Materials 37 - Cross-section, neighborhood of zero forecast errors

Laura Gáti

July 17, 2020

Overview

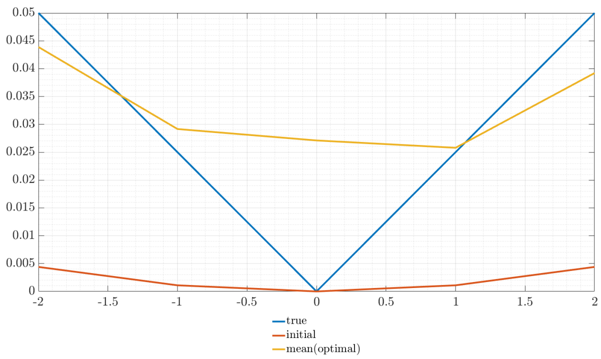
1 Simulated “true” data

(a) In the last materials, we saw that for about $N = 1000$, asymptotic behavior happens. So having a cross-section of N seems to work.

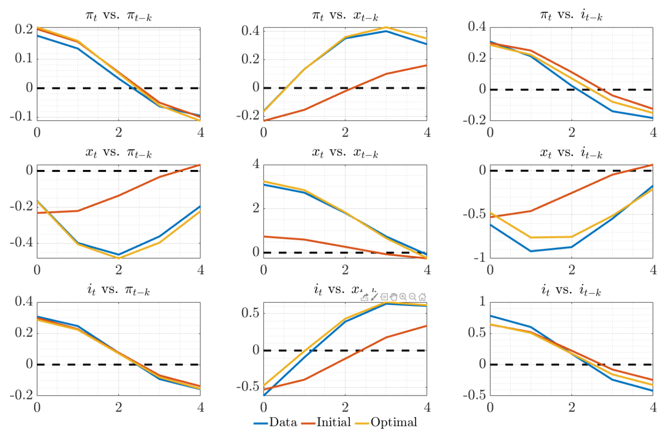
(b) We also saw that in the neighborhood of zero forecast errors the estimates don’t converge to 0. These materials aim to solve this issue. The first two figures recapitulate the problem.

1.1 Odd number of knots (5)

Figure 1: Mean estimates for $N = 100$, imposing convexity with weight 100K, truth with $nfe = 5, fe \in (-2, 2)$



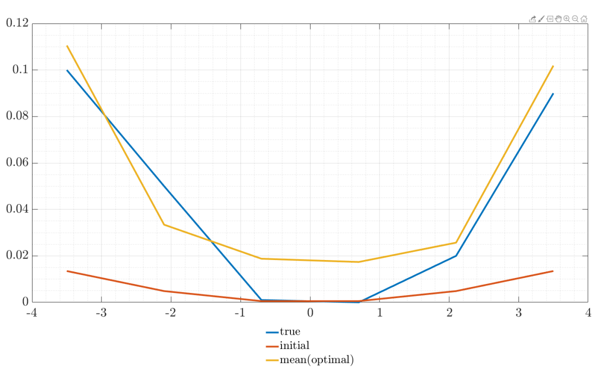
(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$



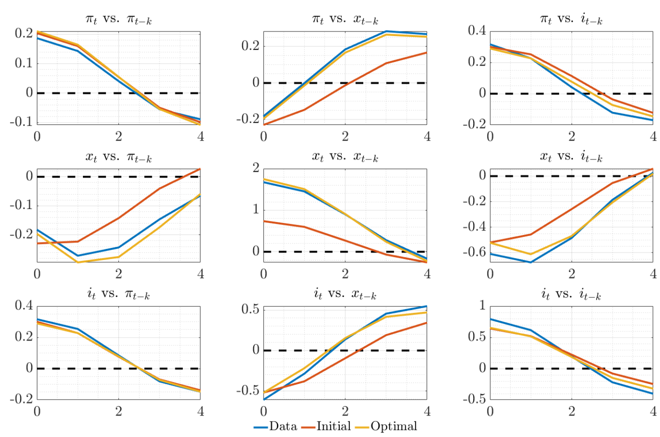
(b) Autocovariogram

1.2 Even number of knots (6)

Figure 2: Mean estimates for $N = 100$, imposing convexity with weight 100K, truth with $nfe = 6, fe \in (-3.5, 3.5)$



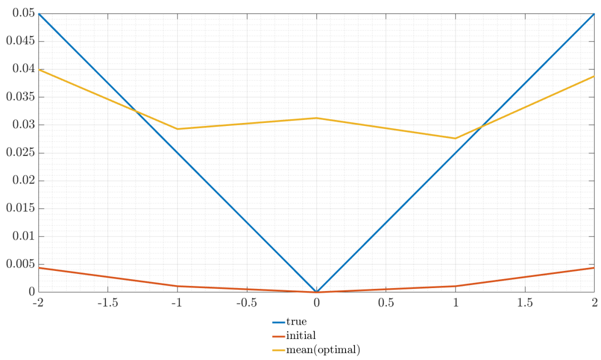
(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$



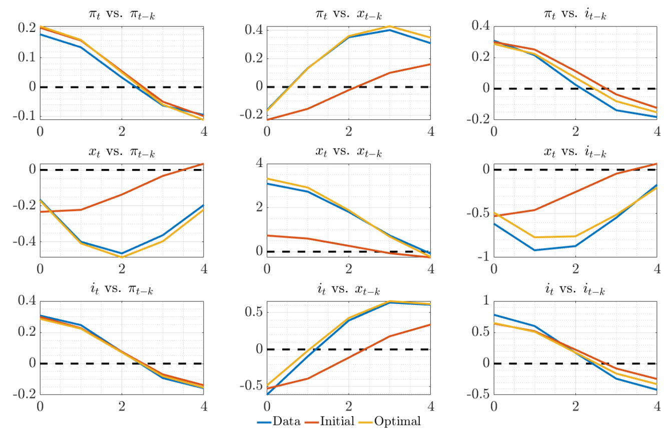
(b) Autocovariogram

1.3 Finer grid in the zero neighborhood - uneven spacing and restriction at 0

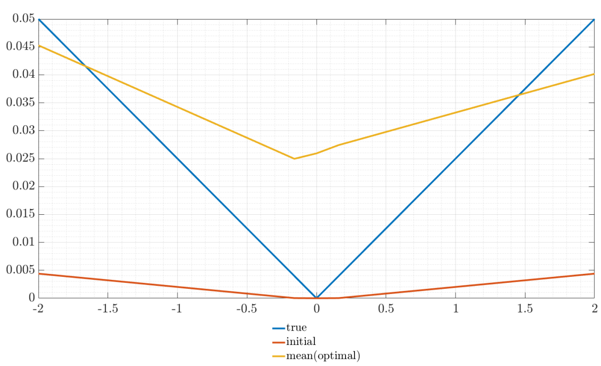
Figure 3: Mean estimates for $N = 100$, 5 breakpoints, imposing convexity w/ weight 10K, truth with $nfe = 5, fe \in (-2, 2)$



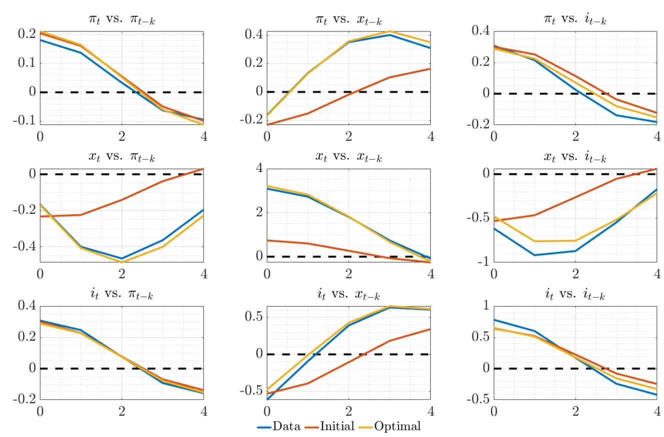
(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$, uniformly spaced



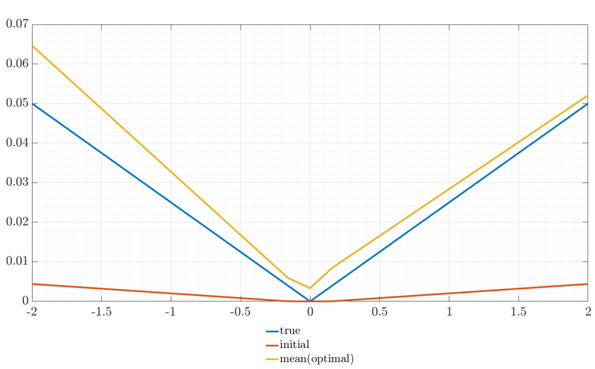
(b) Autocovariogram



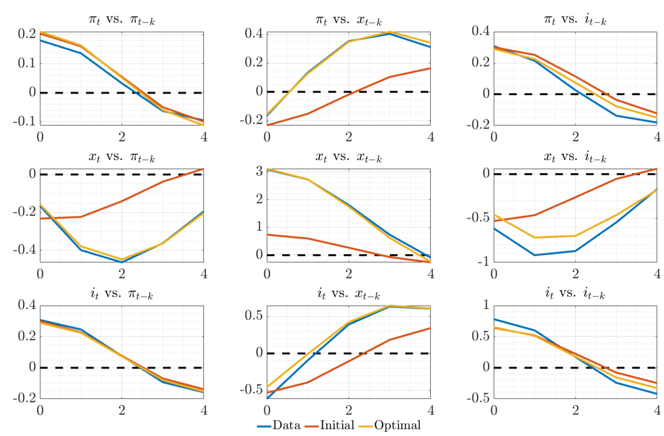
(c) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$, concentrated around 0



