_3 - \beta P)^{-1} \tag{18}$$

$$w = 1 + \sigma\psi_x + \kappa\sigma\psi_\pi \tag{19}$$

5 Learning

In Preston (2005), agents forecast the endogenous variables using the exogenous ones as

$$z_t = a_t + b_t s_t + \epsilon_t \qquad \text{(Preston, p. 101)}$$

which I suspect isn't precise about the timing. Therefore, I write a general PLM of the form

$$z_t = a_{t-1} + b_{t-1} s_{t-1} + \epsilon_t \tag{20}$$

and then $\phi_{t-1} = (a_{t-1}, b_{t-1})$, here 3×4, so that agents learn both a constant and a slope term. This means $\hat{\mathbb{E}}_t z_{t+1} = \phi_t \begin{bmatrix} 1 \\ s_t \end{bmatrix}$. Later, I will simplify here so that agents only learn about the constant, i.e. about CEMP's drift term, but forecast exogenous states rationally:

$$z_t = \bar{z}_{t-1} + b s_{t-1} + \epsilon_t \tag{21}$$

so that $\phi_{t-1} = \bar{z}_{t-1}$. I was initially quite worried about the assumption that agents only learn about the constant because without properly defining the constant slope b, it becomes a permanent deviation from RE that screws up E-stability.

Anticipated utility implies that

$$\hat{\mathbb{E}}_{t-1} \phi_{t+h} = \hat{\mathbb{E}}_{t-1} \phi_t \equiv \phi_{t-1} \quad \forall\, h \geq 0 \tag{22}$$

This is a little tricky. It doesn't only mean that agents today mistakenly believe that they will not update the forecasting rule. It also implies that the belief about ϕ_t was formed at $t-1$. In other words, in the evening of period $t-1$, agents update ϕ to get the ϕ_t they will use in period t. Assuming RE about the exogenous process and anticipated utility, h-horizon forecasts are constructed as:

$$\hat{\mathbb{E}}_t z_{t+h} = a_t + b_t P^{h-1} s_t \quad \forall h \geq 1 \tag{23}$$

Or, if I assume that agents don't learn the slope,

$$\hat{\mathbb{E}}_t z_{t+h} = \bar{z}_t + b P^{h-1} s_t \quad \forall h \geq 1 \tag{24}$$

and the regression coefficients are updated using (for now) a decreasing gain RLS algorithm:

$$\phi_t = \left(\phi_{t-1}' + t^{-1}\mathbf{R_t^{-1}}\begin{bmatrix}\mathbf{1} \\ \mathbf{s_{t-1}}\end{bmatrix}\left(z_t - \phi_{t-1}\begin{bmatrix}1 \\ s_{t-1}\end{bmatrix}\right)'\right)' \tag{25}$$

$$R_t = R_{t-1} + t^{-1}\left(\begin{bmatrix}1 \\ s_{t-1}\end{bmatrix}\begin{bmatrix}1 & s_{t-1}\end{bmatrix} - R_{t-1}\right) \tag{26}$$

R_t is 4×4 and ϕ_t is 3×4. Two questions:

1. The bold $\mathbf{R_t^{-1}}\begin{bmatrix}\mathbf{1} \\ \mathbf{s_{t-1}}\end{bmatrix}$ indicates a difference to CEMP's learning algorithm: these terms are missing in CEMP. Am I right in thinking that that's because in CEMP, agents only learn the constant, and so the data they use is 1 instead of $\begin{bmatrix}1 \\ s_t\end{bmatrix}$, making $R_t = 1 \; \forall \; t$?
2. Can this formulation capture the special case that agents only learn about the constant? $\Leftrightarrow$ Following up on the previous point, it seems to me that when agents learn only the constant, then $\phi_t = \bar{z}_t$ and the learning algorithm boils down to

$$\bar{z}_t = \bar{z}_{t-1} + t^{-1}\underbrace{\left(z_t - (\bar{z}_{t-1} + s_{t-1})\right)}_{\text{fcst error using (24)}} \tag{27}$$

And a note: CEMP is a special case of this model, with the gain switching between decreasing and constant according to the anchoring mechanism.

6 ALMs

6.1 RE

With some abuse of terminology, call the state-space representation the ALM of the RE model:

$$x_t = hx\; x_{t-1} + \eta e_t \tag{28}$$

$$z_t = gx\; x_t \tag{29}$$

Then I can write the "ALM" as

$$z_t = gx\; hx\; x_{t-1} + gx\; \eta e_t \tag{30}$$

Since this ALM implies no constant, I initialize $\bar{z}_0$ as a 3×1 zero vector, and thus $\phi_0 = \begin{bmatrix} \bar{z}_0 & gx\ hx \end{bmatrix}$ (and $hx = P$ for the NK model). Analogously, I initialize R as a $R_0 = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \Sigma_x \end{bmatrix}$, where Σ_x is the VC matrix of the states from the RE solution. For the case where agents only learn the constant, I still initialize $\bar{z}_0$ as a 3×1 zero vector (and R drops).

6.2 EE

I just need to use (23) to evaluate one-period ahead forecasts (for constant-learning only, (24)), and plug those into (6).

