Materials 40 - Still trying to understand why not identified

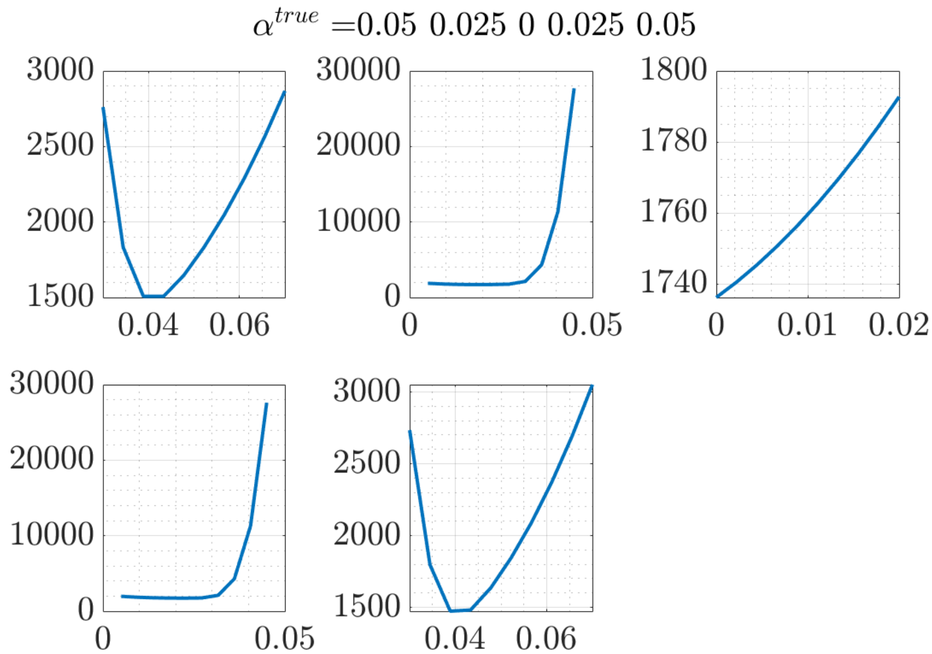
Laura Gáti

August 12, 2020

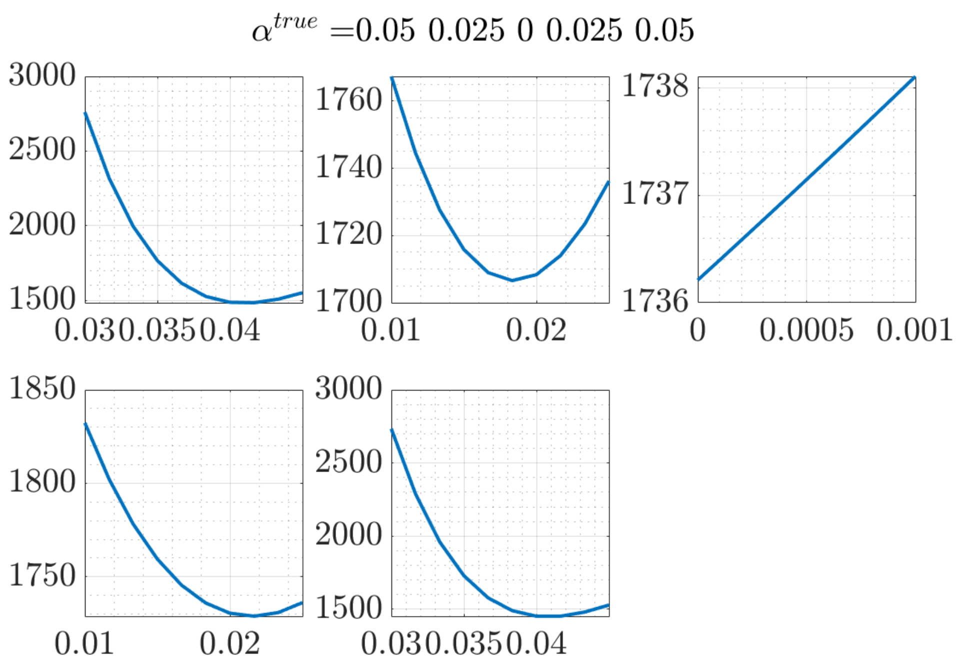
Overview

1 Loss when varying one parameter, more details

Figure 1: Loss for $N = 100$, NOT using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$



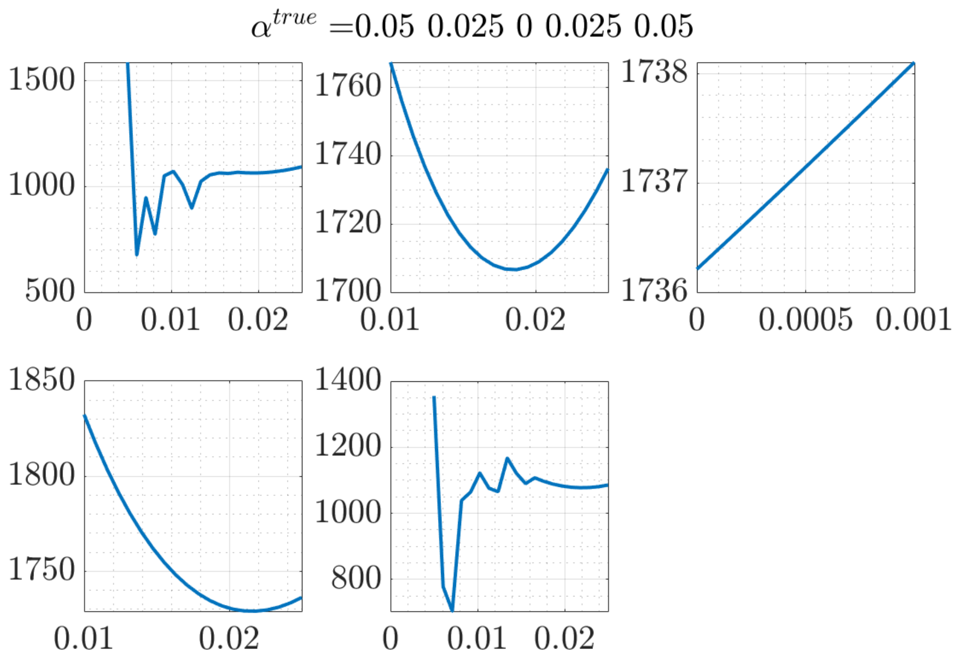
(a) Reference figure from Materials 39: increment around true value 0.02



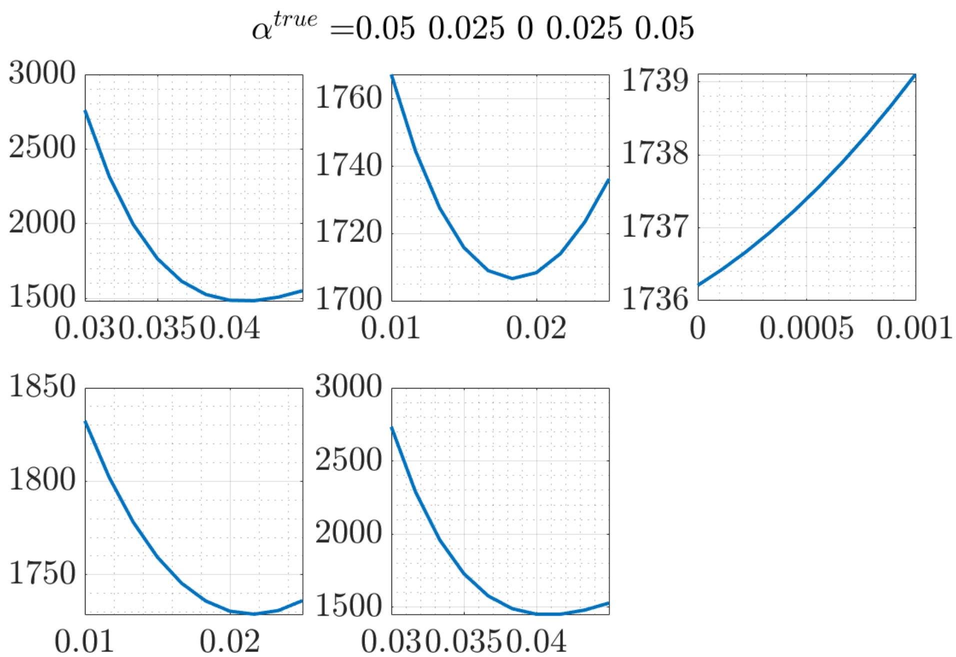
(b) Minima: $(0.041; 0.0187; 0; 0.0218; 0.041)$

1.1 Convexity and 0 at 0 restrictions

Figure 2: Variations I



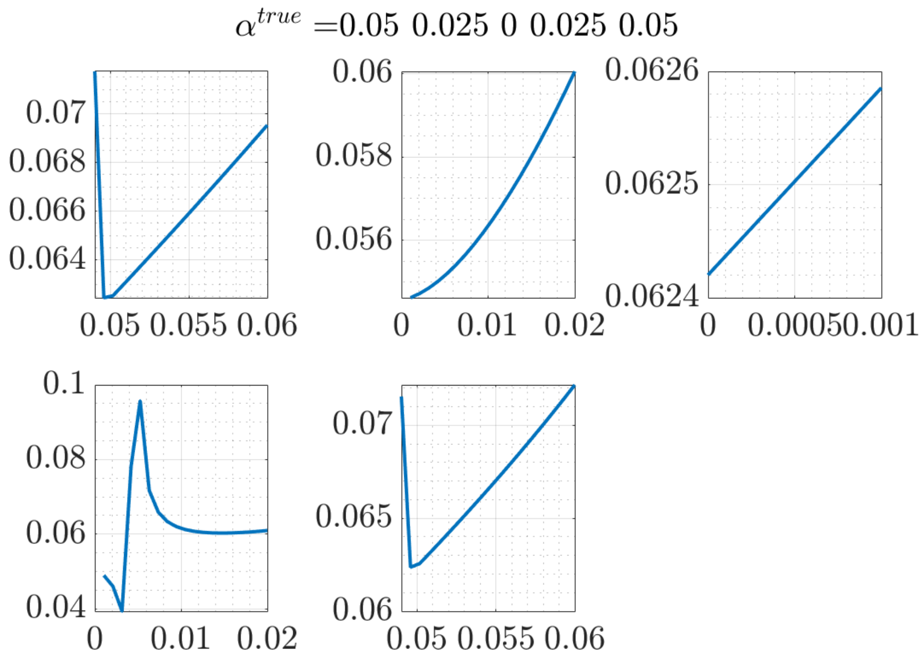
(a) Shutting of the convexity moment (weight 0)



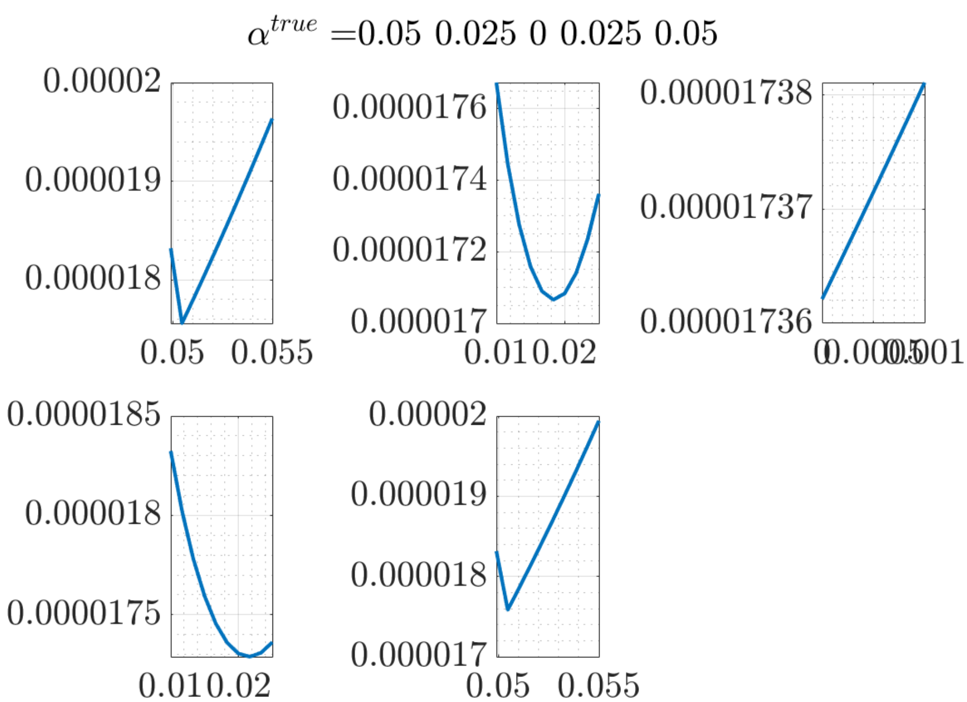
(b) Adding the 0 at 0 restriction (weight 1000)