(d) Autocovariogram

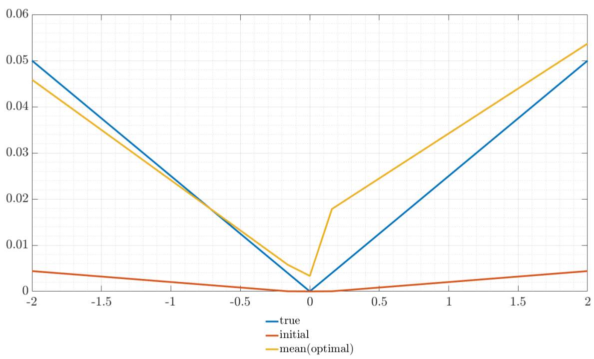


(e) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$,
concentrated around 0, 0 at 0 imposed with weight 1000

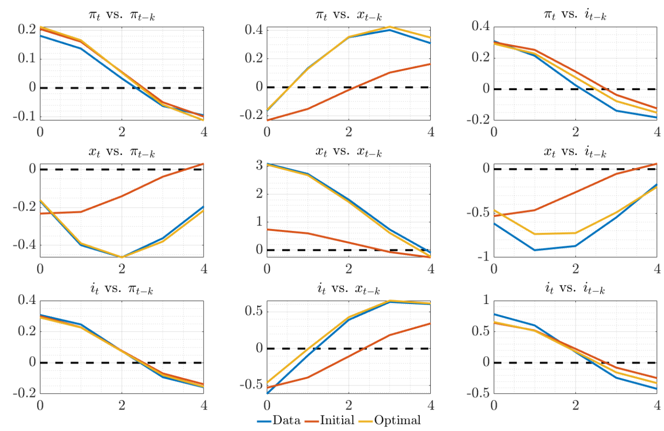


(f) Autocovariogram

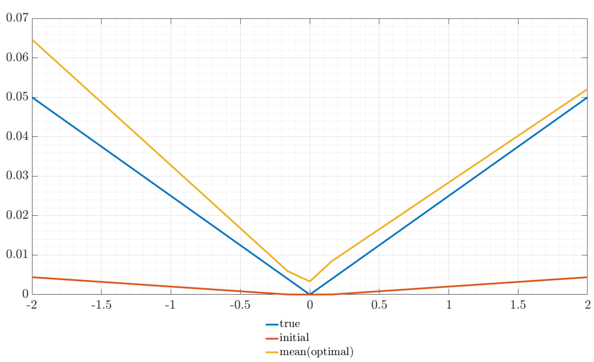
Figure 4: Can I decrease the weight on the convexity moment? Mean estimates for $N = 100$, breakpoints concentrated around 0, 0 at 0 imposed with weight 1000, imposing convexity with variable weight, truth with $nfe = 5, fe \in (-2, 2)$



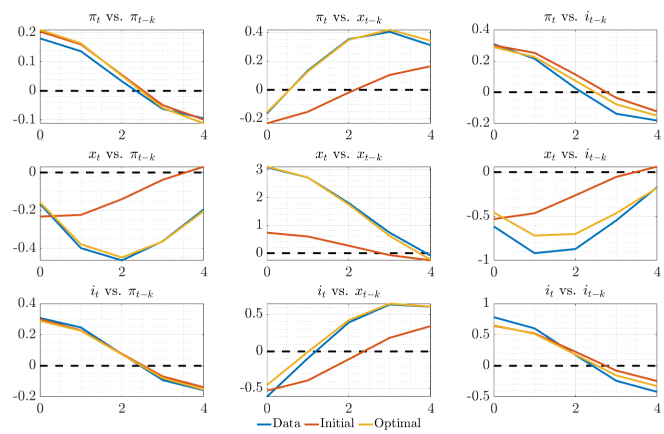
(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$,
5 breakpoints, convexity weight 1000



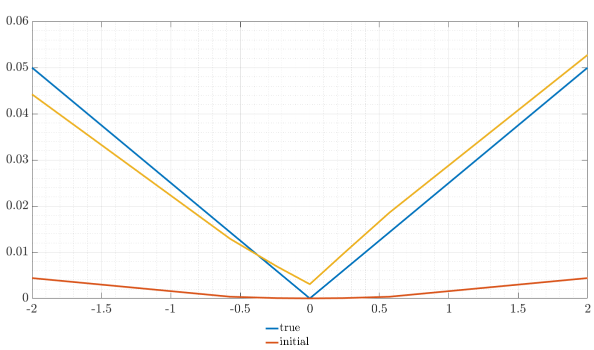
(b) Autocovariogram



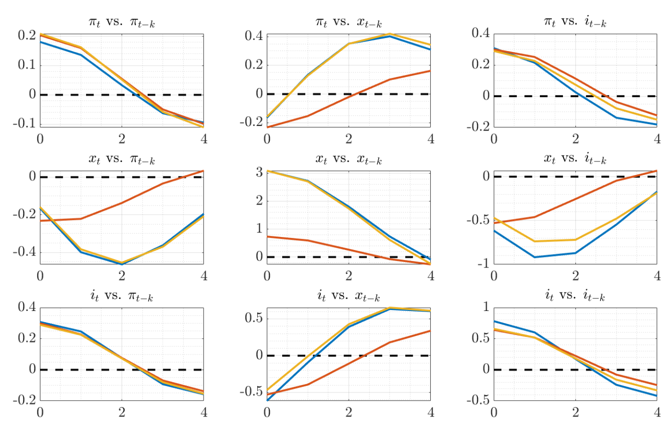
(c) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$,
5 breakpoints, convexity weight 10000



(d) Autocovariogram



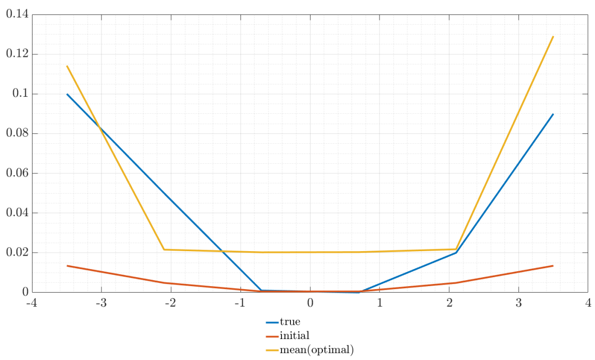
(e) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$,
7 breakpoints, convexity weight 10000



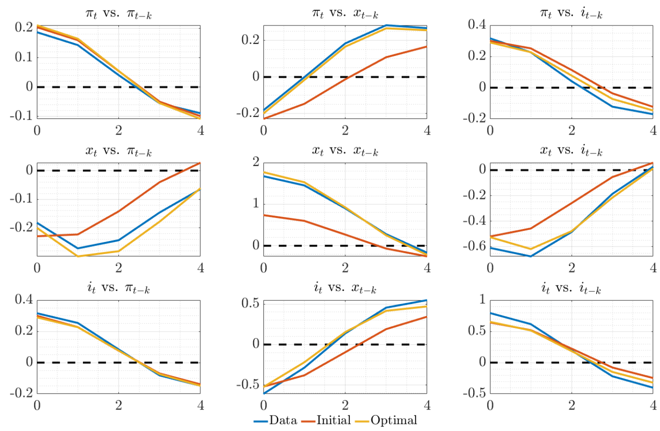
(f) Autocovariogram

1.4 Is the V-shape forced? A U-shaped truth

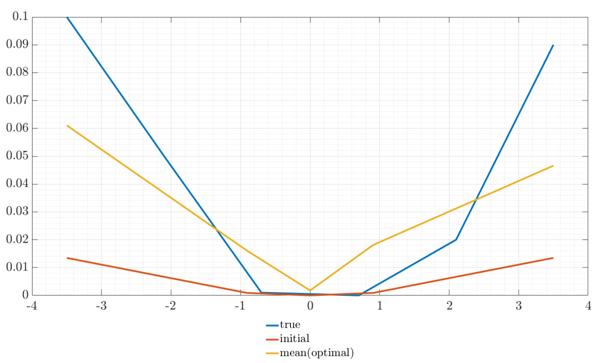
Figure 5: Mean estimates for $N = 100$, imposing convexity with weight 10K, knots concentrated around 0, 0 at 0 imposed with weight 1000; truth with $nfe = 6, fe \in (-3.5, 3.5)$



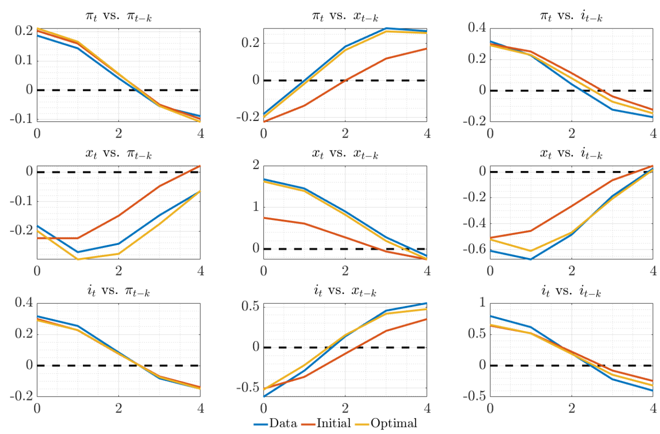
(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$,
6 knots, uniform grid, 0 at 0 restriction not imposed



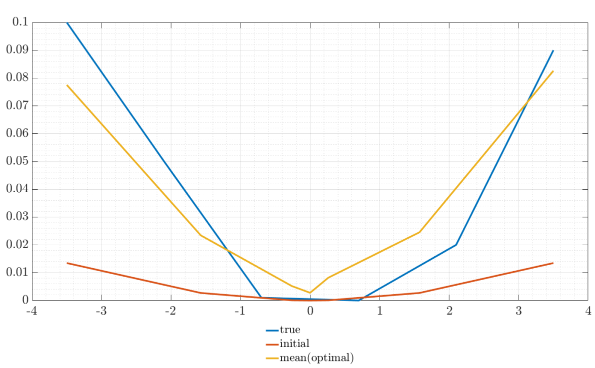
(b) Autocovariogram



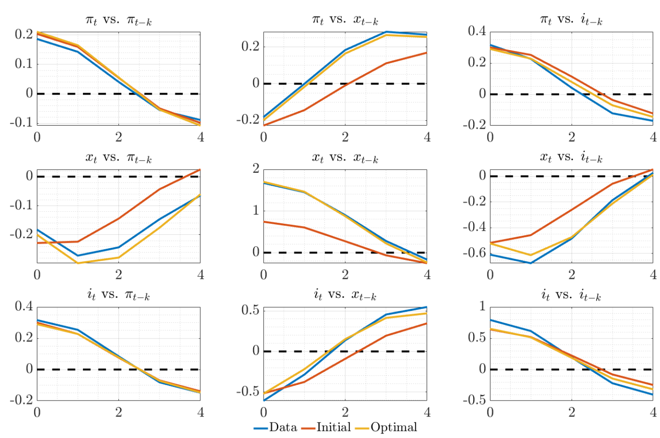
(c) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$, 5 knots