6.3 LR

Evaluate analytical "LR expectations" (9) using the PLM (23),

$$f_a(t) = \frac{1}{1-\alpha\beta} a_t + b_t (I_3 - \alpha\beta P)^{-1} s_t \qquad f_b(t) = \frac{1}{1-\beta} a_t + b_t (I_3 - \beta P)^{-1} s_t \tag{31}$$

and plug them into (7). In the case where agents only learn the constant I use (24):

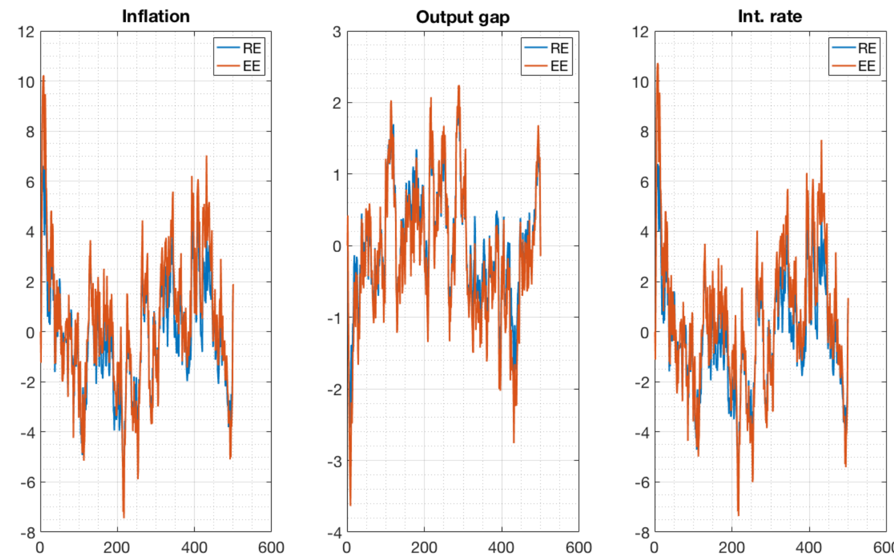
$$f_a(t) = \frac{1}{1-\alpha\beta} \bar{z}_t + b (I_3 - \alpha\beta P)^{-1} s_t \qquad f_b(t) = \frac{1}{1-\beta} \bar{z}_t + b (I_3 - \beta P)^{-1} s_t \tag{32}$$

Alternatively I can evaluate each h-period ahead forecast individually using (23), and then sum H of these terms, discounting appropriately. Earlier, it seemed that already a $H = 100$ is not a bad approximation of ∞-horizons, but now that only holds for f_a. For f_b to be accurate, I need at least $H = 10000$. Why? Does the fact that $\alpha\beta < \beta$ matter so much?