1.2 Measurement error and rescale W

Figure 3: Variations II



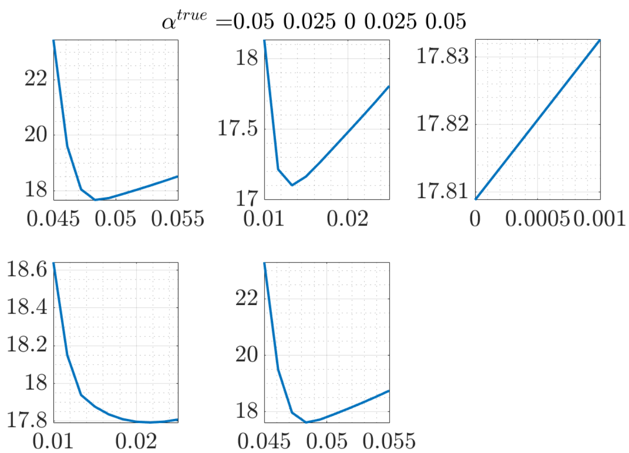
(a) Add measurement error



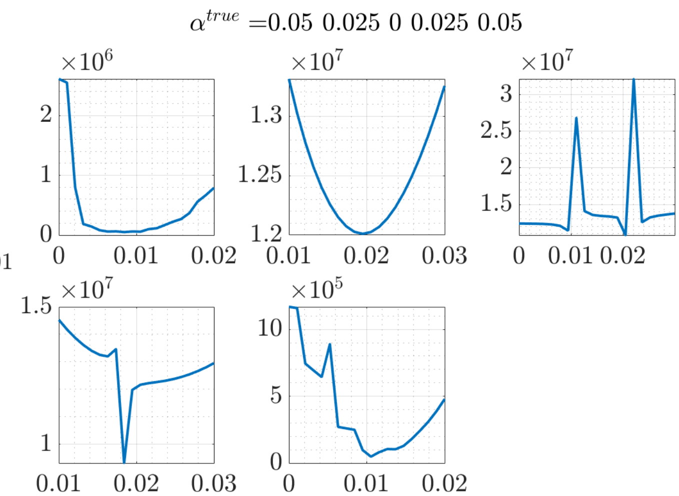
(b) Rescale W

1.3 Add expectations w/ and w/o rescaling W

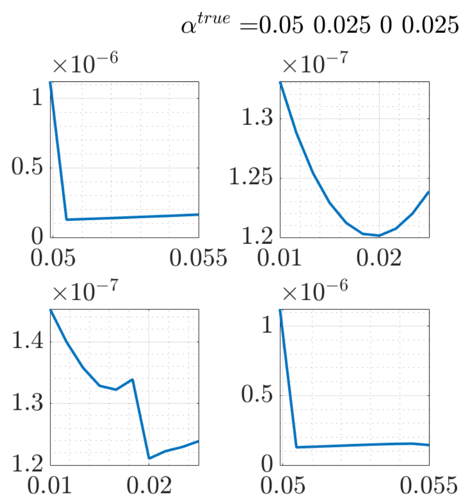
Figure 4: Variations III



(a) A shorter dataset w/o expectations



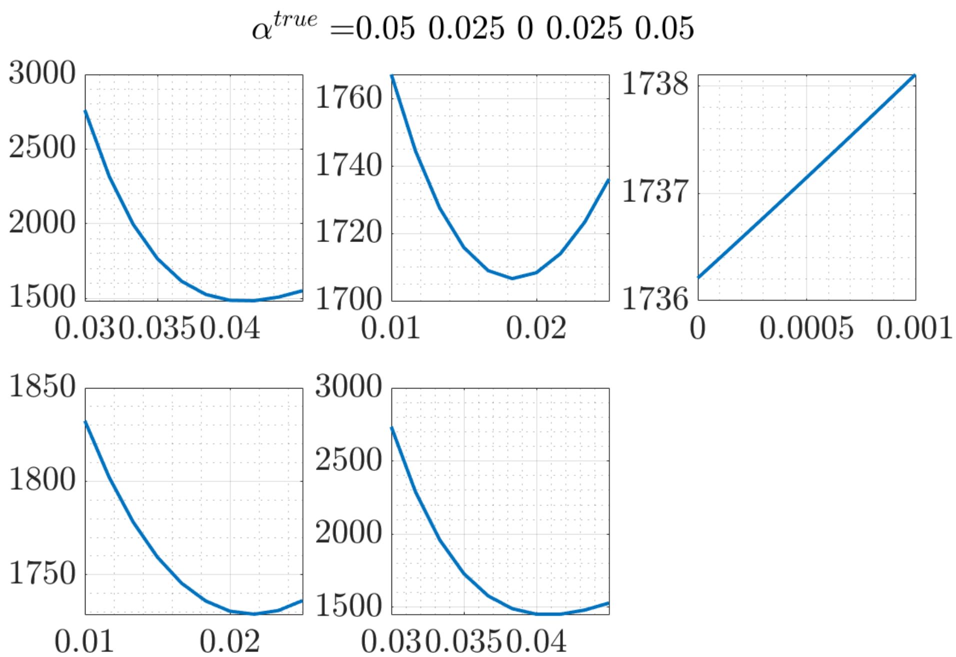
(b) A shorter dataset adding expectations



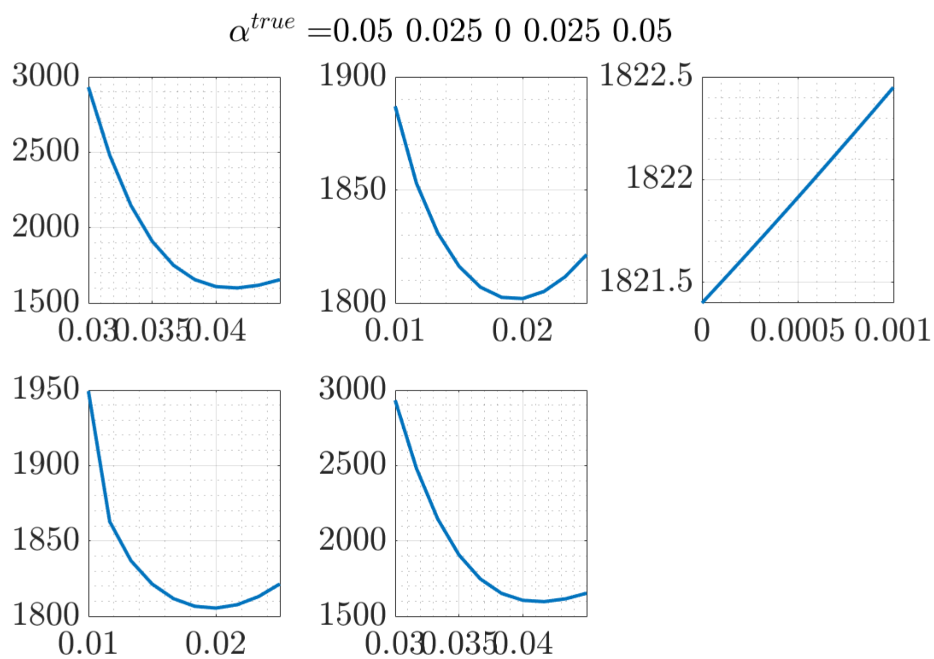
(c) A shorter dataset adding expectations, rescaling W

1.4 Loss w/o expectations but $N = 1000$

Figure 5: Variations IV



(a) Reproduces Fig 1, Panel b



(b) N=1000

2 Take a deep breath: what have I learned?

1. Some indication that the measurement error is screwed up, but I can bypass it, so ignore for now.

2. Rescaling might exit too soon. Main problem is it shouldn't change the *shape* of the loss function, but does. Yet no indication of numerical matrix inversion problems. I don't understand.

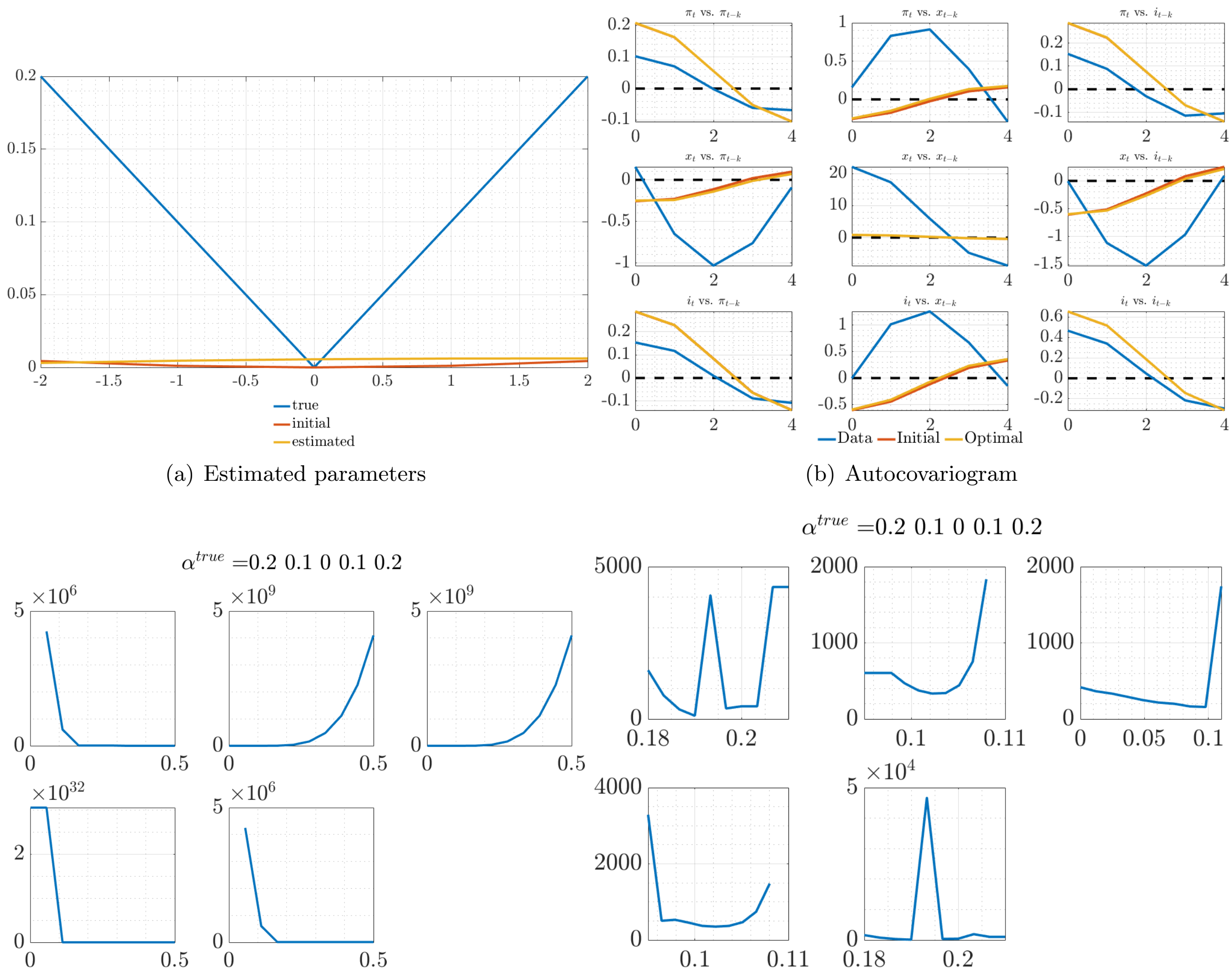
3. Indication that something is screwed up with the expectations, potentially connected with the rescaling. Ridge didn't really help either.