(d) Autocovariogram

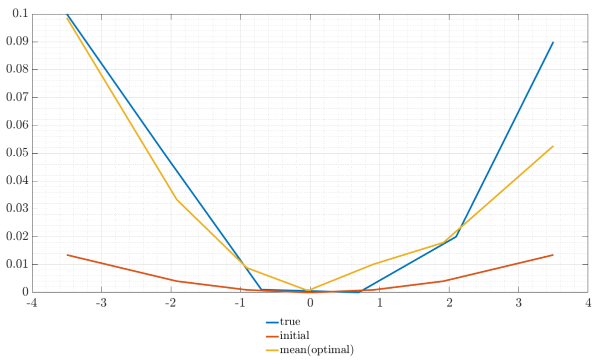


(e) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$, 7 knots

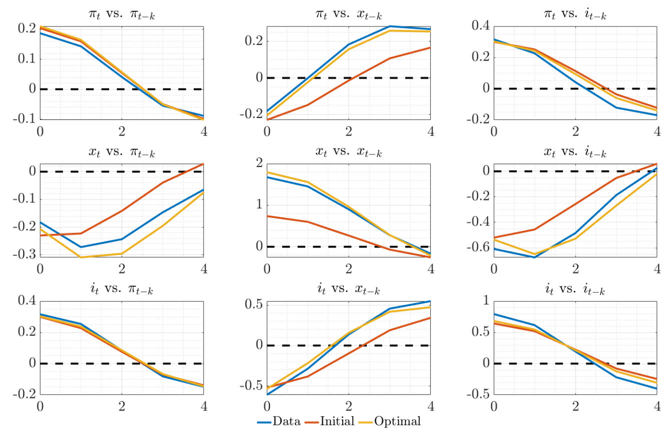


(f) Autocovariogram

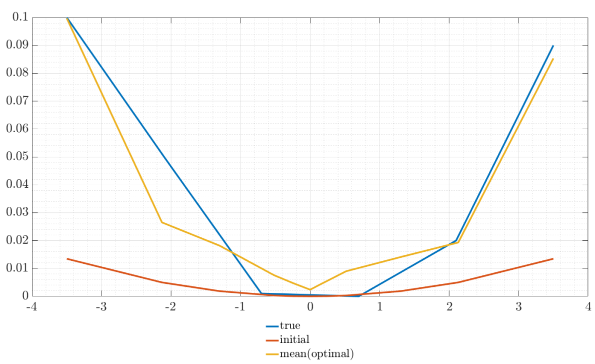
Figure 6: Continued



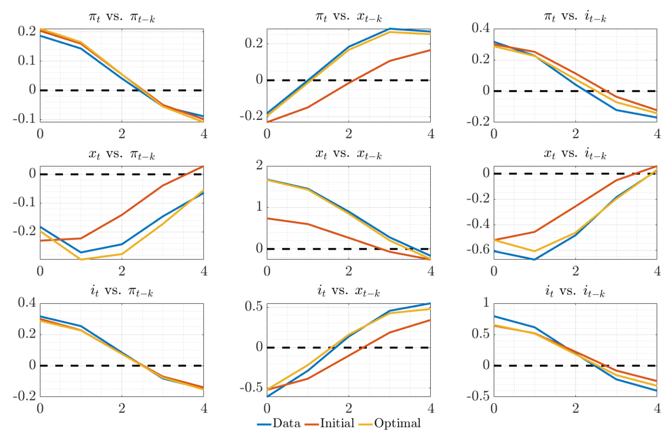
(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$, 9 knots



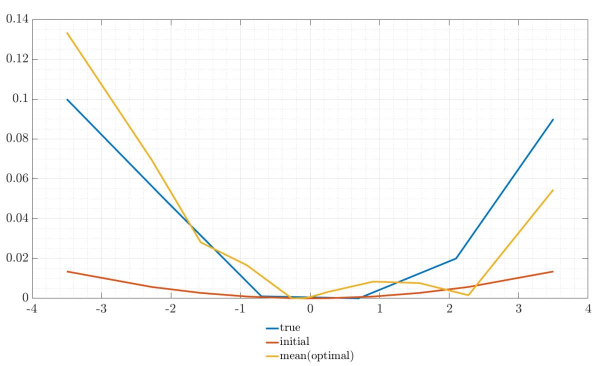
(b) Autocovariogram



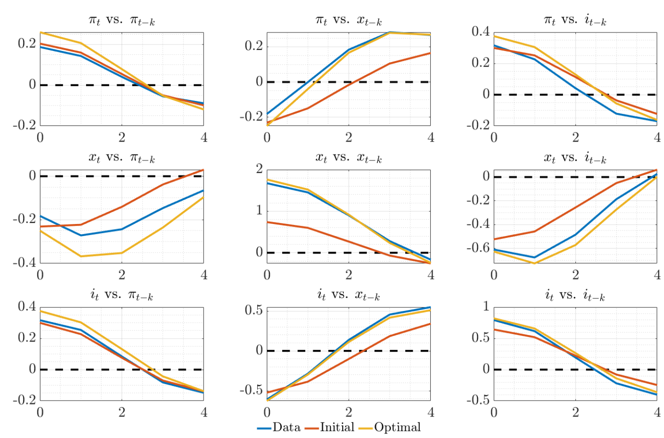
(c) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$, 11 knots



(d) Autocovariogram



(e) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$, 13 knots

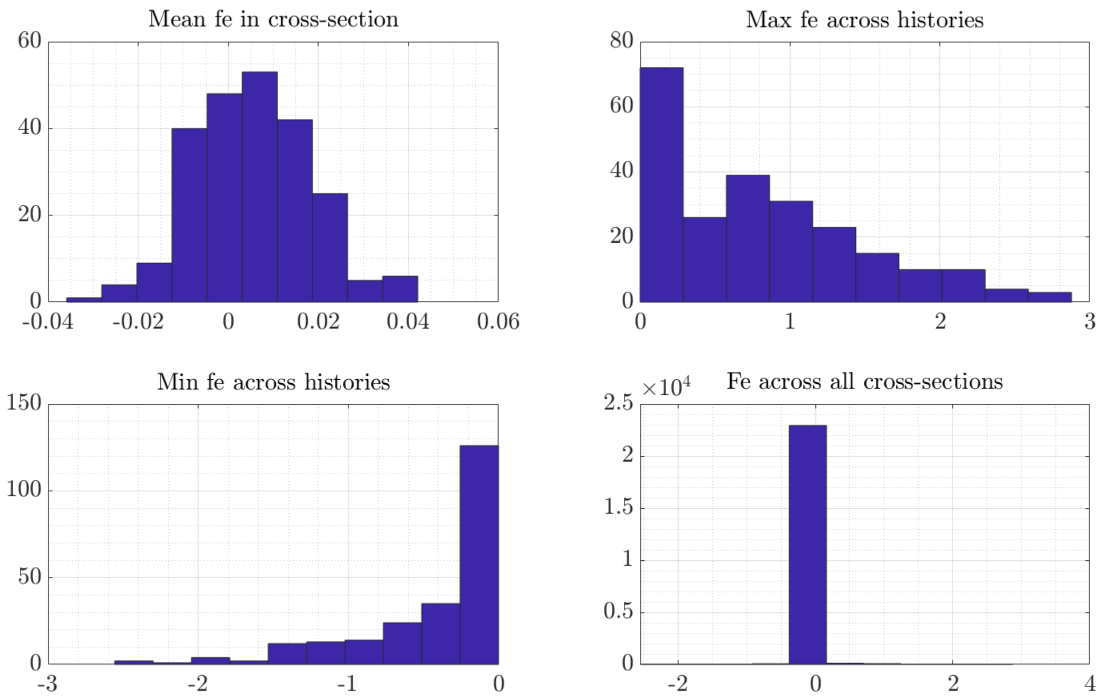


(f) Autocovariogram

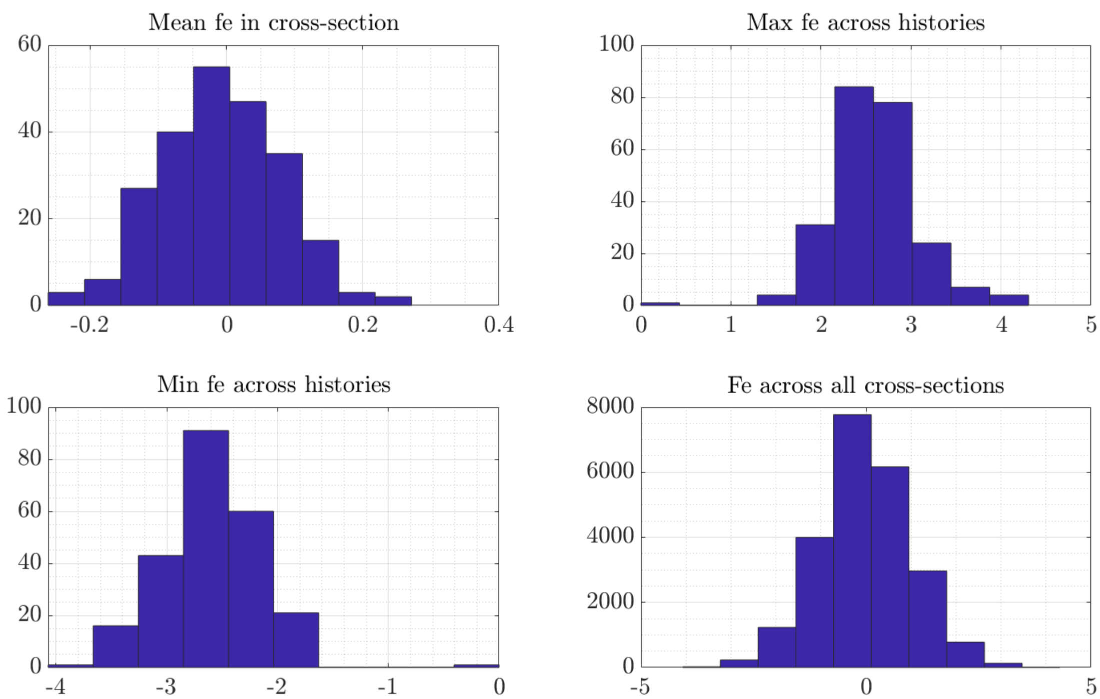
- I still seem to see the issue that the 0-neighborhood isn't well identified, but it doesn't seem grave.
- 11 knots or more seems underidentified, 9 may be the max.