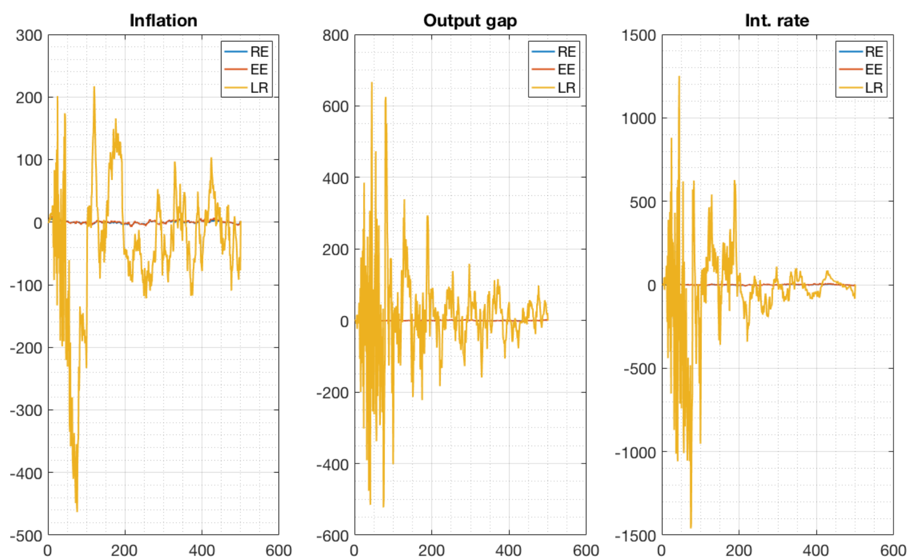
7 Simulations

7.1 Learning slope and constant

Figure 1: Comparing models

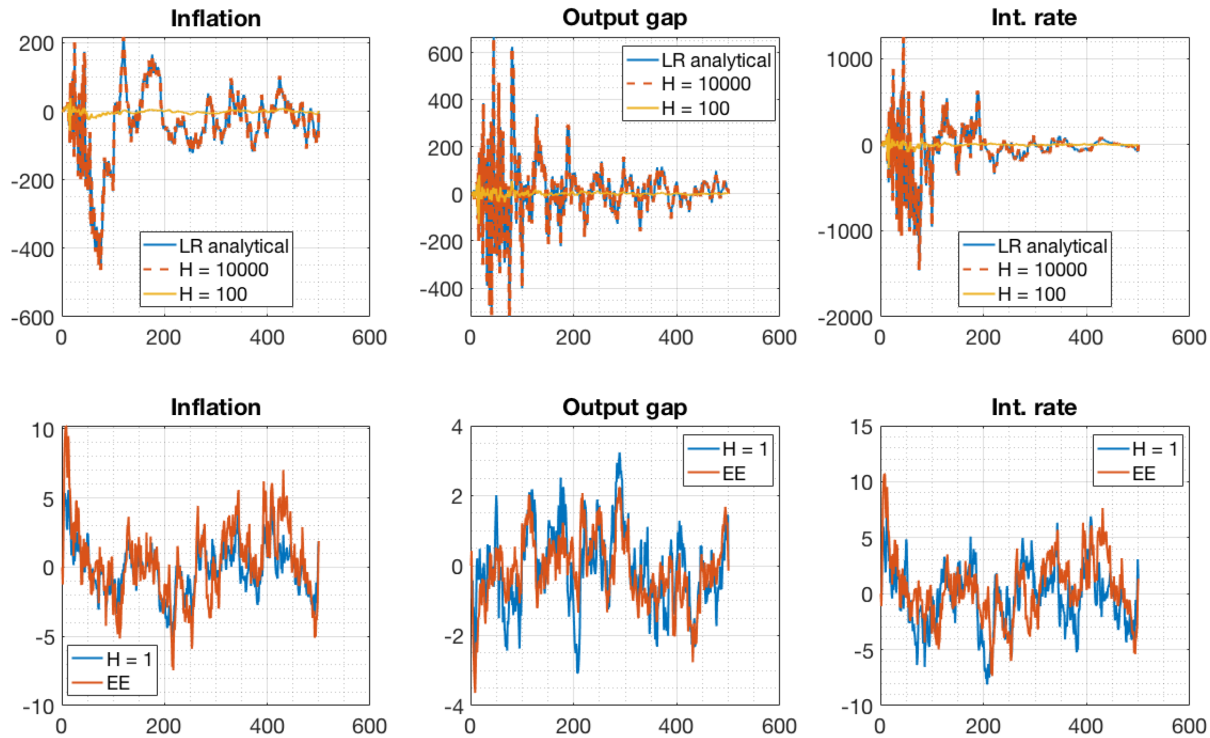


(a) RE and EE models only, learning both slope and constant



(b) RE, EE and LR models, learning both slope and constant, analytical expectations

Figure 2: Comparing horizons



Takeaways:

1. EE learning converges to RE over time, confirming that it's correct. Does LR? It doesn't seem like it (at $T = 100000$, it hasn't converged).

2. LR analytical and truncated expectations coincide for a large enough horizon ($H \sim 10000$)

3. Even for $H = 1$, LR and EE don't coincide; I think this is because the equations do not map onto one another. As argued before, maybe they shouldn't either.

So is it the case that learning isn't converging in the LR model?

→ No! It's clearly converging, albeit slowly, see next fig!

Figure 3: Convergence LR learning

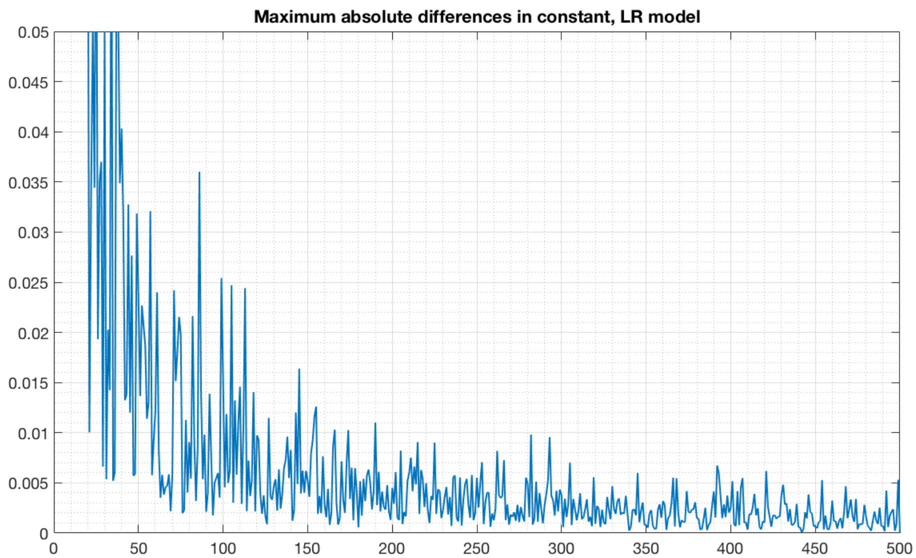
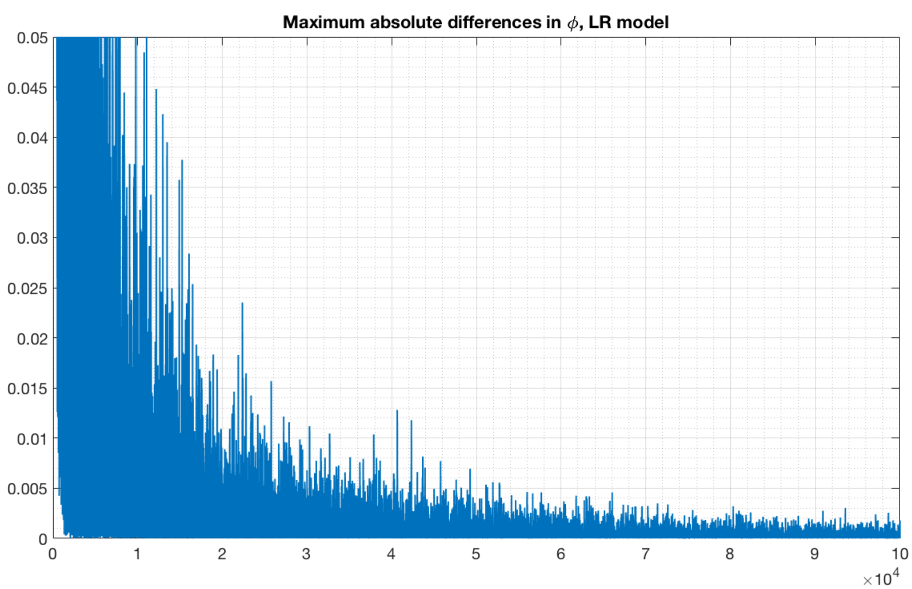