4. Loss function indicates that the parameters *are* identified. However, since loss is greater at true values than at estimated ones, it seems that the truth is a local, not a global min. I need to i) use some tricks to find this min ii) understand why this min isn't the global. I have a hypothesis:

   I think expectations in the true data aren't very large, and thus also aren't fluctuating enough. This screws up the moments somehow, but it also means that the estimation wants to set the αs corresponding to large forecast errors to a low value, b/c otherwise it would cause fluctuations that aren't there in the data. Combined with the zero-neighborhood problem, the flat estimate is the result.

3 Truth with more action in expectations

Figure 6: Not using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$ true αs scaled up by 4



(a) Estimated parameters (b) Autocovariogram

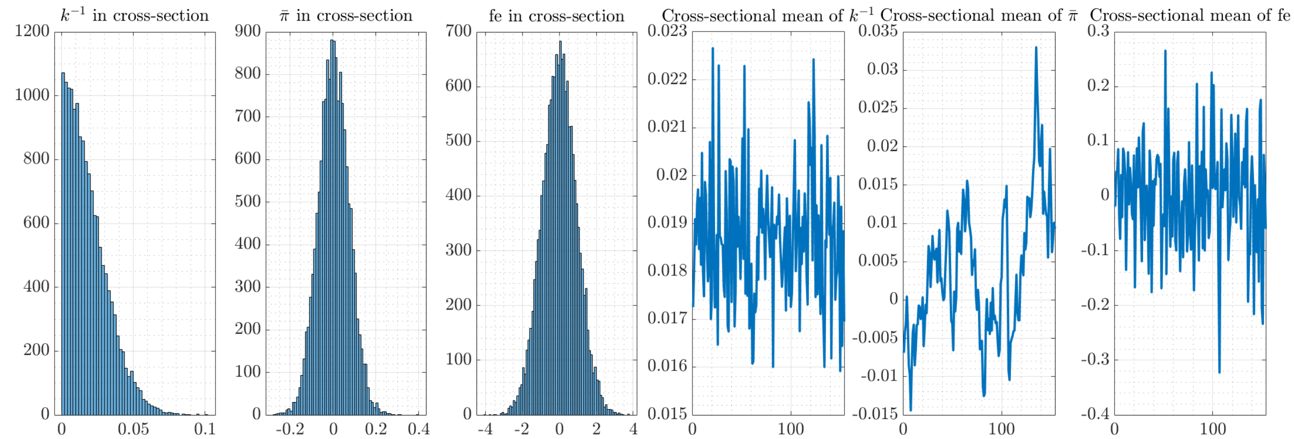
(c) Loss (d) Loss, zooming in

For every evaluation of the loss function, very many simulations explode.

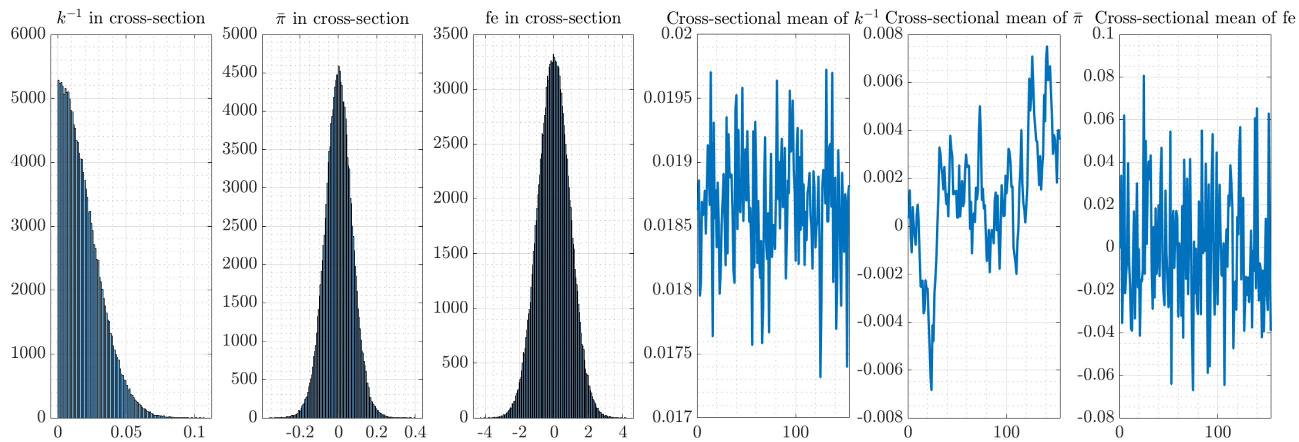
4 Look into behavior of simulation as a function of shocks and αs

4.1 Cross-section for α^{true}

Figure 7: Summary statistics of simulation in a cross-section, $\alpha = (0.05, 0.025, 0, 0.025, 0.05)$, $nfe = 5, fe \in (-2, 2)$, `rng(1)`



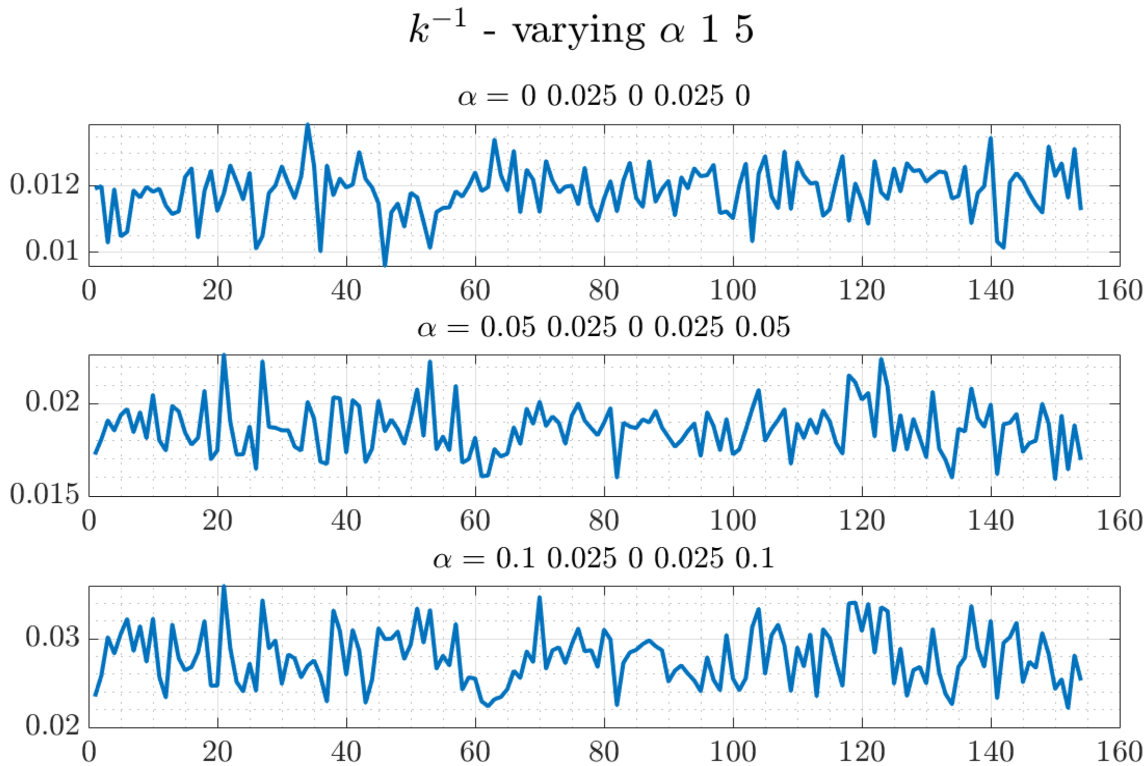
(a) Distributions, $N = 100$ (b) Means, $N = 100$



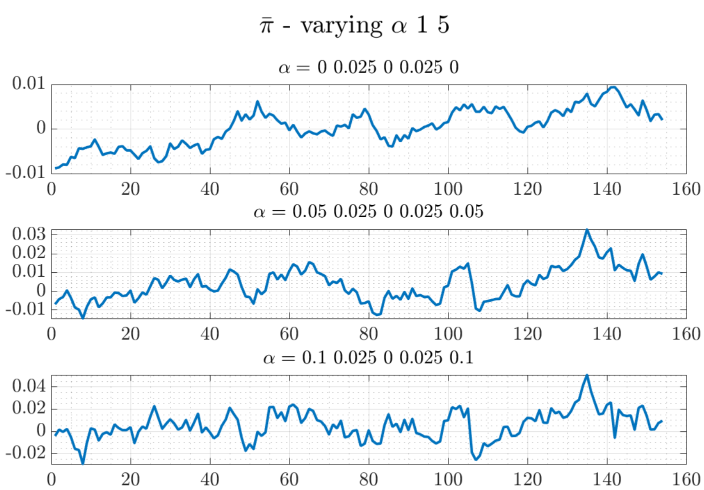
(c) Distributions, $N = 1000$ (d) Means, $N = 1000$

4.2 Cross-sectional simulations moving particular elements of α at a time

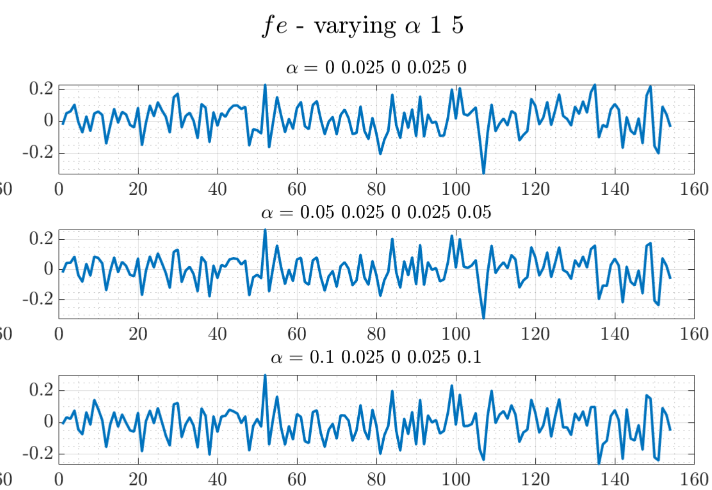
Figure 8: Cross-sectional means when varying $\alpha_{1,5}$, $N = 100$, $\alpha = (0.05, 0.025, 0, 0.025, 0.05)$, $nfe = 5$, $fe \in (-2, 2)$, `rng(1)`



(a)

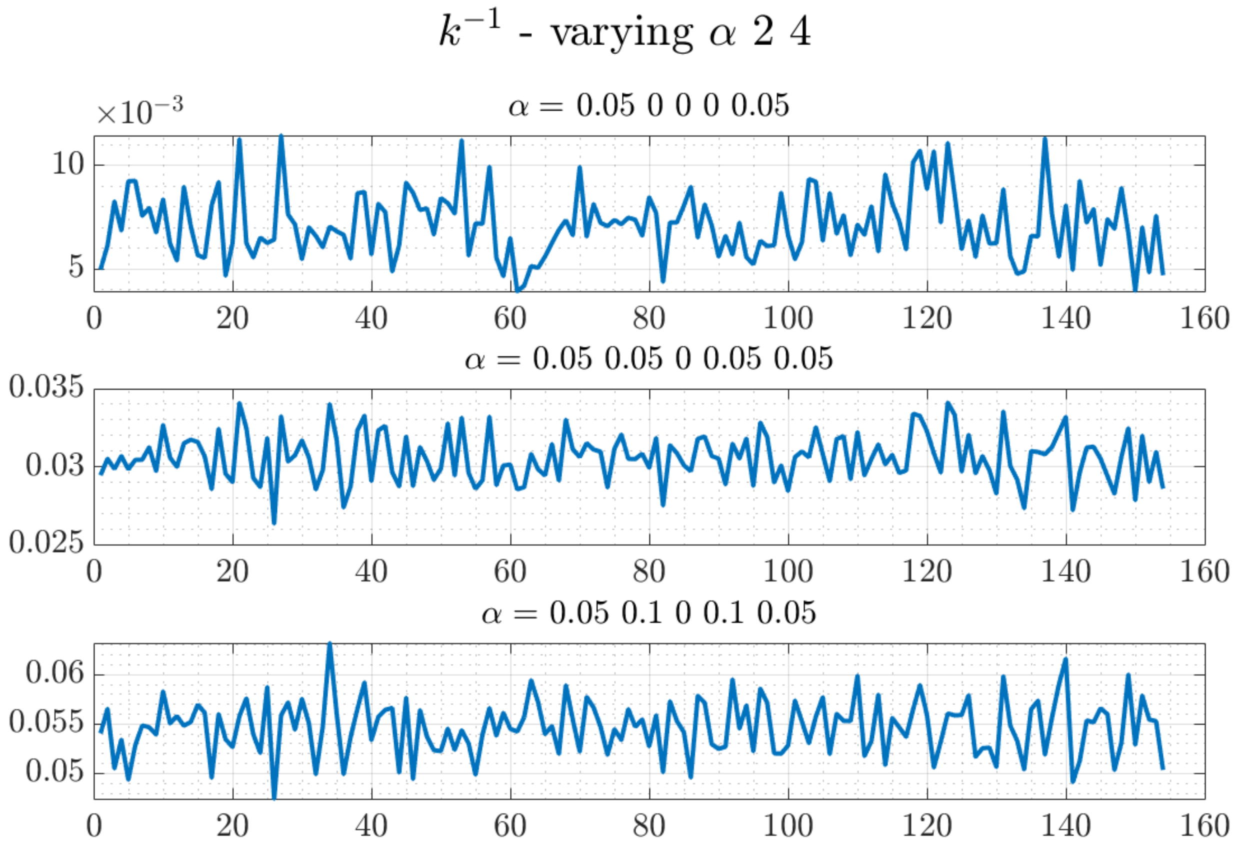


(b)