1.5 Are there no forecast errors in the 0-neighborhood?

Figure 7: Distribution of forecast errors in the cross-section, $N = 100$



(a) truth with $nfe = 6, fe \in (-3.5, 3.5)$ 13 knots



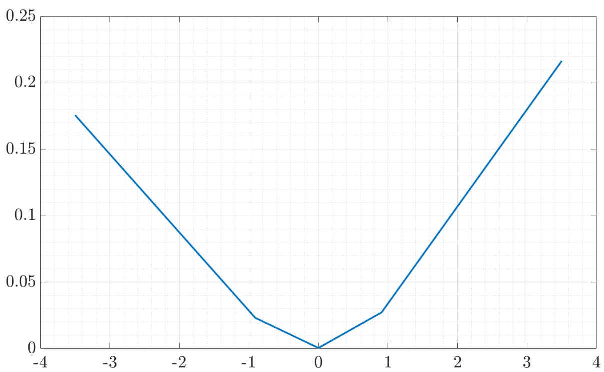
(b) truth with $nfe = 6, fe \in (-3.5, 3.5)$ 5 knots

→ Yes there are! In fact, too many knots make forecast errors more concentrated around 0.

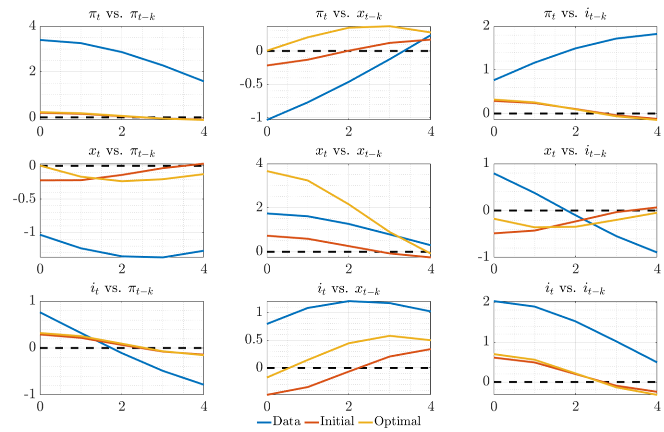
2 Autocovariogram for real data

2.1 # knots

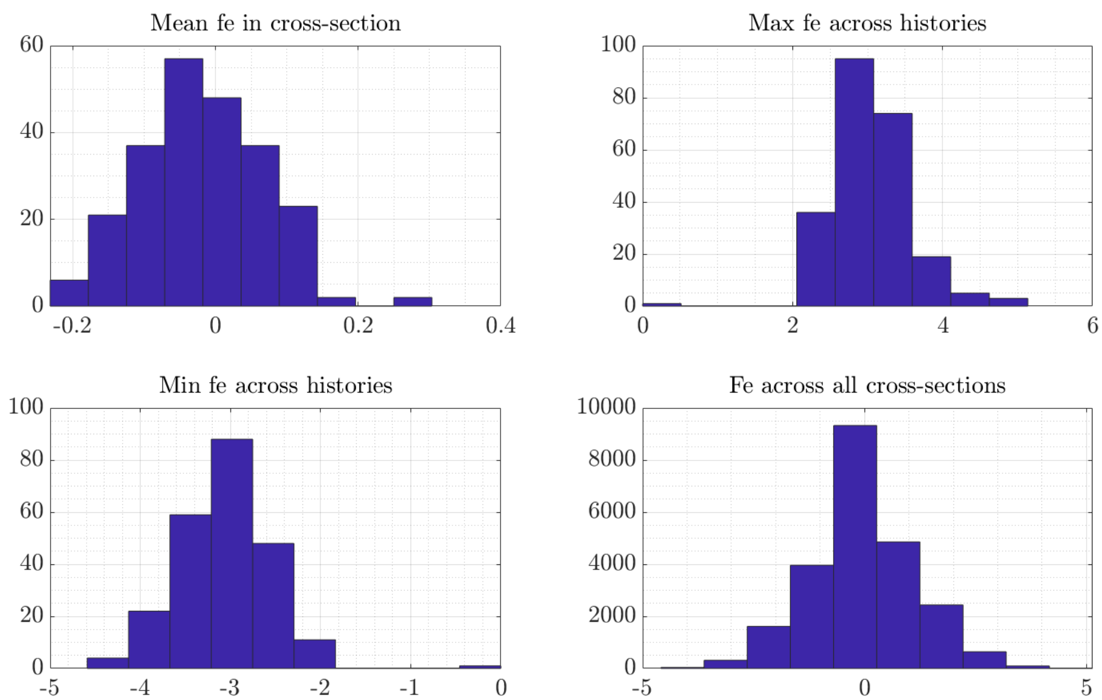
Figure 8: Mean estimated parameters, autocovariogram and forecast errors for 5 breakpoints



(a) $\bar{\bar{\alpha}}$



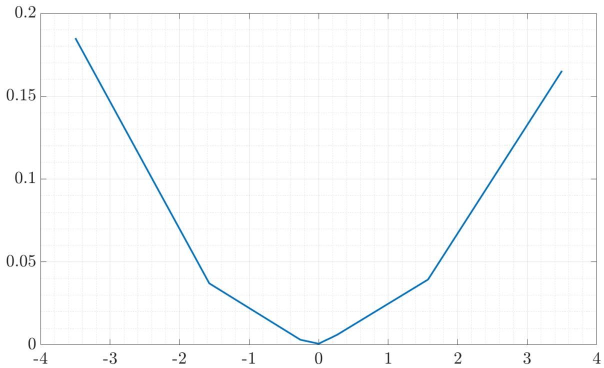
(b) Autocovariogram at an average of the N shock histories



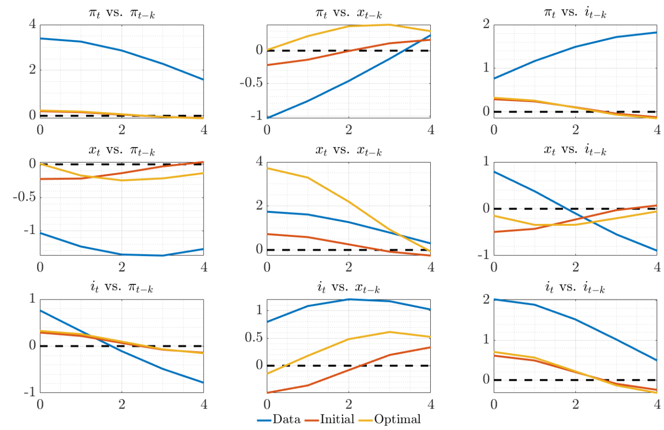
(c) Distribution of forecast errors

Constant estimation parameters: cross-section of size $N = 100, fe \in (-3.5, 3.5)$, convexity imposed with weight 10K, points unevenly spaced (denser at 0), 0 at 0 imposed with weight 1000, mean moment not imposed

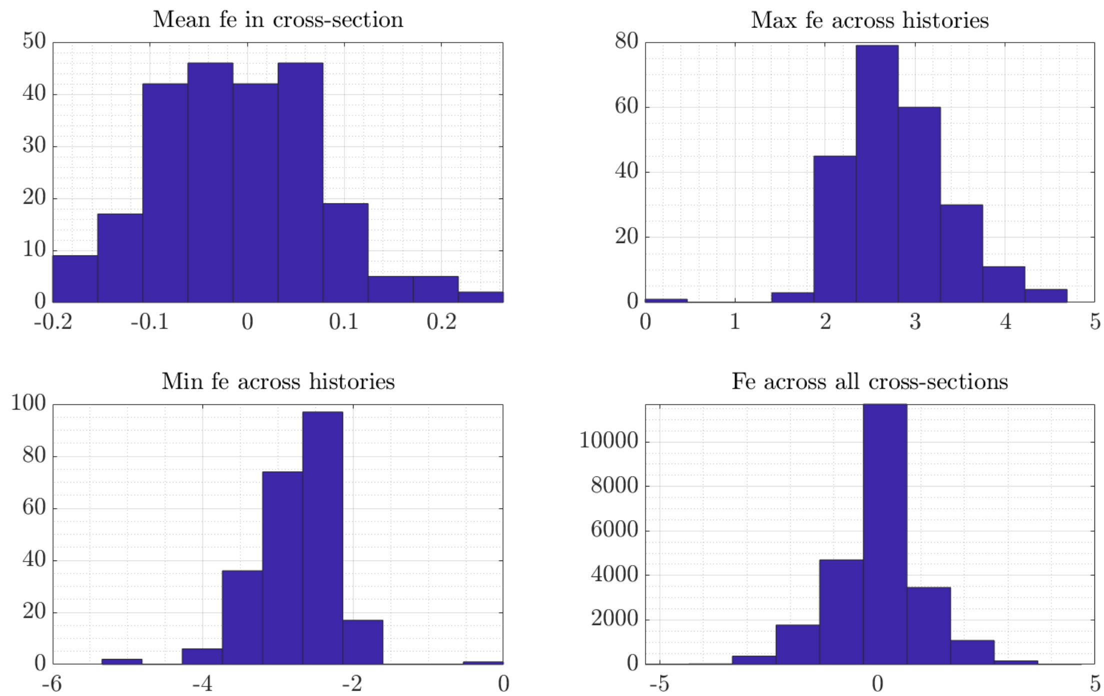
Figure 9: Mean estimated parameters, autocovariogram and forecast errors for 7 breakpoints