Figure 4: Convergence LR learning, $T = 100000$

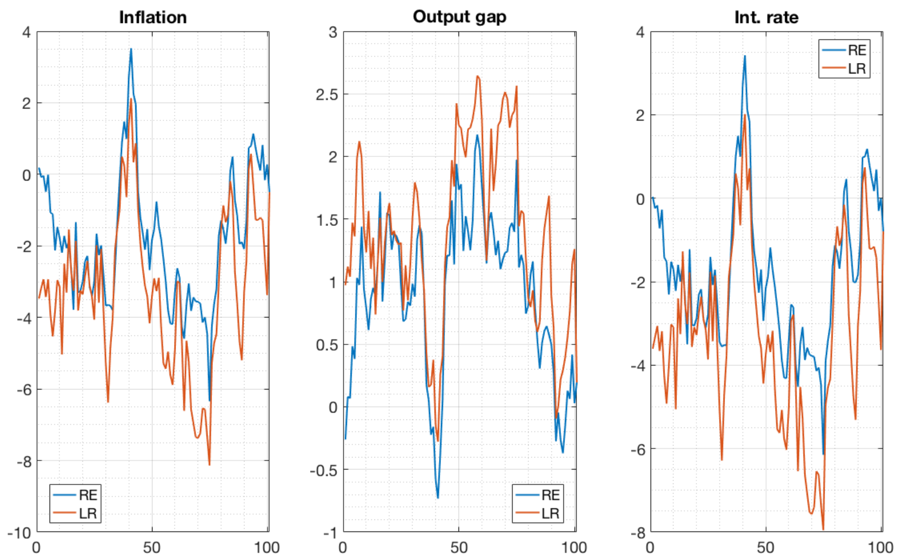


See: clearly converging!

... but, are they converging to the RE observables?

Yes, after a million periods, they are ... !

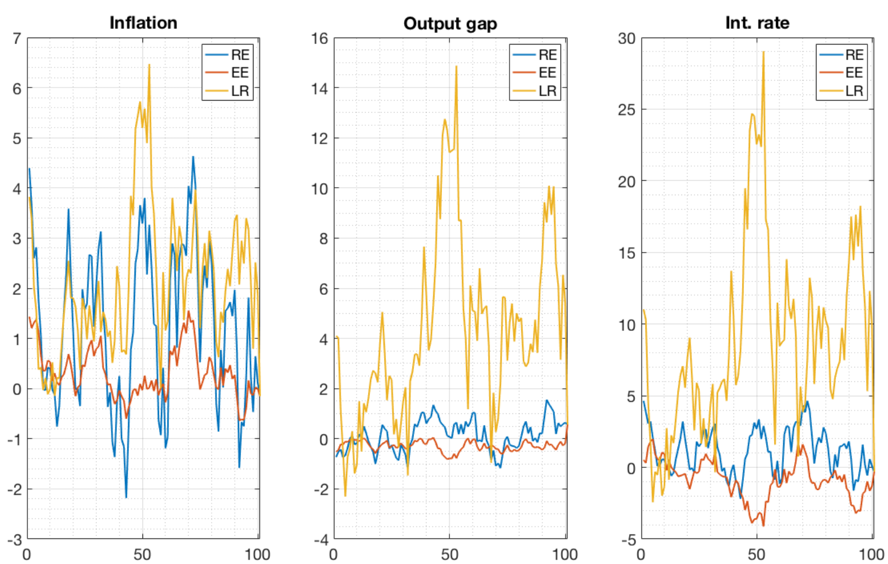
Figure 5: RE and LR models, $T = 100000$, last 100 periods



Side note: here, max abs differences in ϕ are on the order of magnitude of 10^{-5}.) After 2 million periods, they nearly overlap, but still not quite. (Diffs are at 10^{-6} now. That takes 4 min to run though!)

7.2 Learning constant only

Figure 6: All models, $T = 200000$, last 100 periods



As I feared, for $b = I_3$, this doesn't look E-stable: even after 1 million periods, neither EE nor LR converges to the RE solution. And this is despite learning clearly converging: max abs differences in the constant are $< 10^{-5}$. Or are they converging, just *much* slower? After 10 million periods, still not closer. So no.