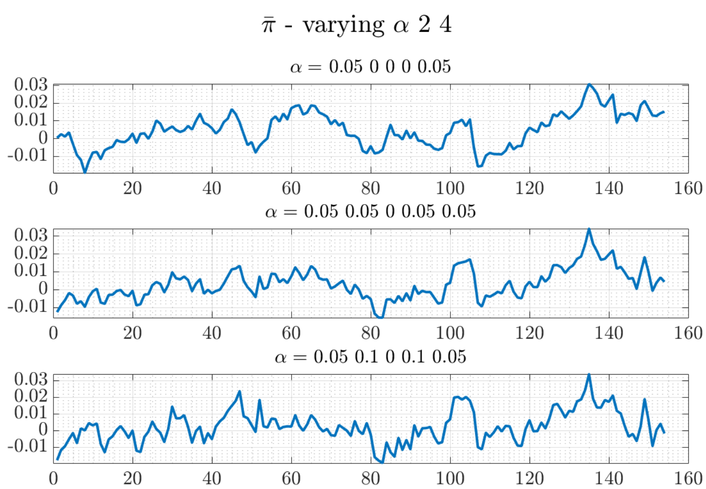


(c)

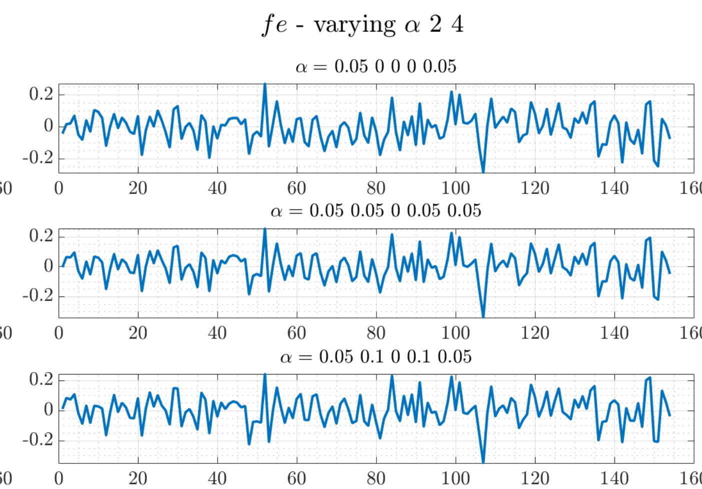
Figure 9: Cross-sectional means when varying $\alpha_{2,4}$, $N = 100$, $\alpha = (0.05, 0.025, 0, 0.025, 0.05)$, $nfe = 5, fe \in (-2, 2)$, `rng(1)`



(a)

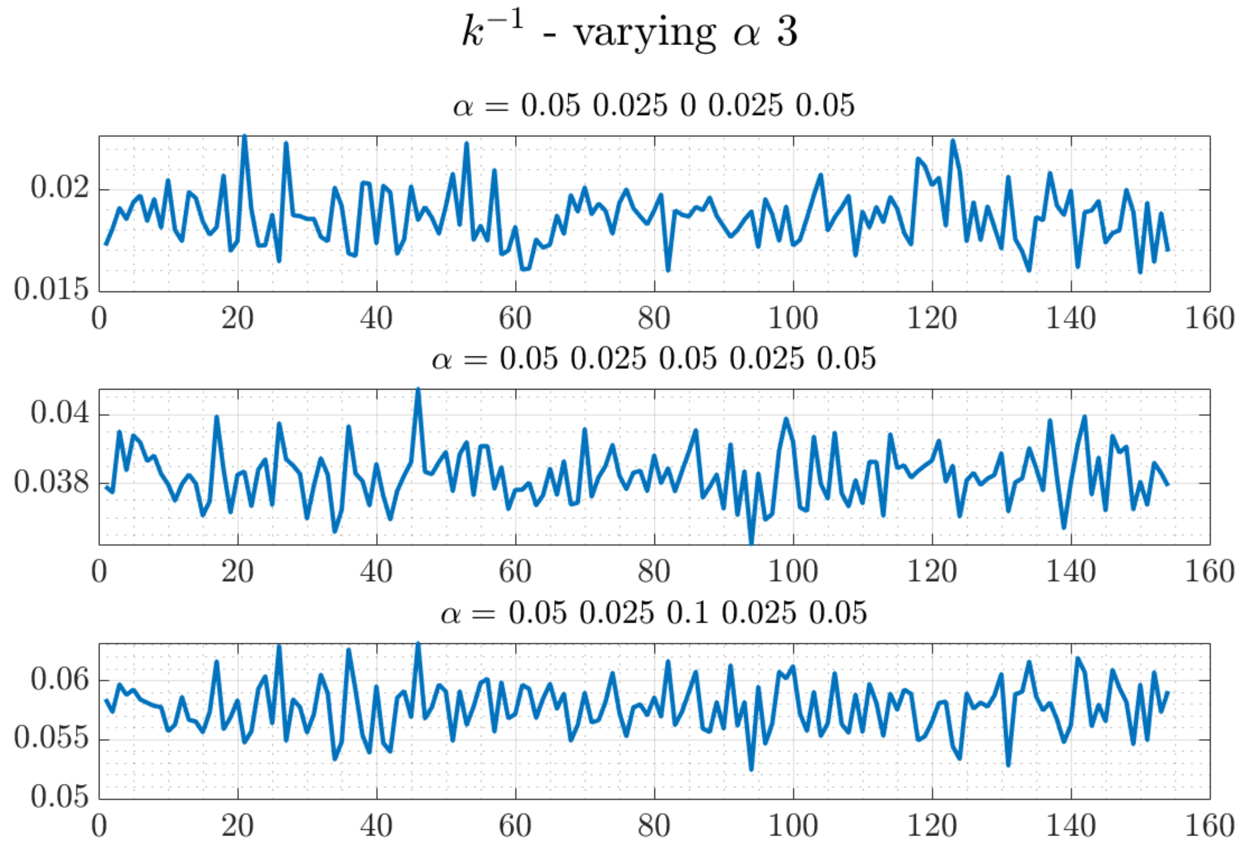


(b)

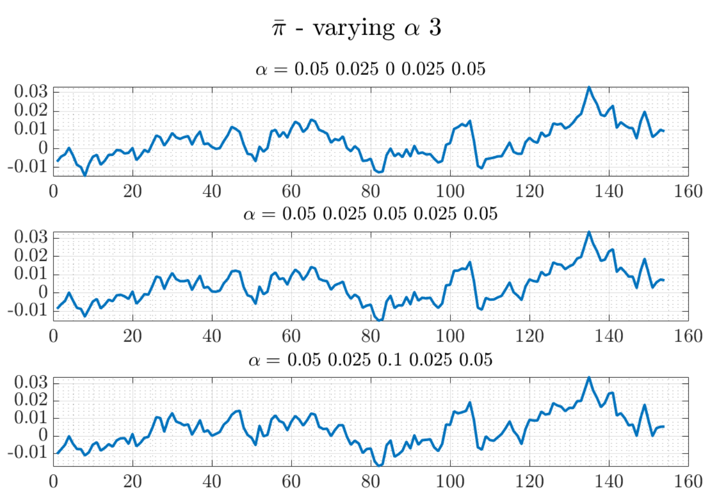


(c)

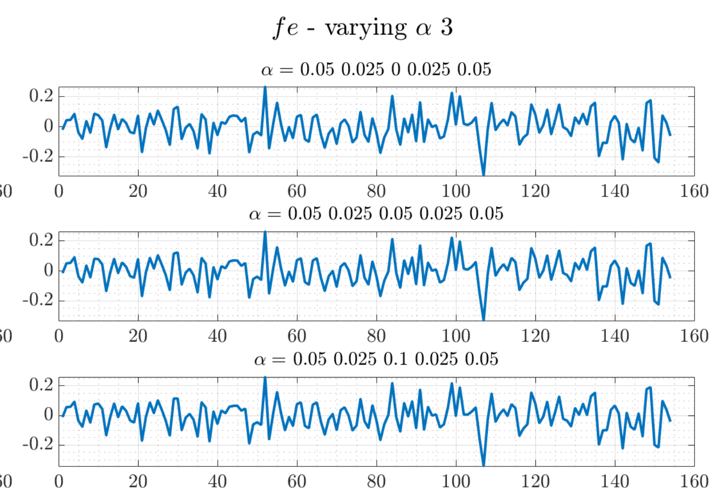
Figure 10: Cross-sectional means when varying α_3, $N = 100$, $\alpha = (0.05, 0.025, 0, 0.025, 0.05)$, $nfe = 5$, $fe \in (-2, 2)$, `rng(1)`



(a)



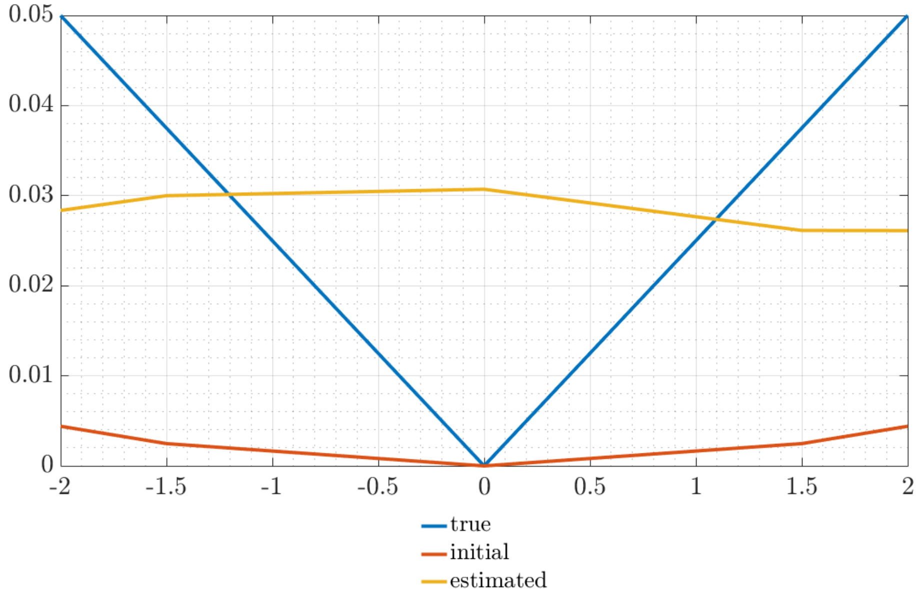
(b)