(a) $\bar{\bar{\alpha}}$



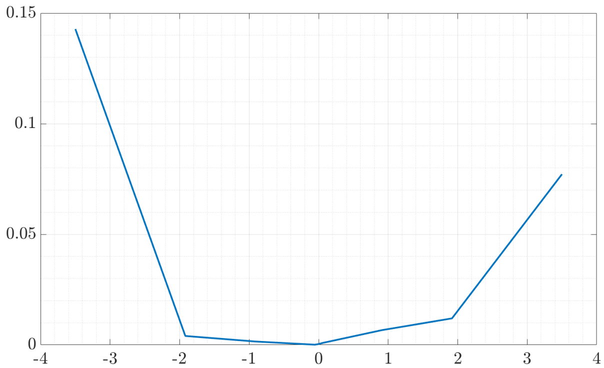
(b) Autocovariogram at an average of the N shock histories



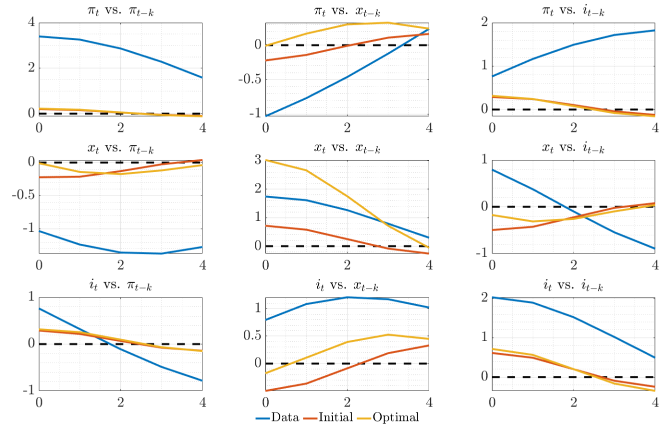
(c) Distribution of forecast errors

Constant estimation parameters: cross-section of size $N = 100, fe \in (-3.5, 3.5)$, convexity imposed with weight 10K, points unevenly spaced (denser at 0), 0 at 0 imposed with weight 1000, mean moment not imposed

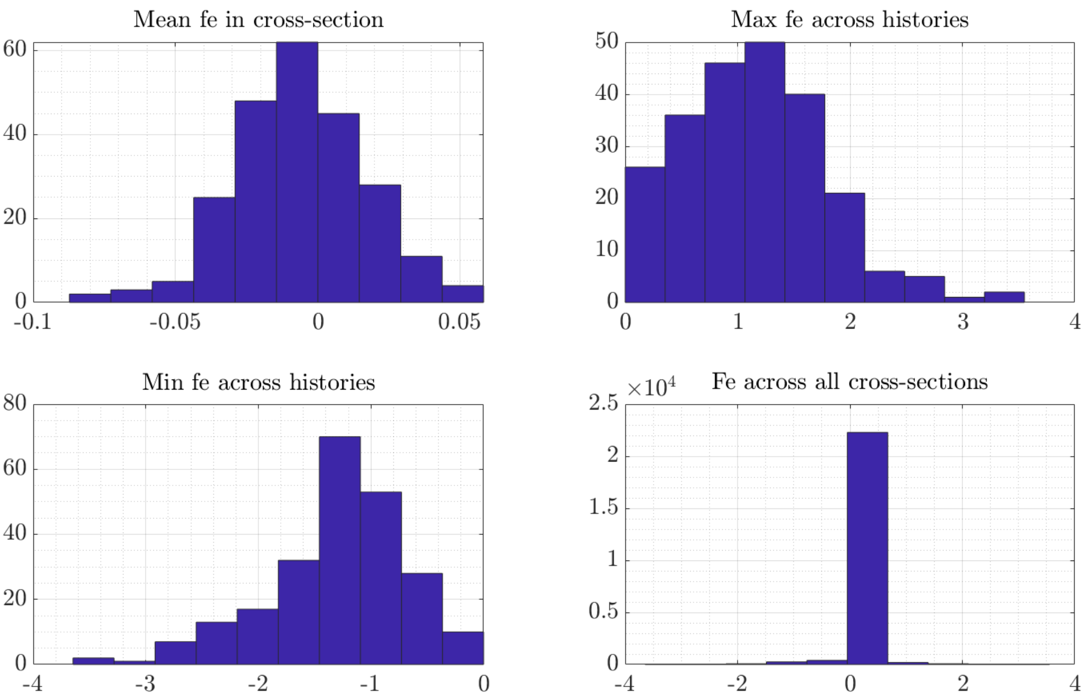
Figure 10: Mean estimated parameters, autocovariogram and forecast errors for 9 breakpoints



(a) $\bar{\bar{\alpha}}$



(b) Autocovariogram at an average of the N shock histories



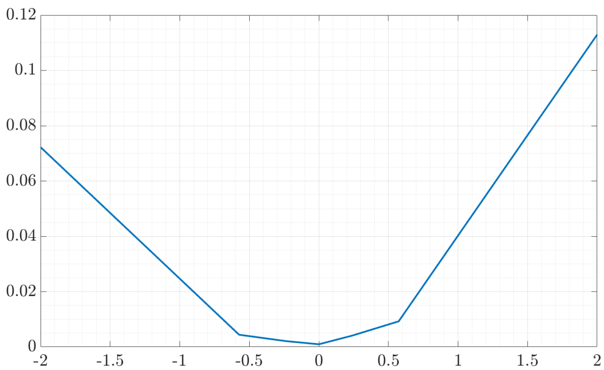
(c) Distribution of forecast errors

Constant estimation parameters: cross-section of size $N = 100, fe \in (-3.5, 3.5)$, convexity imposed with weight 10K, points unevenly spaced (denser at 0), 0 at 0 imposed with weight 1000, mean moment not imposed

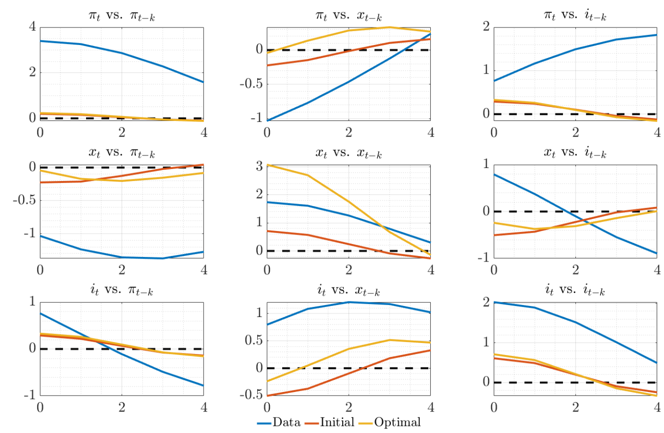
Note: only 6/100 converged $\rightarrow$ seems underidentified. Pick 7 knots as default.

2.2 Support of forecast errors

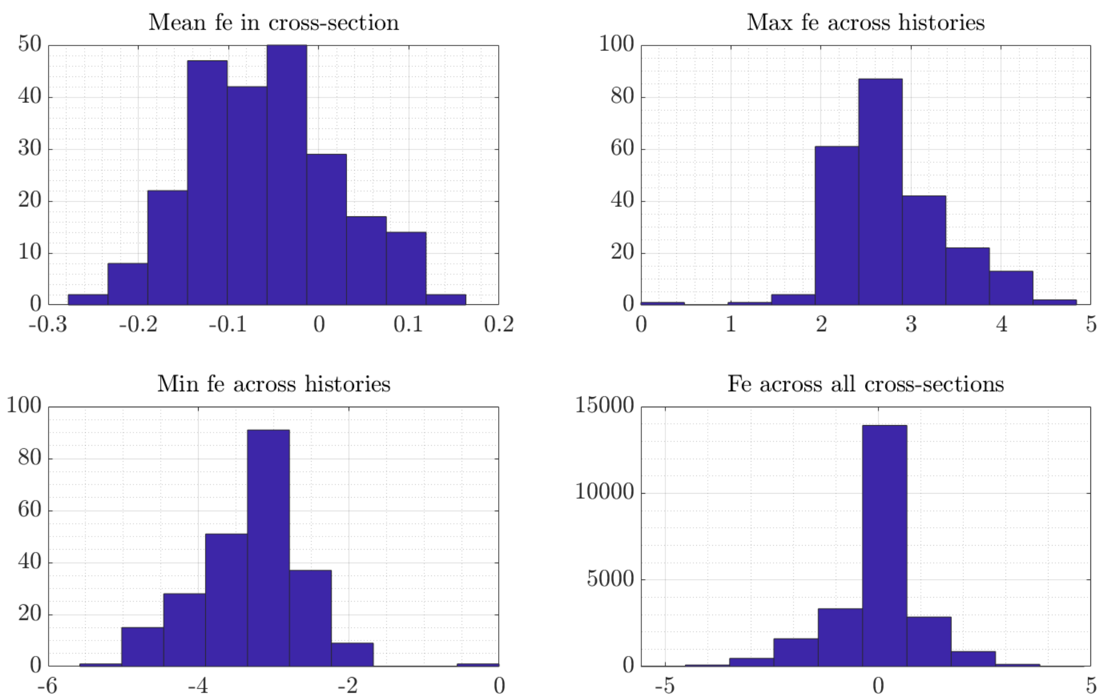
Figure 11: Mean estimated parameters, autocovariogram and forecast errors for $fe \in (-2, 2)$