Therefore I change the PLM such that it nests the REE. In other words, I can't set $b = I_3$, because that's not what the REE implies. Instead, I set $b = gx\ hx$, i.e. the value at which I initialized learning earlier. This means that the PLM, instead of (23), is:

$$\hat{\mathbb{E}}_t z_{t+h} = a_t + bP^{h-1}s_t \quad \forall h \geq 1 \quad \text{and} \quad b = gx\ hx \tag{33}$$

and the LR expectations, instead of (31), are given by

$$f_a(t) = \frac{1}{1-\alpha\beta}\bar{z}_t + b(I_3 - \alpha\beta P)^{-1}s_t \qquad\qquad f_b(t) = \frac{1}{1-\beta}\bar{z}_t + b(I_3 - \beta P)^{-1}s_t \tag{34}$$

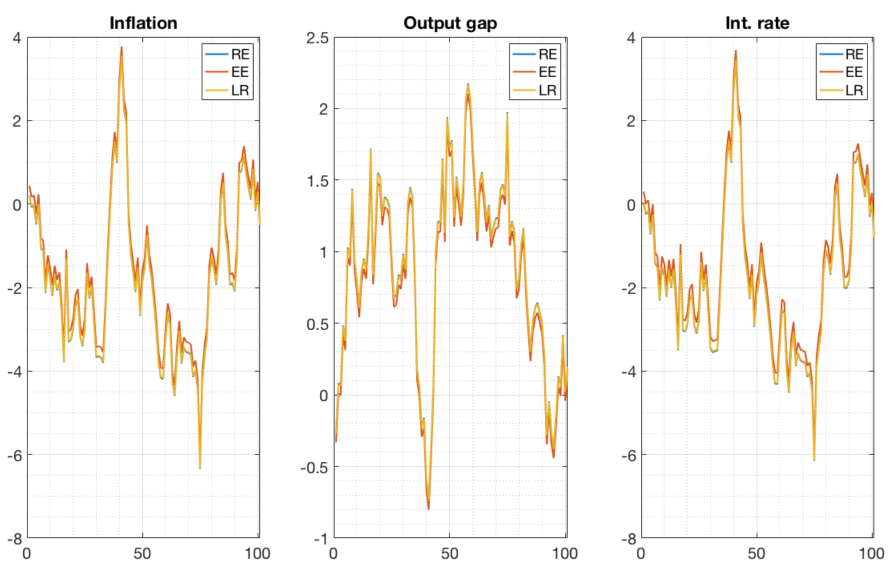
again with $b = gx\ hx$ and I'm using a and $\bar{z}$ interchangagbly.

And the learning algorithm is:

$$\bar{z}_t = \bar{z}_{t-1} + t^{-1}\big(z_t - (\bar{z}_{t-1} + bs_{t-1})\big) \qquad\qquad b = gx\ hx \tag{35}$$

When I do that, here's the last 100 periods for all three models that I get: Yay!

Figure 7: All models, $T = 100000$, last 100 periods



8 Extensions

8.1 Vector anchoring: add anchoring about all observables to LR model where agents learn only the constant

1. Take the LR learning model where agents learn the constant only
2. Add CEMP's anchoring
3. As a first step, just compare dynamics with the decreasing gain LR learning model
4. As a second step, implement the "learn only about 1 element" formulation

So the PLM is still (33), LR expectations are still (34), but the learning algorithm modifies to

$$\bar{z}_t = \bar{z}_{t-1} + k_t^{-1}\big(z_t - (\bar{z}_{t-1} + bs_{t-1})\big) \qquad b = gx\ hx \tag{36}$$

$$k_t = \mathbb{I}(k_{t-1}+1) + (1-\mathbb{I})\bar{g}^{-1} \tag{37}$$

$$\mathbb{I} = \begin{cases} 1 & \text{if } \theta_t \le \bar{\theta} \\ 0 & \text{otherwise.} \end{cases} \tag{38}$$

$$\theta_t = |\hat{\mathbb{E}}_{t-1} z_t - \mathbb{E}_{t-1}\, z_t| / (\sigma_r + \sigma_i + \sigma_u) \tag{39}$$

and I denote by the function $\mathbf{f_k}$ the anchoring mechanism given by (37)-(39).

Let's evaluate the criterion $\theta_t(\sigma_r + \sigma_i + \sigma_u)$ for this PLM and ALM:

$$\hat{\mathbb{E}}_{t-1} z_t = \bar{z}_{t-1} + bs_{t-1} \tag{40}$$

$$\begin{aligned}
\mathbb{E}_{t-1}\, z_t &= \Big(A_a f_a(t) + A_b f_b(t) + A_s s_t\Big) \\
&= \mathbb{E}_{t-1}\left(A_a \frac{1}{1-\alpha\beta}\bar{z}_t + A_a b(I_3 - \alpha\beta P)^{-1} s_t + A_b \frac{1}{1-\beta}\bar{z}_t + A_b b(I_3 - \beta P)^{-1} s_t + A_s s_t\right) \\
&= \left(A_a \frac{1}{1-\alpha\beta} + A_b \frac{1}{1-\beta}\right)\mathbb{E}_{t-1}\,\bar{z}_t + \Big(A_a b(I_3-\alpha\beta P)^{-1} + A_b b(I_3 - \beta P)^{-1} + A_s\Big)\mathbb{E}_{t-1}\, s_t \\
&= \left(A_a \frac{1}{1-\alpha\beta} + A_b \frac{1}{1-\beta}\right)\underbrace{\bar{z}_{t-1}}_{\text{discuss!}} + \Big(A_a b(I_3-\alpha\beta P)^{-1} + A_b b(I_3 - \beta P)^{-1} + A_s\Big) P s_{t-1}
\end{aligned} \tag{41}$$