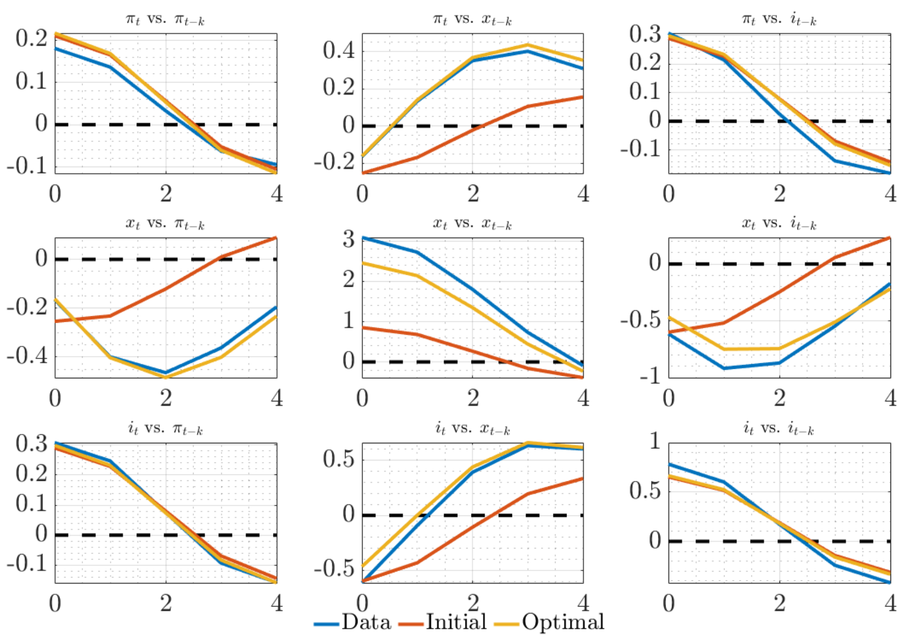
(c)

5 Only α corresponding to large forecast errors

Figure 11: Not using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$, gridpoints $= [-2, -1.5, 0, 1.5, 2]$

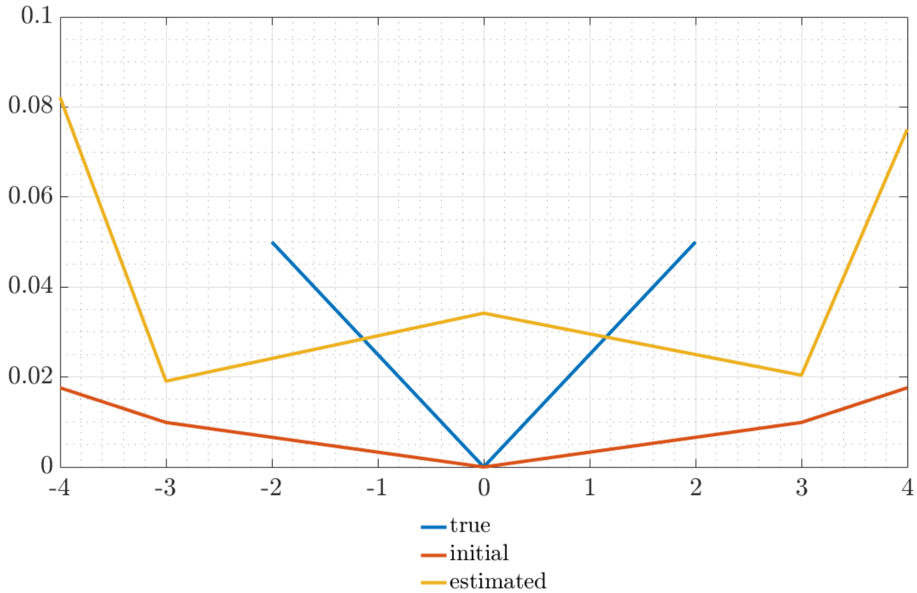


(a) Estimated parameters

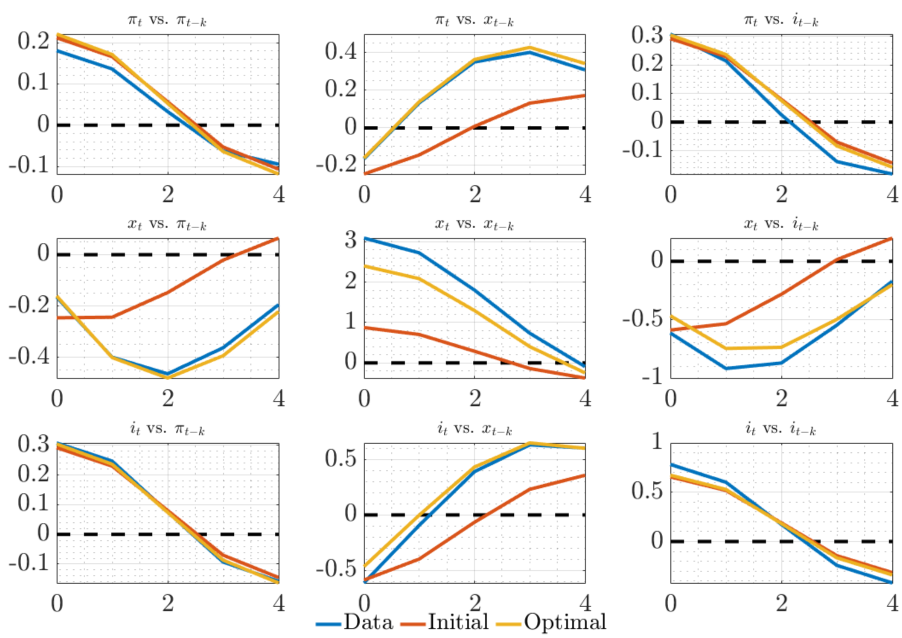


(b) Autocovariogram

Figure 12: Not using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$, gridpoints $= [-4, -3, 0, 3, 4]$

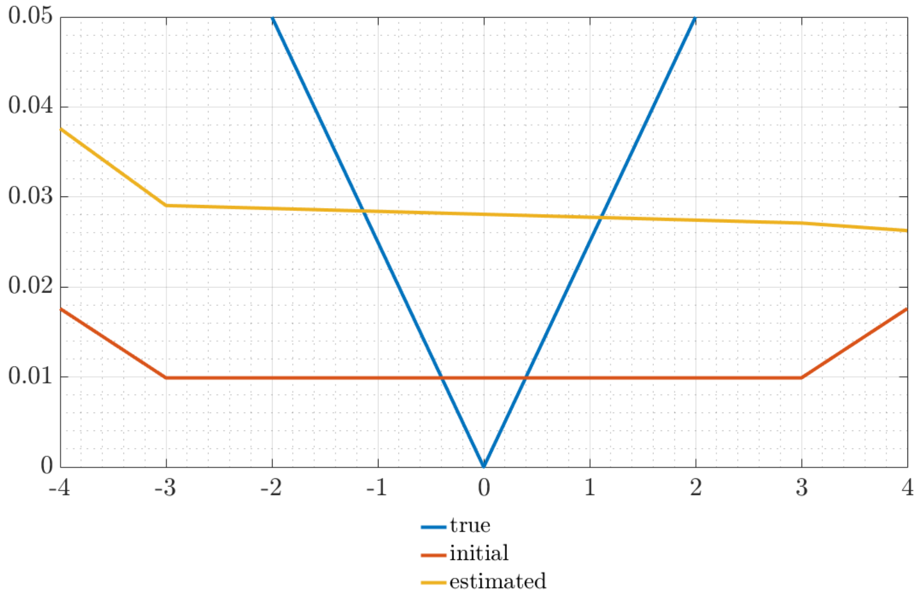


(a) Estimated parameters

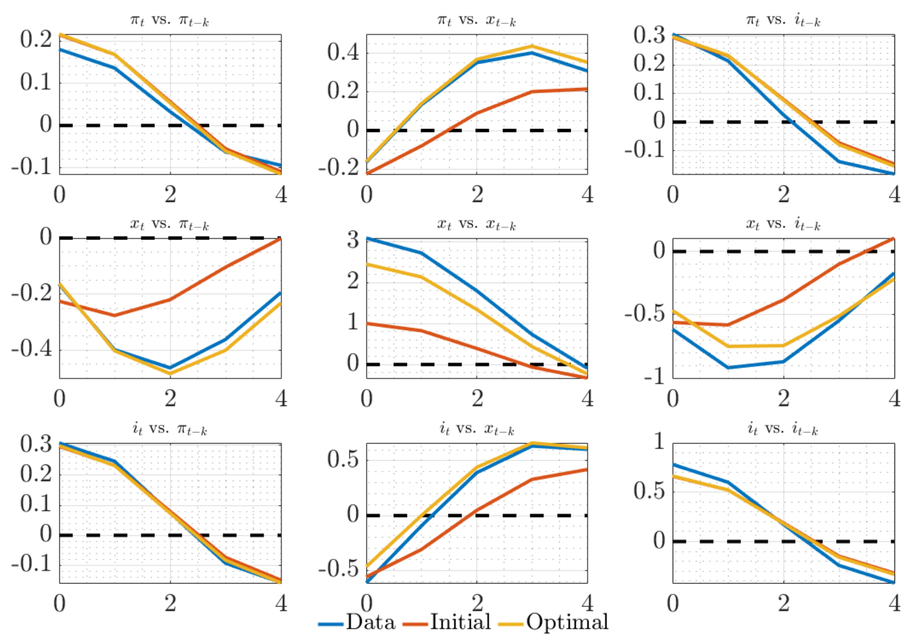


(b) Autocovariogram

Figure 13: Not using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$, gridpoints $= [-4, -3, 3, 4]$ (no zero-point)

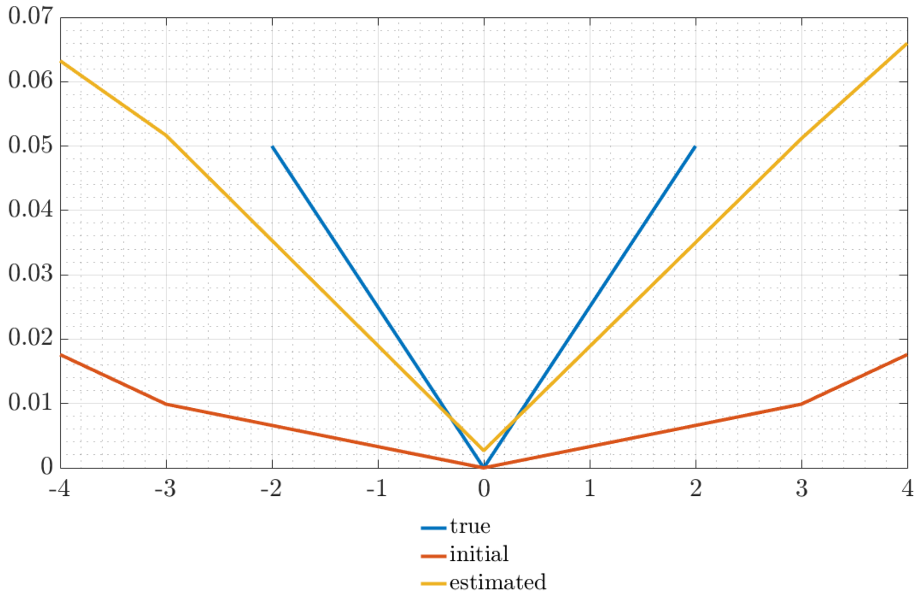


(a) Estimated parameters

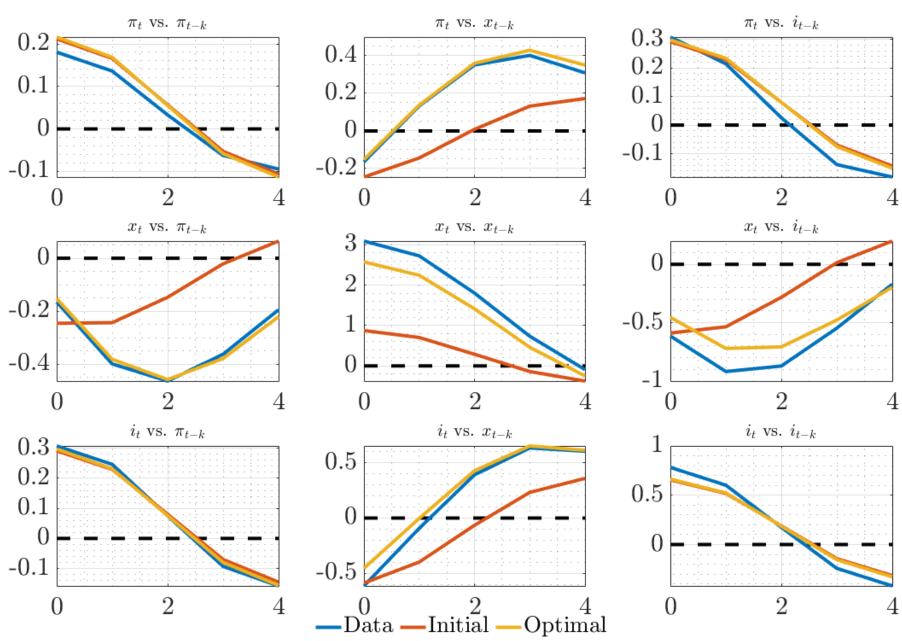


(b) Autocovariogram

Figure 14: Not using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$, gridpoints $= [-4, -3, 0, 3, 4]$ with 0 at 0 imposed with weight 1000