(a) $\bar{\tilde{\alpha}}$



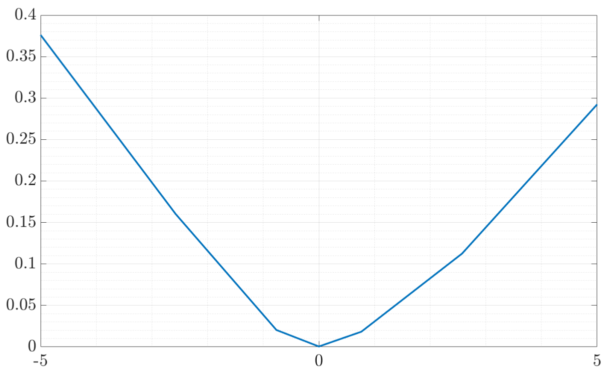
(b) Autocovariogram at an average of the N shock histories



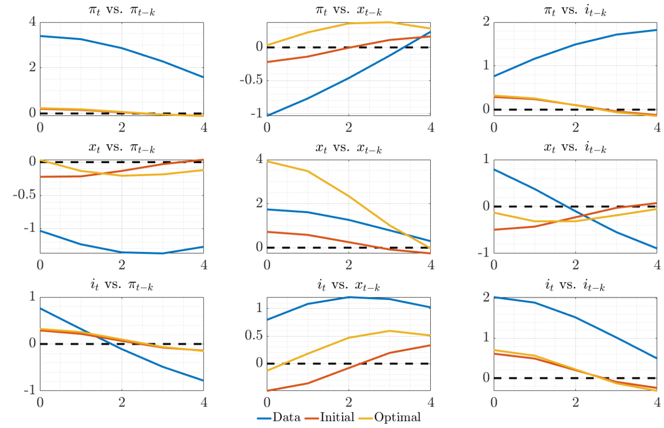
(c) Distribution of forecast errors

Constant estimation parameters: cross-section of size $N = 100$, 7 knots, convexity imposed with weight 10K, points unevenly spaced (denser at 0), 0 at 0 imposed with weight 1000, mean moment not imposed

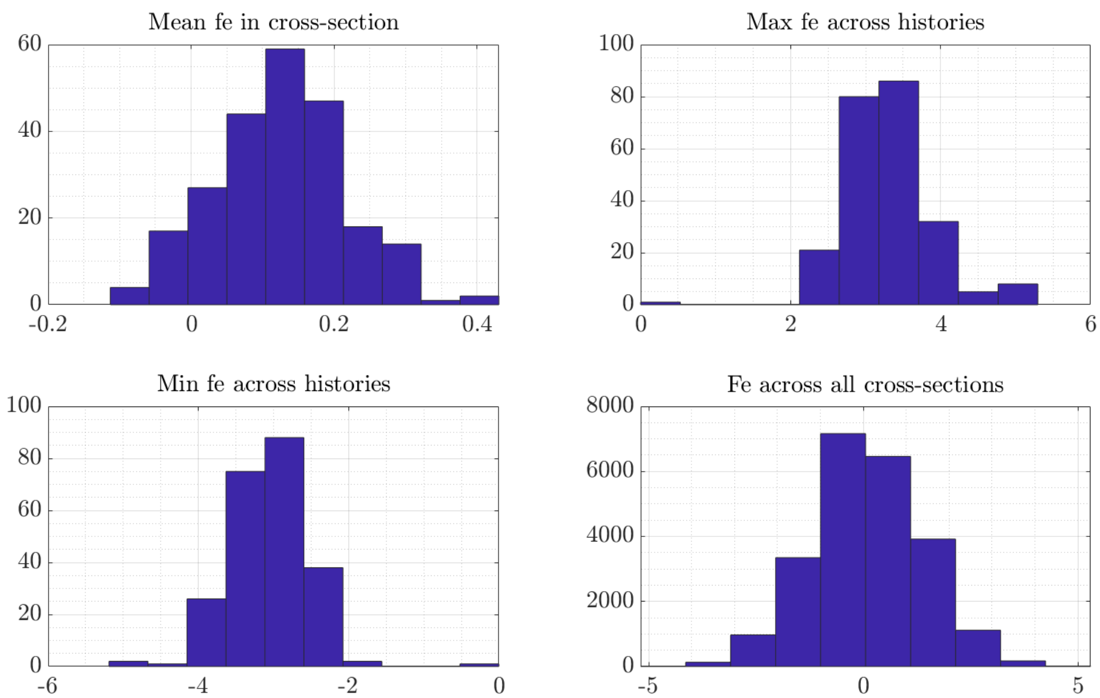
Figure 12: Mean estimated parameters, autocovariogram and forecast errors for $fe \in (-5, 5)$



(a) $\bar{\bar{\alpha}}$



(b) Autocovariogram at an average of the N shock histories



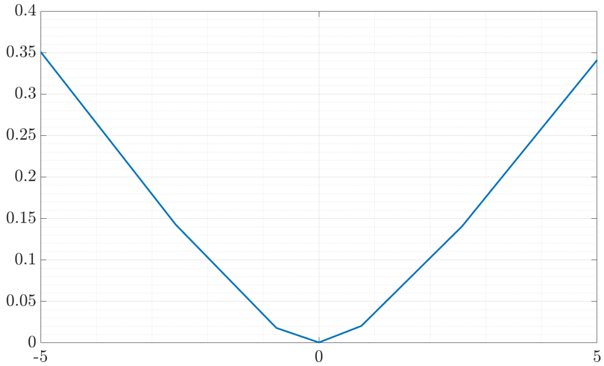
(c) Distribution of forecast errors

Constant estimation parameters: cross-section of size $N = 100$, 7 knots, convexity imposed with weight 10K, points unevenly spaced (denser at 0), 0 at 0 imposed with weight 1000, mean moment not imposed

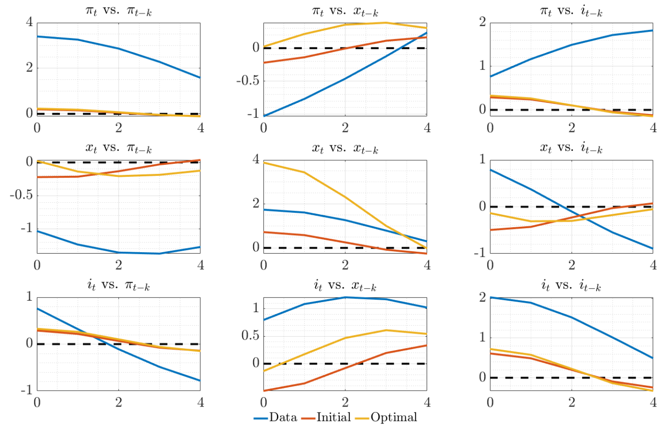
$\rightarrow$ It seems robust to the choice of forecast-error-support!

2.3 Repeat the last specification with $N = 1000$

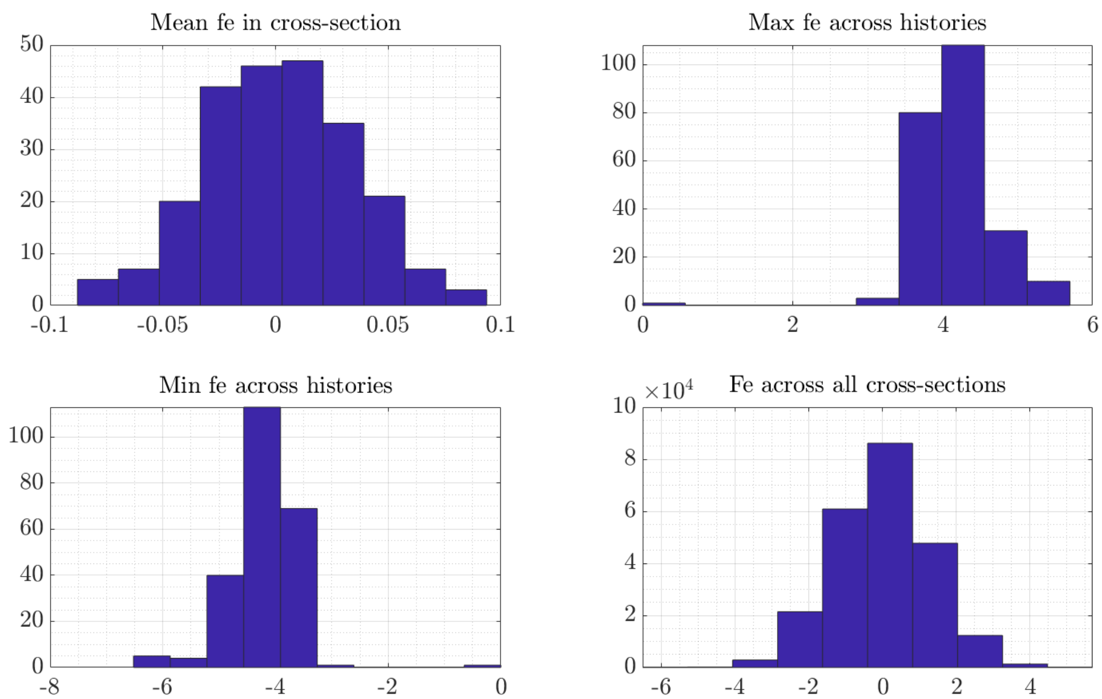
Figure 13: Mean estimated parameters, autocovariogram and forecast errors for cross-section of size $N = 1000$



(a) $\bar{\hat{\alpha}}$



(b) Autocovariogram at an average of the N shock histories



(c) Distribution of forecast errors

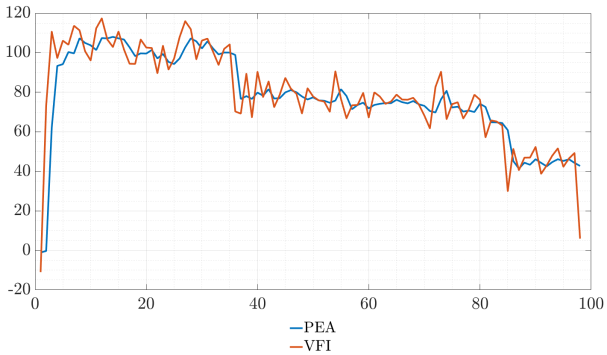
Constant estimation parameters: 7 knots, $fe \in (-5, 5)$,convexity imposed with weight 10K, points unevenly spaced (denser at 0), 0 at 0 imposed with weight 1000, mean moment not imposed