This relies on the assumption that

$$\mathbb{E}_{t-1}\,\bar{z}_t = \bar{z}_{t-1} \tag{42}$$

Is this an ok assumption?

So subtracting objective expectations (41) from subjective ones (40), and taking absolute values gives us the criterion times noise:

$$\theta_t(\sigma_r + \sigma_i + \sigma_u) = |\left(I_3 - A_a \frac{1}{1-\alpha\beta} - A_b \frac{1}{1-\beta}\right)\bar{z}_{t-1} \\ + \left(b - A_a b(I_3 - \alpha\beta P)^{-1}P - A_b b(I_3 - \beta P)^{-1}P - A_s P\right)s_{t-1}| \tag{43}$$

Here is a simulation with $\bar{\theta} = 5$, still higher than CEMP's 0.029, but lower than 20 that I needed before to get decreasing gains. (3-4 seems to be the threshold value: for lower $\bar{\theta}$, gains are always constant).

Figure 8: All models, $T = 500$, last 100 periods

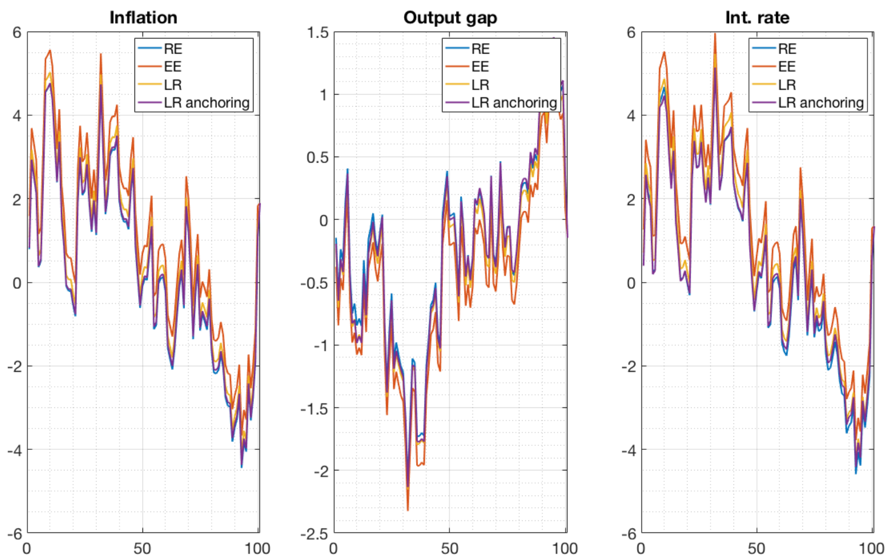
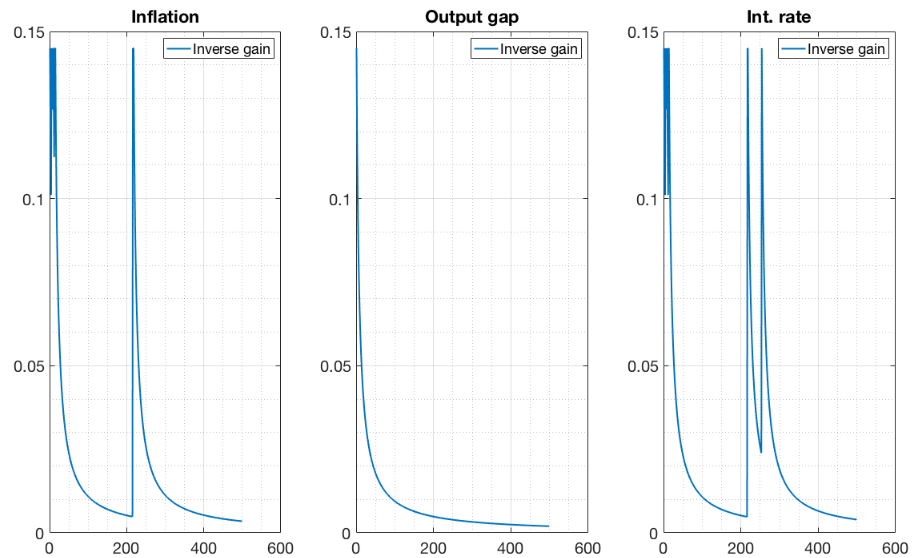


Figure 9: Gains for the anchoring model



8.2 Scalar anchoring: learning about the constant in inflation only

Let the model still be given by equation (7). But now assume that agents use the following PLM to forecast:

$$\hat{\mathbb{E}}_t z_{t+1} = \begin{bmatrix} \bar{\pi}_t \\ 0 \\ \xi \end{bmatrix} + b_1 s_t \tag{44}$$

where b_1 is the first row of b. Equation (44) reflects that agents know that the output gap doesn't have a drift, but they think that inflation does. What I'm unsure about is what I should assume for what agents think about the drift in interest rates, ξ. Two options:

- $\xi = 0$, agents also know that there's no drift in interest rates;
- $\xi = \bar{\pi}$, agents think interest rates will share the inflation drift.