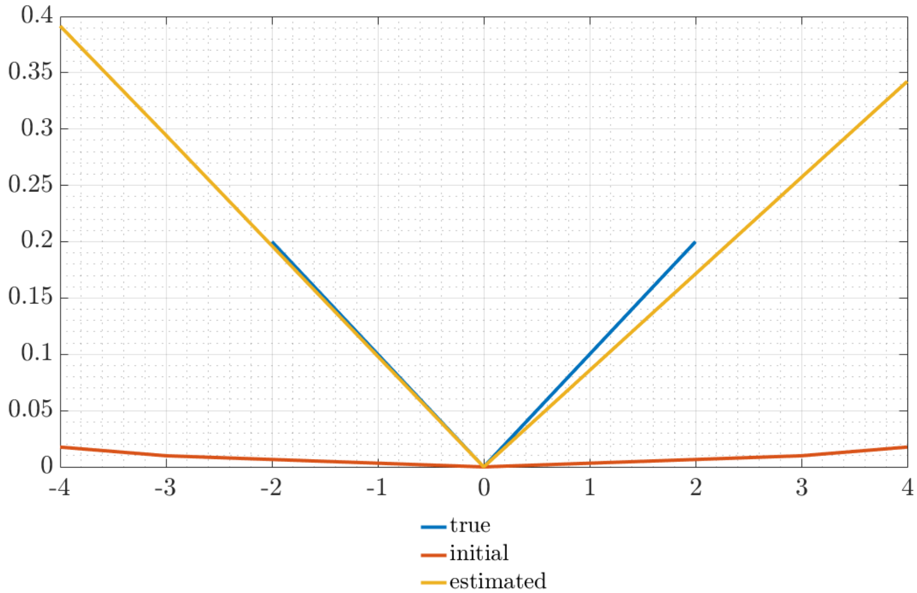


(a) Estimated parameters

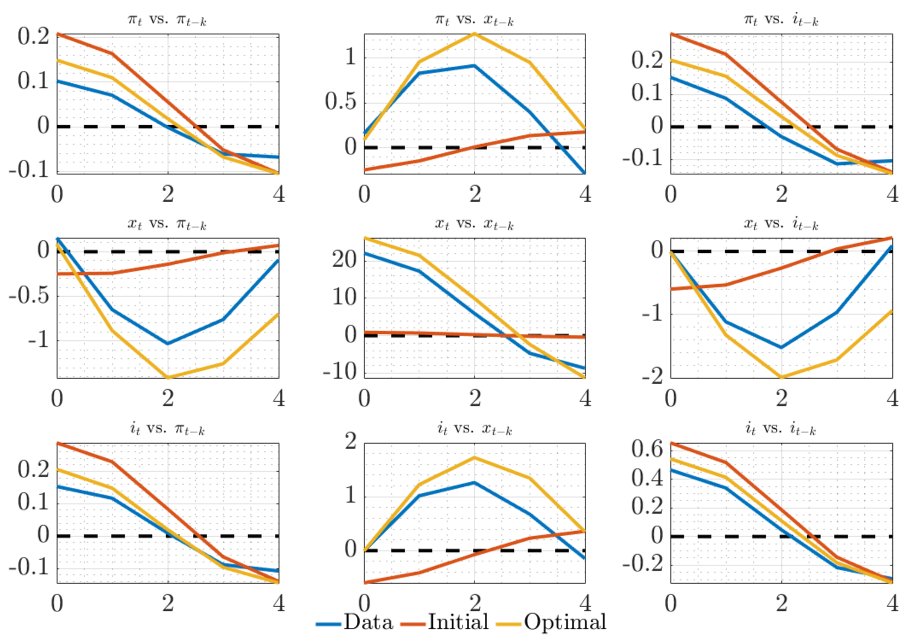


(b) Autocovariogram

Figure 15: Not using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$, gridpoints = $[-4, -3, 0, 3, 4]$ with 0 at 0 imposed with weight 1000, true parameters scaled up by 4



(a) Estimated parameters



(b) Autocovariogram

6 Notes on the GMM weighting matrix

The loss function takes the form:

$$L = (\Omega^{data} - \Omega^{model})W(\Omega^{data} - \Omega^{model})' \tag{1}$$

where Ω are the moments and W is the weighting matrix. Let $\Sigma \equiv Var(\Omega^{data,bootstrap})$. Then $W = \Sigma^{-1}$.

For `lsqnonlin`, I write the residuals of the objective function as

$$R = (vec(\Omega^{data}) - vec(\Omega^{model}))diag(W) \tag{2}$$

I implement additional restrictions by adding elements to the residuals vector R with a manually specified weight. For example, the convexity restriction is added as follows:

- For each guess α, compute g_i'', numerical second derivatives of the anchoring function. There are always `numel(alpha)-2` of those, so if I estimate 5 knots, there are 3 second derivatives, $i = 1, 2, 3$.

- For $i = 1, 2, 3$, add the following elements to the residuals vector R:

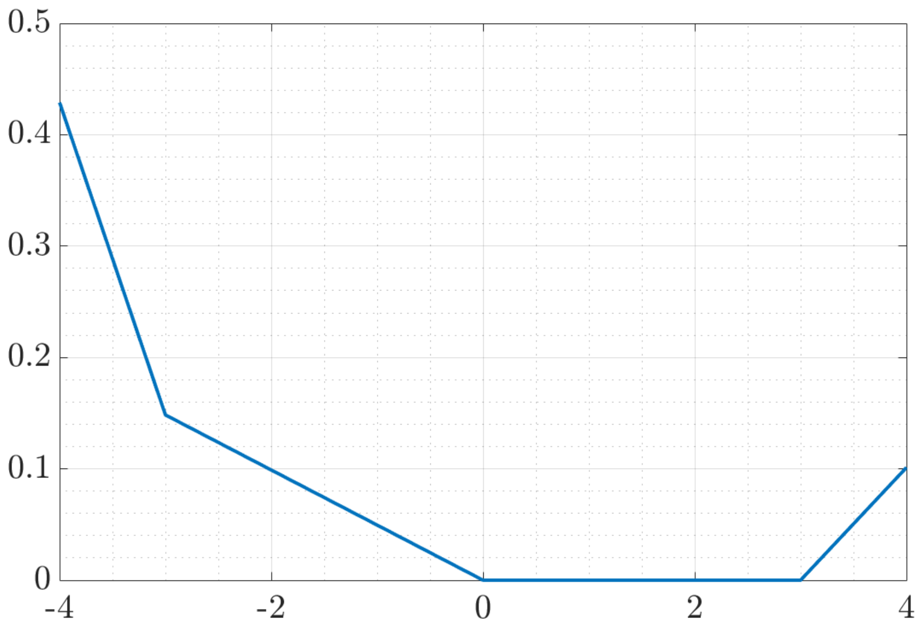
$$\begin{cases} (g_i'')^2 W^{convexity} & \text{if} \quad g_i'' < 0 \\ 0 \quad \text{otherwise} \end{cases} \tag{3}$$

  where $W^{convexity}$ is a manually specified weight.

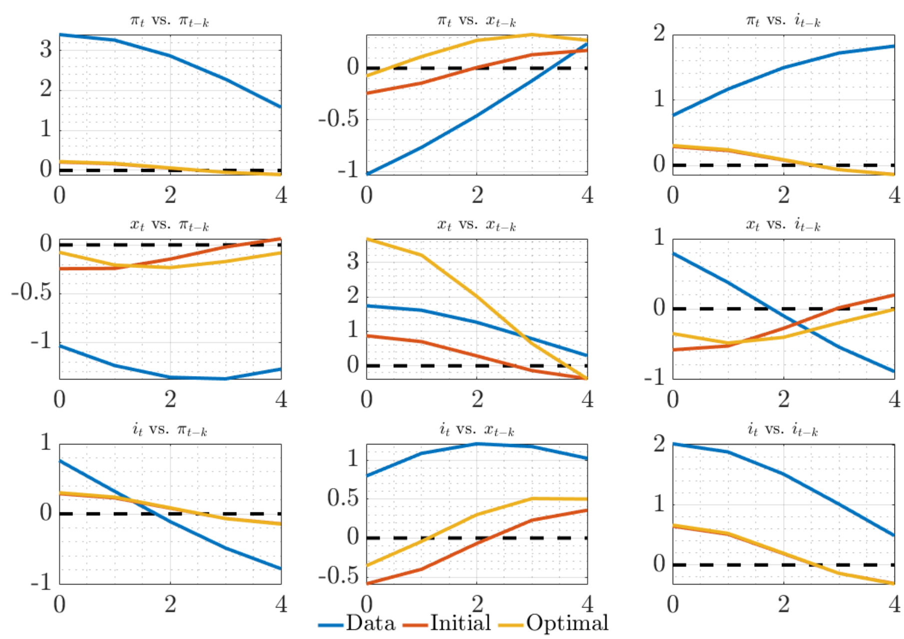
7 Some additional estimation exercises for the N simulations strategy with settings preceding those of the previous section

7.1 Real data with αs out in the edges, as above

Figure 16: Not using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, real data, gridpoints $= [-4, -3, 0, 3, 4]$ with 0 at 0 imposed with weight 1000, $N = 100$