- Took 96.33 min.
- Seems robust.
- $\text{mean}(\hat{\alpha}) = (0.3509; 0.1427; 0.0180; 0.0005; 0.0204; 0.1406; 0.3411)$

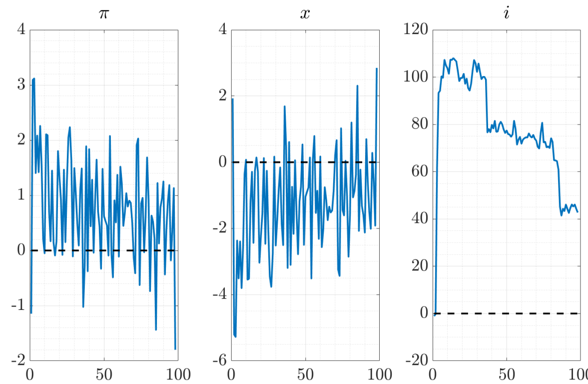
3 Redo PEA

using $\alpha = \text{mean}(\hat{\alpha}) = (0.3509; 0.1427; 0.0180; 0.0005; 0.0204; 0.1406; 0.3411)$

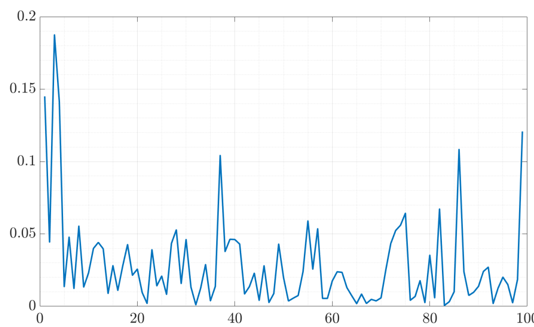
Figure 14: Optimal policy solution conditional on one particular evolution of shocks



(a) Policy function: PEA against VFI



(b) Observables



(c) Gain

Using the same parameter values as for estimation, using $\text{mean}(\hat{\alpha}) = (0.3509; 0.1427; 0.0180; 0.0005; 0.0204; 0.1406; 0.3411)$ as approximation for the anchoring function, no monpol shocks because Taylor rule is not in effect, thus it is also assumed not to be known. Shock sequence 2×100 generated using `rng(0)`.

3.1 Why is policy nonstationary?

Figure 15: Tracking long-run expectations? Evolution of $\bar{\pi}$

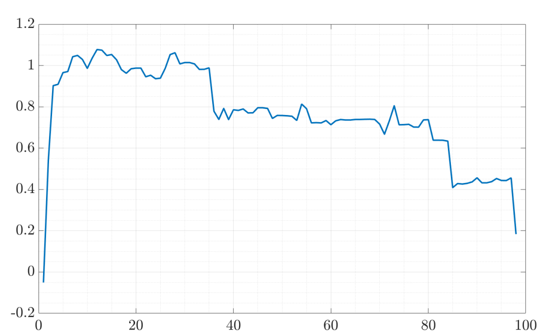
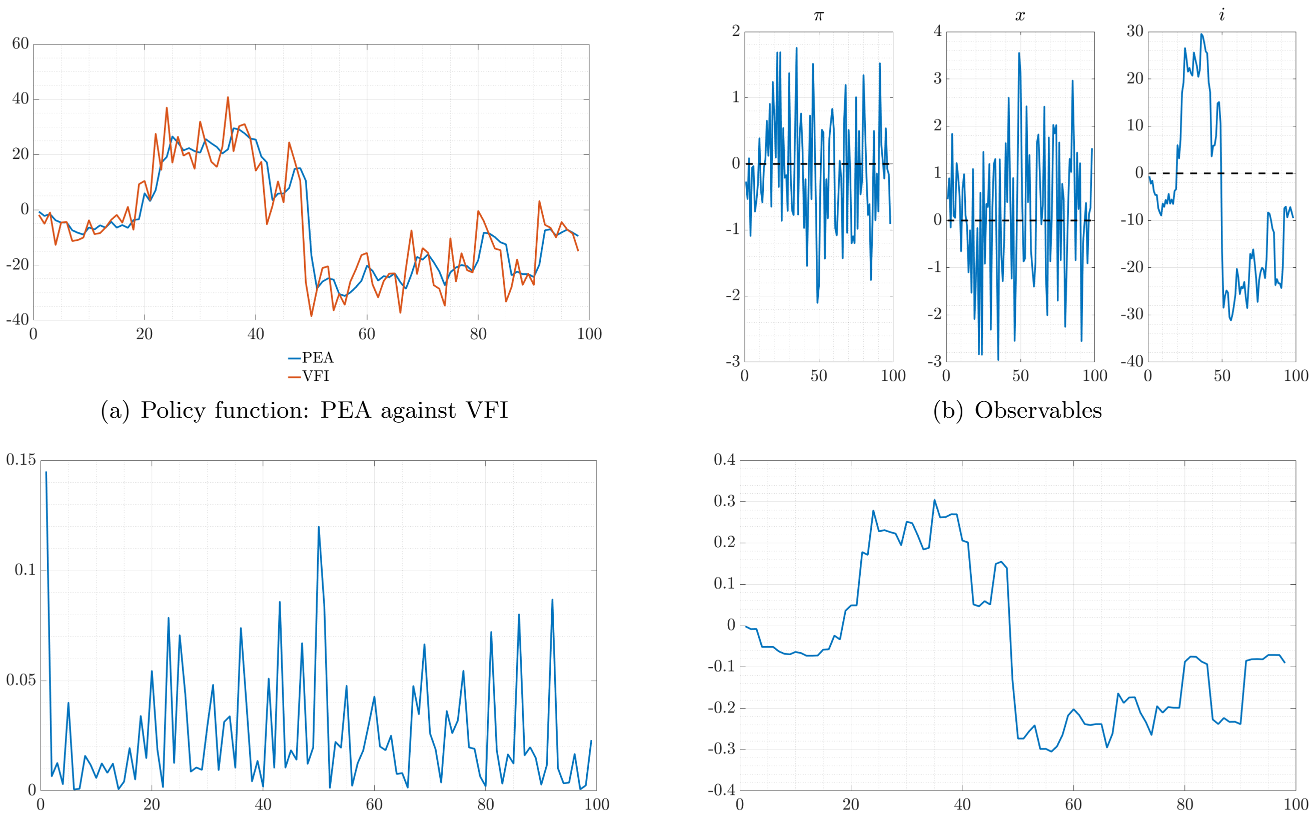


Figure 16: Optimal policy solution conditional on a different, `rng(2)` evolution of shocks



(a) Policy function: PEA against VFI

(b) Observables

(c) Gain

(d) $\bar{\pi}$

- It sure looks like policy was chasing long-run expectations.
- Running PEA for different PLMs, in particular, constant-only-inflation-only (default) vs. constant-only, all observables learning, I find that the nonstationary policy only occurs for my default policy. This seems to stem from the fact that if long-run expectations of x and i are also allowed to fluctuate, they, taken together with $\bar{\pi}$ cancel each other out. In particular, the long-run expectation of i I think does the stabilization that i itself needs to do if expectations of i aren't allowed to fluctuate.

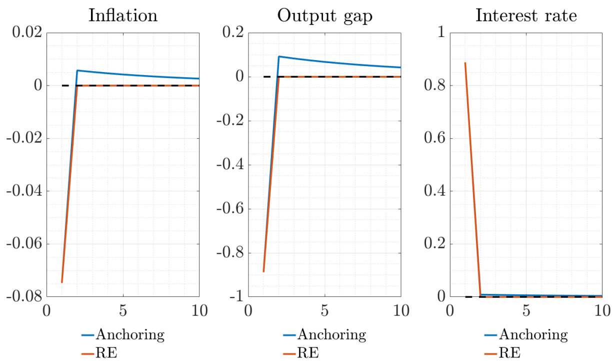
3.2 Why is VFI-policy more volatile than that of PEA?

Is this because the spline, as a spectral approximation method, has a tendency to oscillate around kinks? Or because the approximation is very rough (4 gridpoints (or 6, or 8) for $\bar{\pi}$, 2 each for the two shocks at t and at $t-1$)?

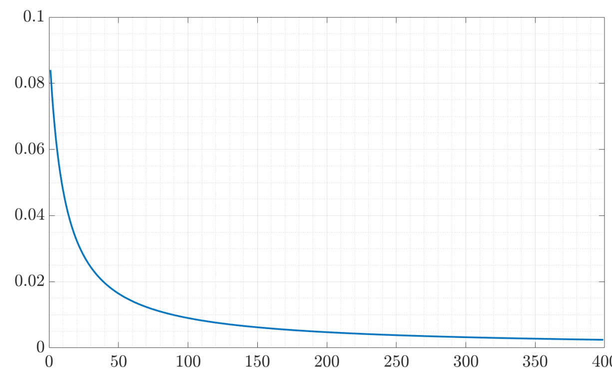
4 Impulse responses to iid monpol shocks across a wide range of learning models

$T = 400, N = 100, n_{drop} = 5$, shock imposed at $t = 25$, calibration as above, Taylor rule assumed to be known, PLM = learn constant only, of inflation only.