For now I'll go with the first option.

Agents estimate the inflation drift using the CEMP anchoring mechanism, but now in its scalar form, the way it was designed by CEMP. Thus the criterion θ_t of equation (39) will now just be the first element of the old θ_t in (43).

Figure 10: Gains for the anchoring model, inflation drift only

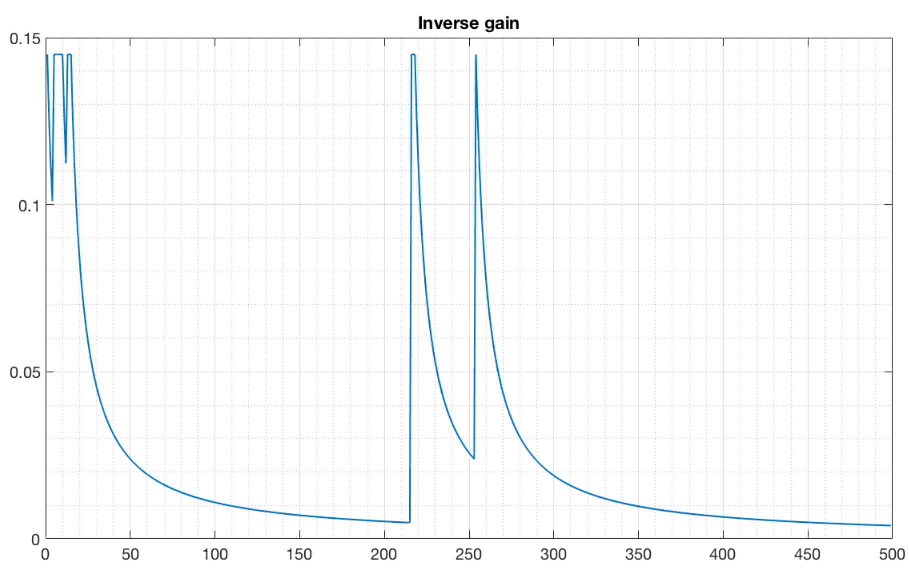


Figure 11: All models, $T = 500$, full sample

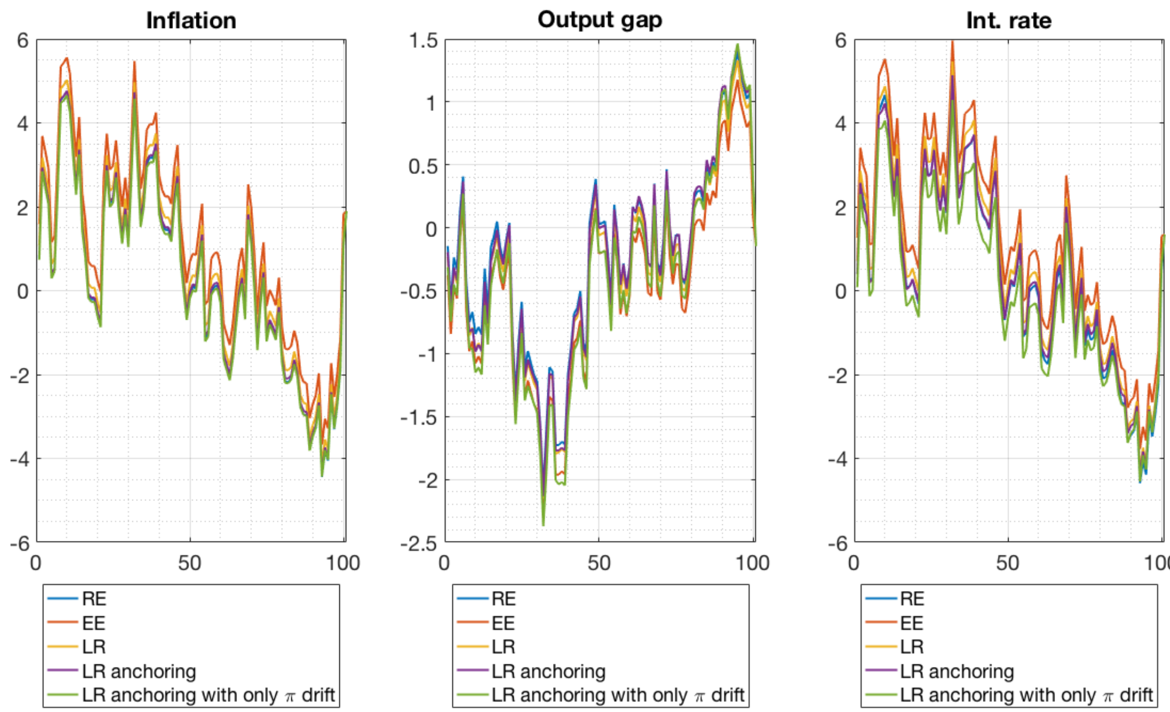
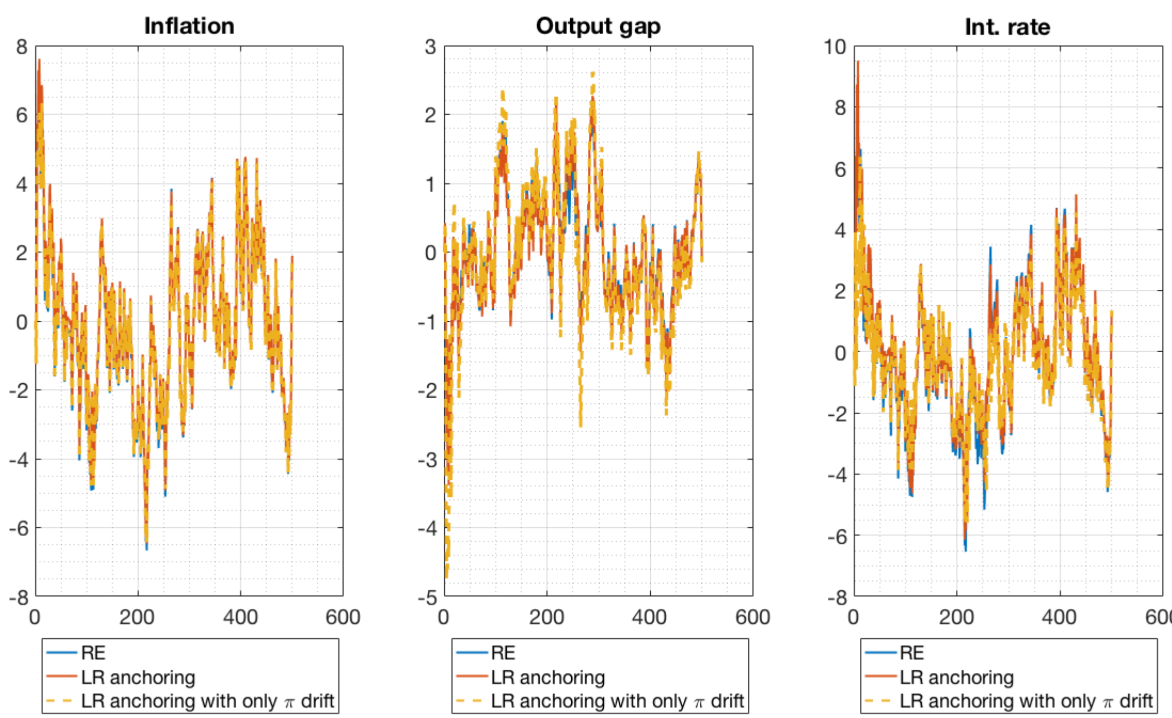


Figure 12: RE vs the vector and scalar LR anchoring models, $T = 500$, last 100 periods



9 What I've learned so far

- The more things there is to learn, the slower the convergence to RE
  - Learning slope and intercept converges slower than just intercept, c.p.
- LR expectations make convergence to RE slower and transition more volatile if agents learn slope and constant (Liam Graham gets the opposite)
  - As I add more and more period-ahead forecasts, convergence becomes slower as the economy is initially extremely volatile, so that the ALM is far from the RE solution.
  - This result reverses if agents are only learning the intercept. I think I see why: because the bulk of their LR expectations are still going to be correct then. ("A bigger fraction of their expectation will be the correct (slope) part, and the intercept will count to a smaller fraction.")
- Anchoring seems to improve the performance of the LR model, you converge faster.
  - I was expecting the opposite, and maybe that is true as well if you spend more time being deanchored ⇒ I hypothesize that whether anchoring actually improves convergence depends on the series of innovations you're subject to, such that convergence is slower the more time you spend being de-anchored. (*How should I write this btw? With or w/o hyphen? CEMP write both "unanchored" and "un-anchored," but the latter only occurs once. They never use the "de-" prefix.*) (Liam Graham finds that constant gain learning doesn't have much business cycle effects, so similar.)
- Having intercepts in the learning rule for inflation only vs. all the variables doesn't seem to matter much.
  - I expected the "scalar anchoring" specification to have faster convergence, or at least less volatility in the transition than the "vector anchoring" counterpart because the PLM is closer to the RE solution to start with. But I don't see a lot of that.
- For the same TR coefficients as for the other learning models (1.5 for both), anchoring models become explosive at around $\bar{\theta} \approx 4$.
  - I haven't had to invoke the projection facility for LR learning because I haven't had explosions.

- I'm a little unsure how to do projection when agents are only learning a constant.

- A note on volatility and convergence: I don't think that there is an iff. relationship, but I do think there is an if relationship. High volatility (relative to the RE path) does imply slower convergence because inference is based on realizations that are far from the RE path. However, slow convergence doesn't imply high volatility, because you can be close to the RE path and yet only approach it slowly (small updating steps).

10 Plan for DW presentation

- Present only LR model with anchoring
- Scalar anchoring: message is easier to convey (one gain plot instead of three)
- Or should I present simulations that contrast with LR learning with decreasing gains?
- Don't expect to have any optimal monetary policy results yet, so just plan to present it as future work