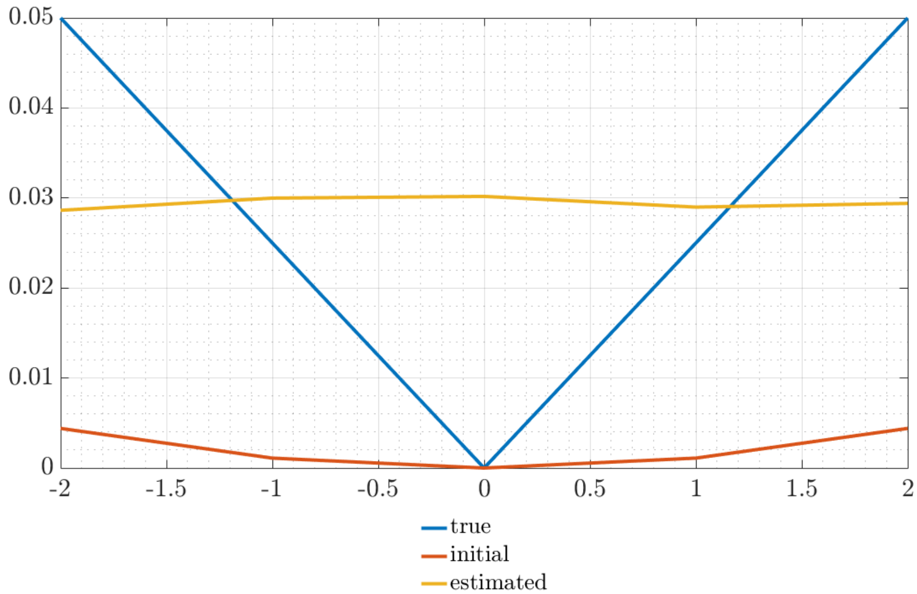
(a) Estimated parameters



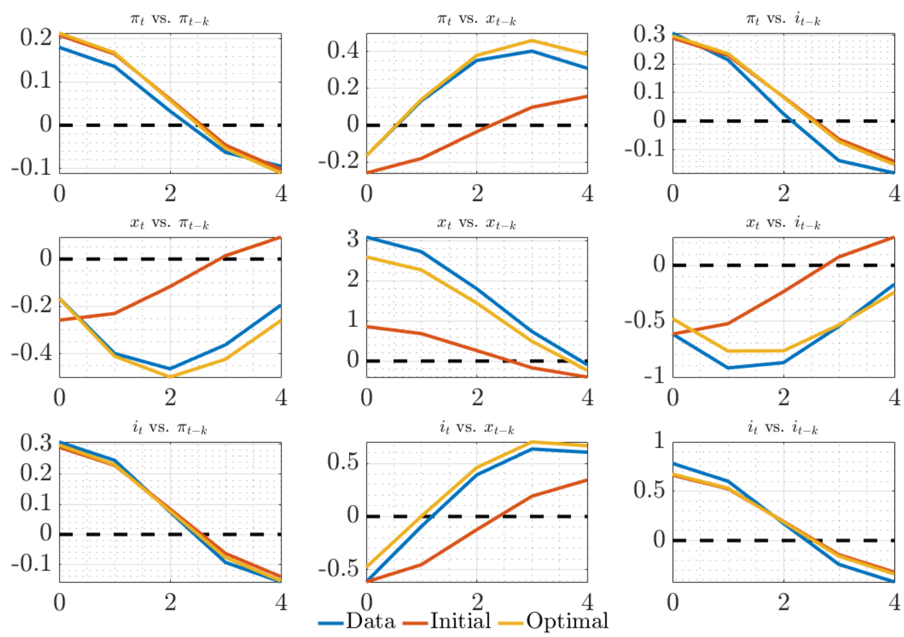
(b) Autocovariogram

7.2 $N = 1000$

Figure 17: NOT using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$



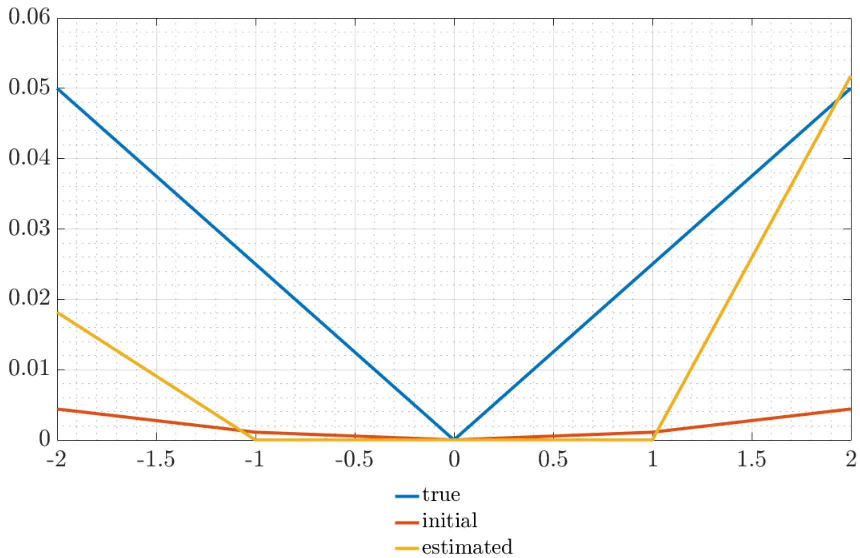
(a) Estimated parameters



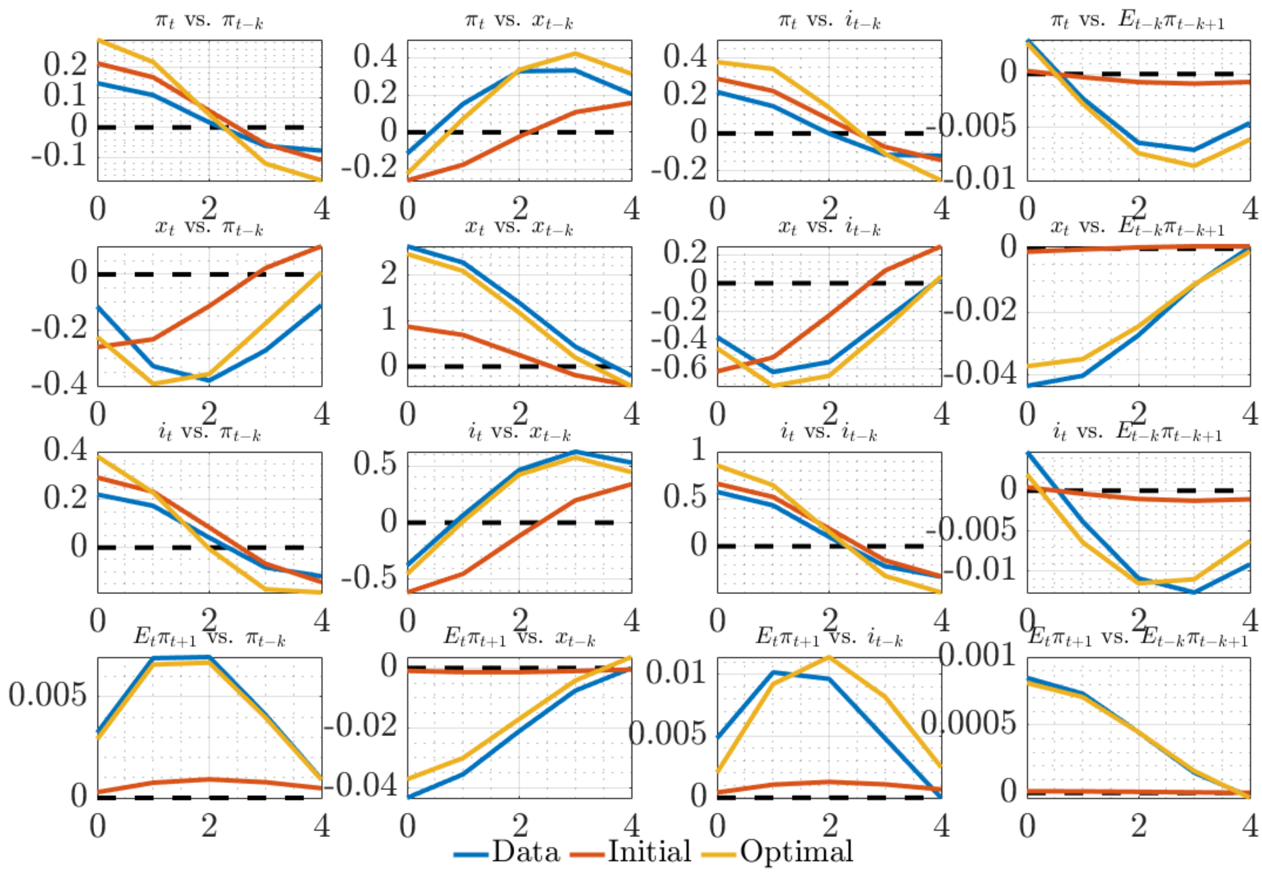
(b) Autocovariogram

7.3 Expectations

Figure 18: using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$



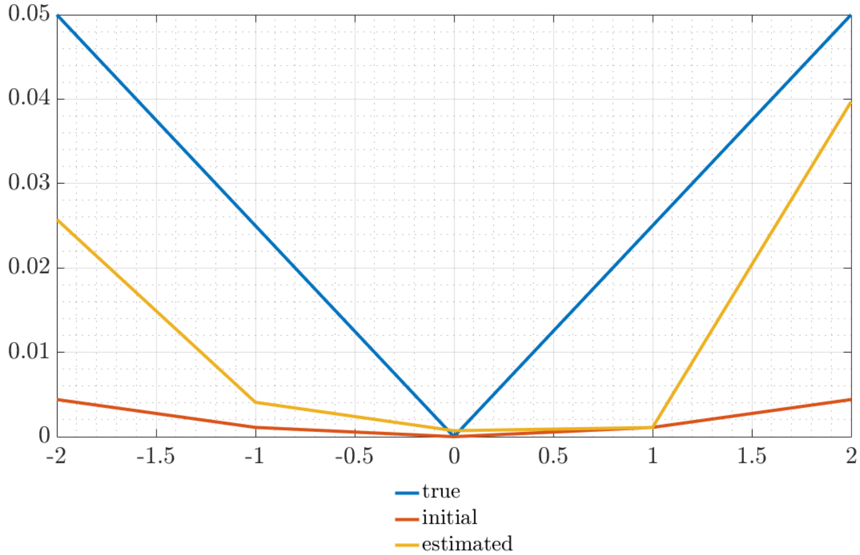
(a) Estimated parameters



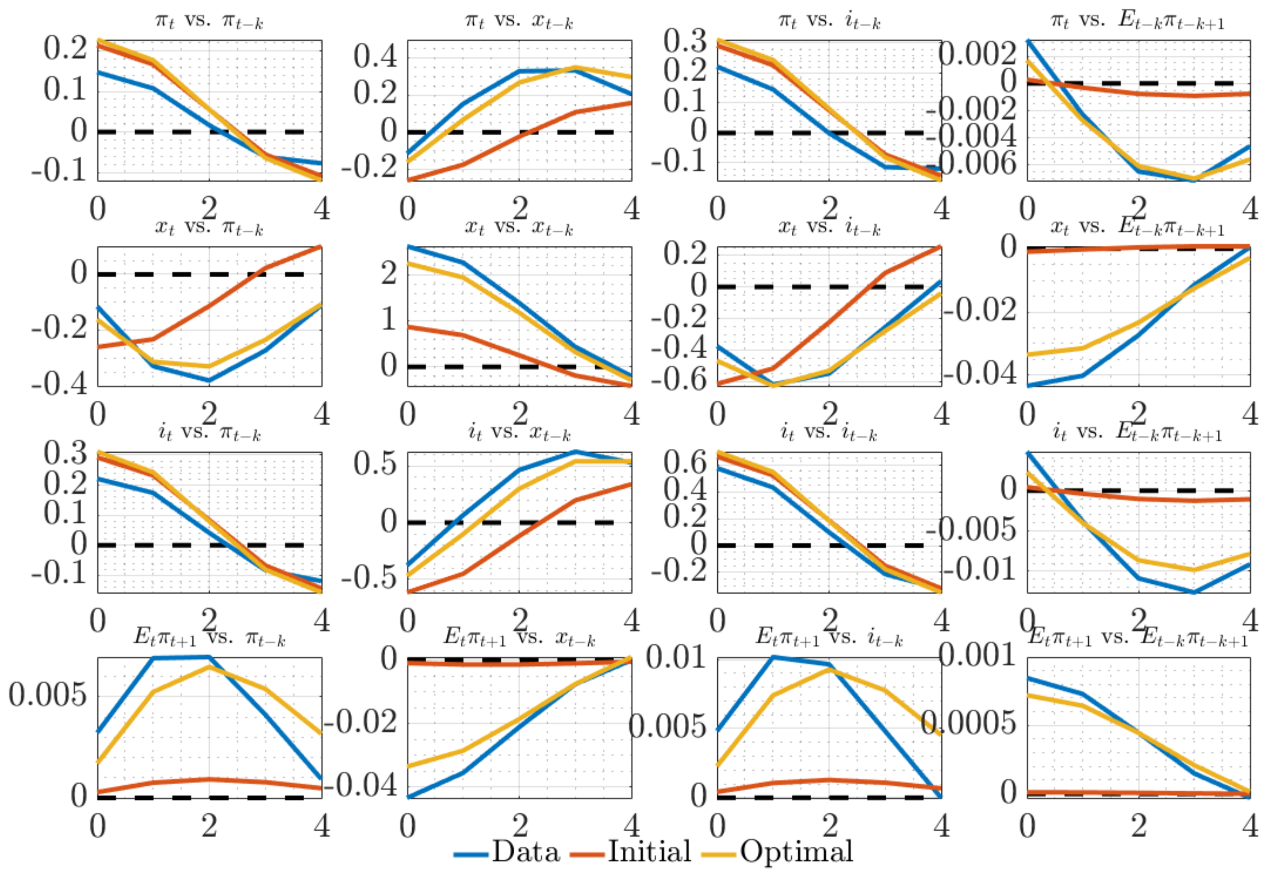
(b) Autocovariogram

7.4 Expectations and rescaling

Figure 19: using 1-step ahead forecasts of inflation, rescaling W, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$