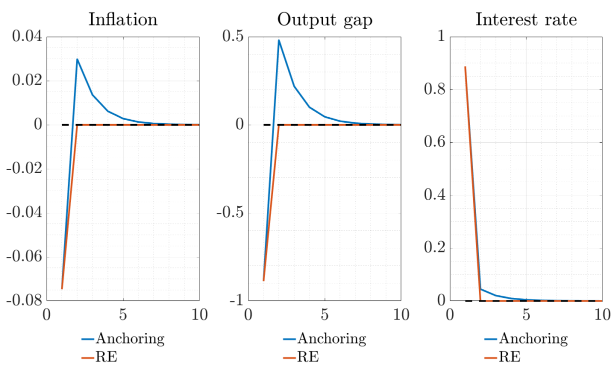
Figure 17: IRFs and gain history (sample means)



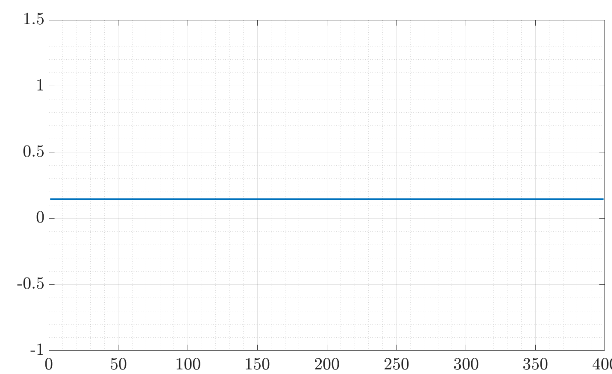
(a) Decreasing gain learning



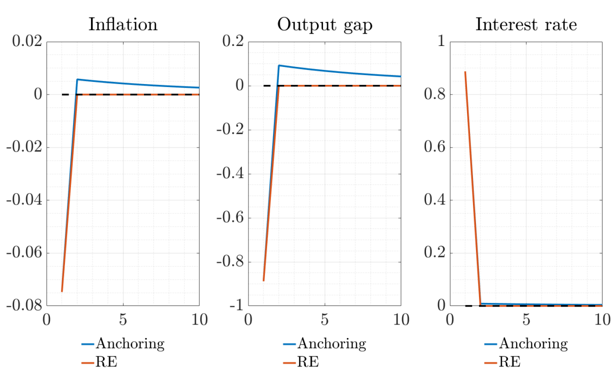
(b) Mean gain



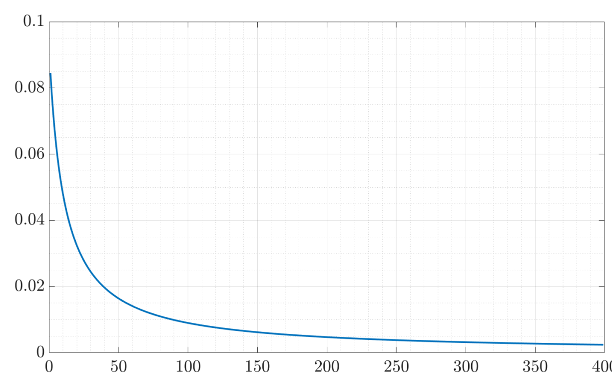
(c) Constant gain learning



(d) Mean gain

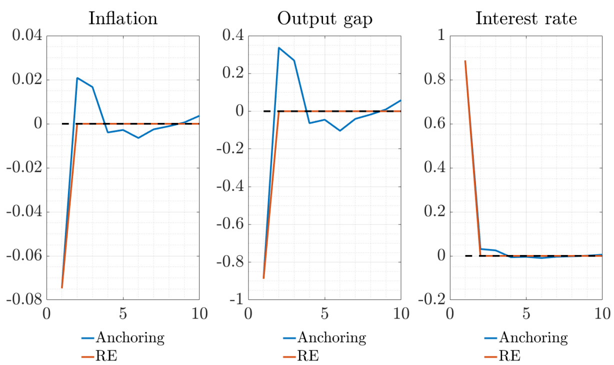


(e) CEMP criterion (vector)

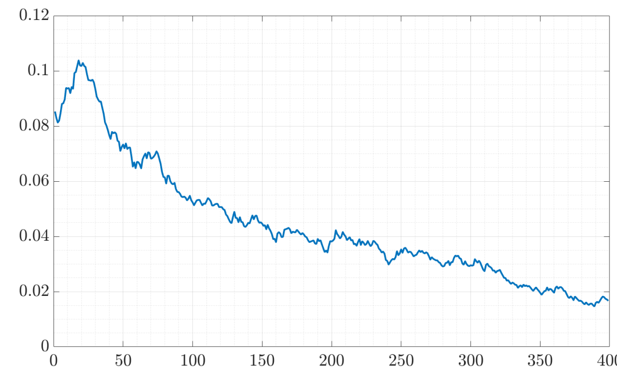


(f) Mean gain

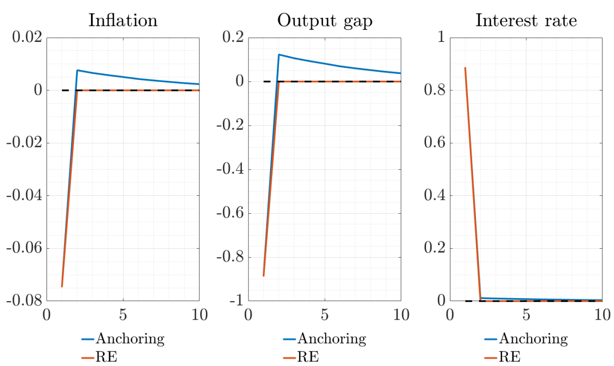
Figure 18: IRFs and gain history (sample means), continued



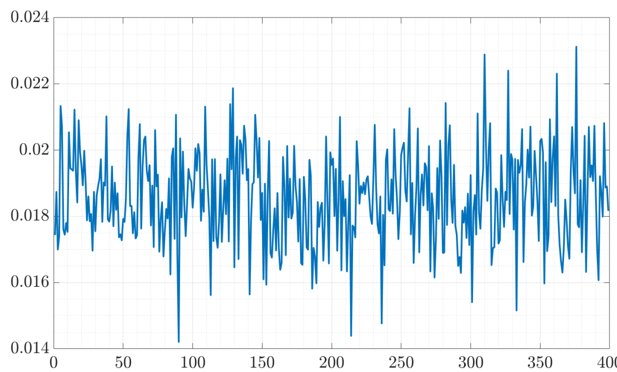
(a) CUSUM criterion (vector)



(b) Mean gain



(c) Smooth criterion, approximated, using $\alpha^{true} = (0.05; 0.025; 0; 0.025; 0.05)$, on $fe \in (-2, 2)$.



(d) Mean gain

A Model summary

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

$$\pi_t = \kappa x_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t}\big(\kappa\alpha\beta x_{T+1} + (1-\alpha)\beta\pi_{T+1} + u_T\big) \tag{A.2}$$

$$i_t = \psi_\pi \pi_t + \psi_x x_t + \bar{i}_t \qquad \text{(if imposed)} \tag{A.3}$$

$$\text{PLM:} \qquad \hat{\mathbb{E}}_t z_{t+h} = a_{t-1} + b h_x^{h-1} s_t \quad \forall h \geq 1 \qquad b = g_x\, h_x \tag{A.4}$$

$$\text{Updating:} \qquad a_t = a_{t-1} + k_t^{-1}\big(z_t - (a_{t-1} + b s_{t-1})\big) \tag{A.5}$$

$$\text{Anchoring function:} \qquad k_t^{-1} = \rho_k k_{t-1}^{-1} + \gamma_k f e_{t-1}^2 \tag{A.6}$$

$$\text{Forecast error:} \qquad f e_{t-1} = z_t - (a_{t-1} + b s_{t-1}) \tag{A.7}$$

$$\text{LH expectations:} \qquad f_a(t) = \frac{1}{1-\alpha\beta} a_{t-1} + b(\mathbb{I}_{nx} - \alpha\beta h)^{-1} s_t \qquad f_b(t) = \frac{1}{1-\beta} a_{t-1} + b(\mathbb{I}_{nx} - \beta h)^{-1} s_t \tag{A.8}$$

This notation captures vector learning (z learned) for intercept only. For scalar learning, $a_t = \begin{pmatrix}\bar{\pi}_t & 0 & 0\end{pmatrix}'$ and b_1 designates the first row of b. The observables (π, x) are determined as:

$$x_t = -\sigma i_t + \begin{bmatrix}\sigma & 1-\beta & -\sigma\beta\end{bmatrix} f_b + \sigma \begin{bmatrix}1 & 0 & 0\end{bmatrix} (\mathbb{I}_{nx} - \beta h_x)^{-1} s_t \tag{A.9}$$

$$\pi_t = \kappa x_t + \begin{bmatrix}(1-\alpha)\beta & \kappa\alpha\beta & 0\end{bmatrix} f_a + \begin{bmatrix}0 & 0 & 1\end{bmatrix} (\mathbb{I}_{nx} - \alpha\beta h_x)^{-1} s_t \tag{A.10}$$

B Target criterion

The target criterion in the simplified model (scalar learning of inflation intercept only, $k_t^{-1} = \mathbf{g}(f e_{t-1})$):

$$\pi_t = -\frac{\lambda_x}{\kappa}\Bigg\{ x_t - \frac{(1-\alpha)\beta}{1-\alpha\beta}\Big(k_t^{-1} + ((\pi_t - \bar{\pi}_{t-1} - b_1 s_{t-1}))\mathbf{g}_\pi(t)\Big)$$
$$\Bigg(\mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i} \prod_{j=0}^{i-1} (1 - k_{t+1+j}^{-1} - (\pi_{t+1+j} - \bar{\pi}_{t+j} - b_1 s_{t+j})\mathbf{g}_{\bar{\pi}}(t+j))\Bigg)\Bigg\} \tag{B.1}$$

where I'm using the notation that $\prod_{j=0}^{0} \equiv 1$. For interpretation purposes, let me rewrite this as follows:

$$\pi_t = -\frac{\lambda_x}{\kappa} x_t + \frac{\lambda_x}{\kappa}\frac{(1-\alpha)\beta}{1-\alpha\beta}\Big(k_t^{-1} + f e_{t|t-1}^{eve}\mathbf{g}_\pi(t)\Big)\mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i}$$
$$-\frac{\lambda_x}{\kappa}\frac{(1-\alpha)\beta}{1-\alpha\beta}\Big(k_t^{-1} + f e_{t|t-1}^{eve}\mathbf{g}_\pi(t)\Big)\Big(\mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i} \prod_{j=0}^{i-1} (k_{t+1+j}^{-1} + f e_{t+1+j|t+j}^{eve})\mathbf{g}_{\bar{\pi}}(t+j)\Big) \tag{B.2}$$

Interpretation: tradeoffs from discretion in RE + effect of current level and change of the gain on future tradeoffs + effect of future expected levels and changes of the gain on future tradeoffs