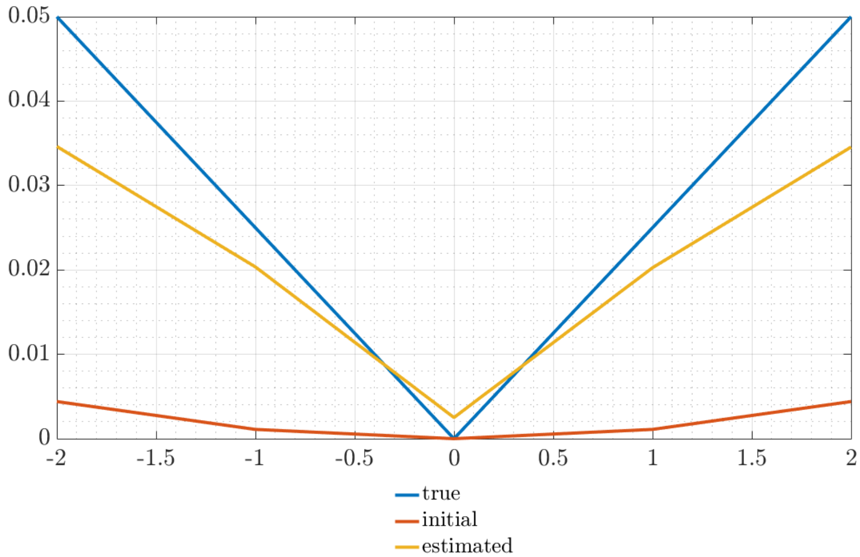
(a) Estimated parameters



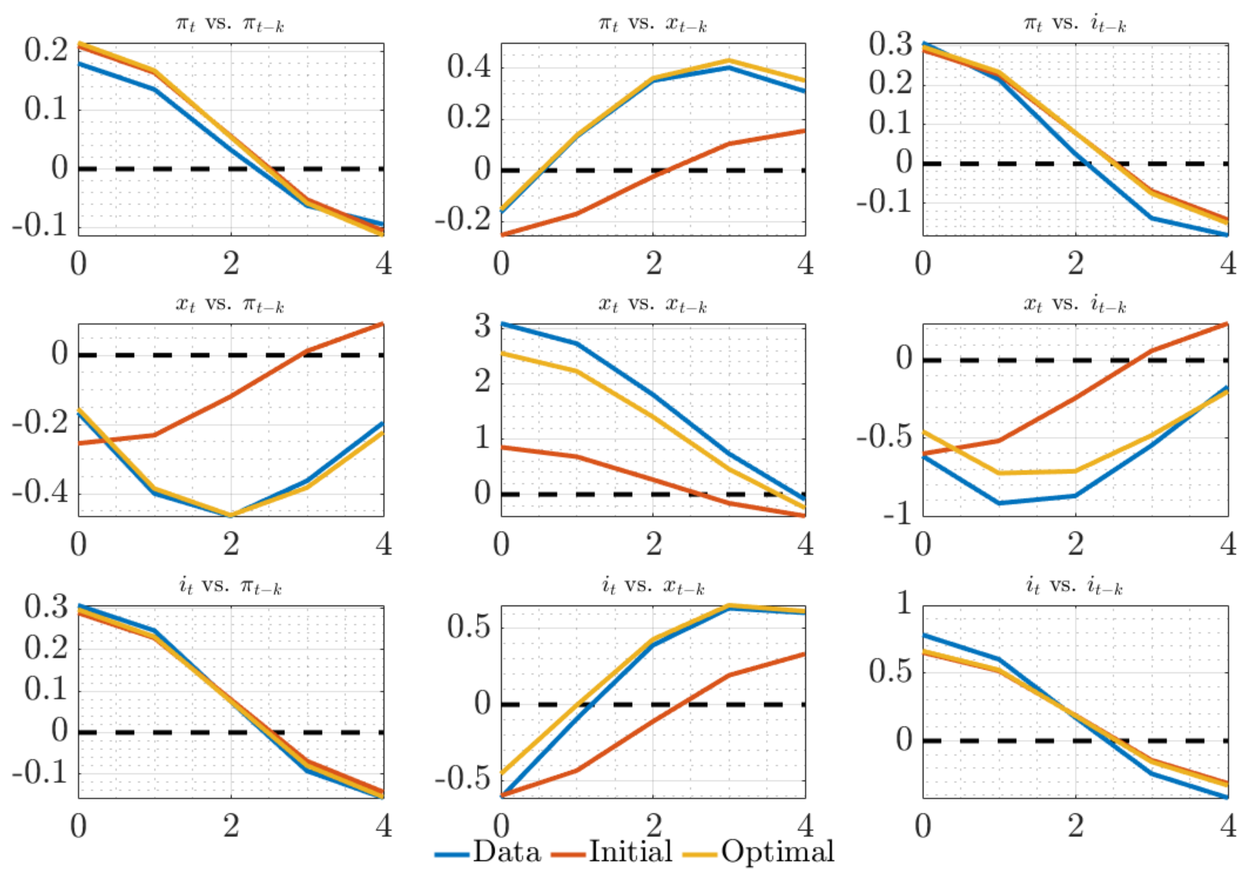
(b) Autocovariogram

7.5 0 at 0

Figure 20: 0 at 0 imposed with weight 1000 not using 1-step ahead forecasts of inflation, not rescaling W, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$



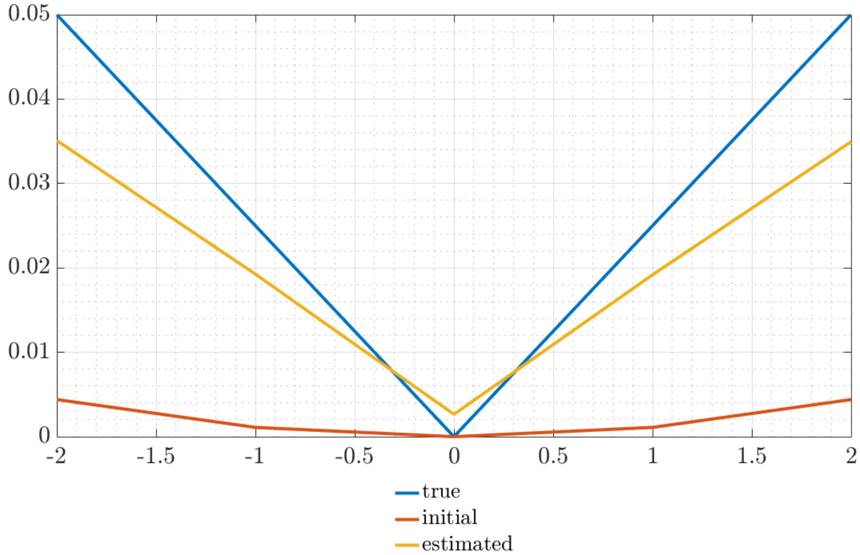
(a) Estimated parameters



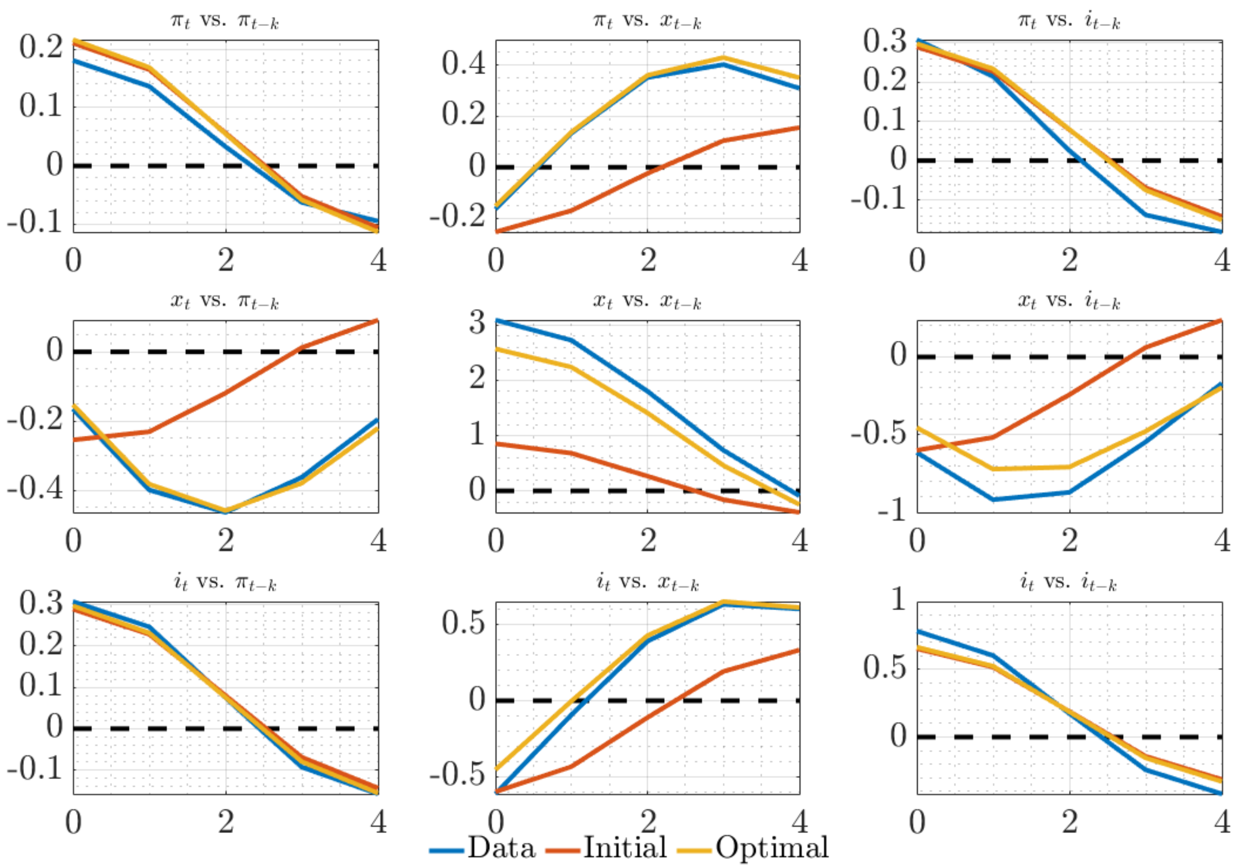
(b) Autocovariogram

7.6 0 at 0, more convexity

Figure 21: 0 at 0 imposed with weight 1000 not using 1-step ahead forecasts of inflation, not rescaling W, estimate mean moments once, imposing convexity with weight 1000K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$



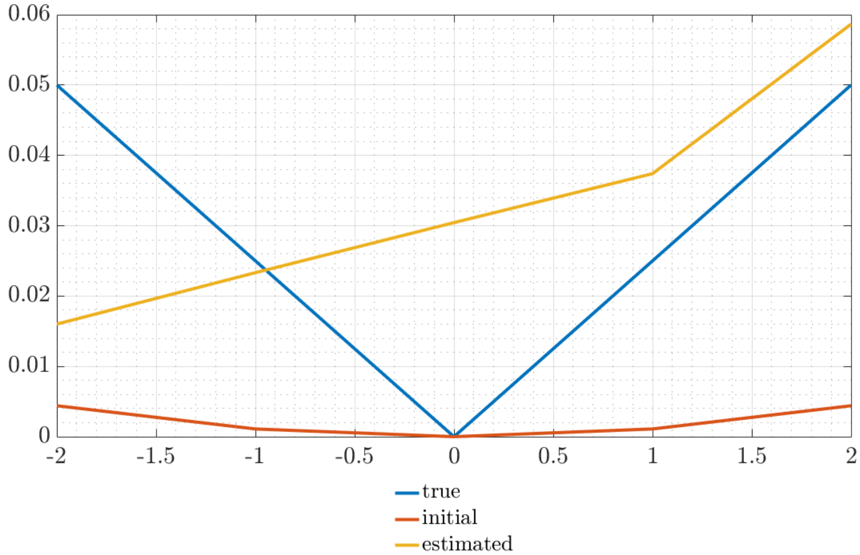
(a) Estimated parameters



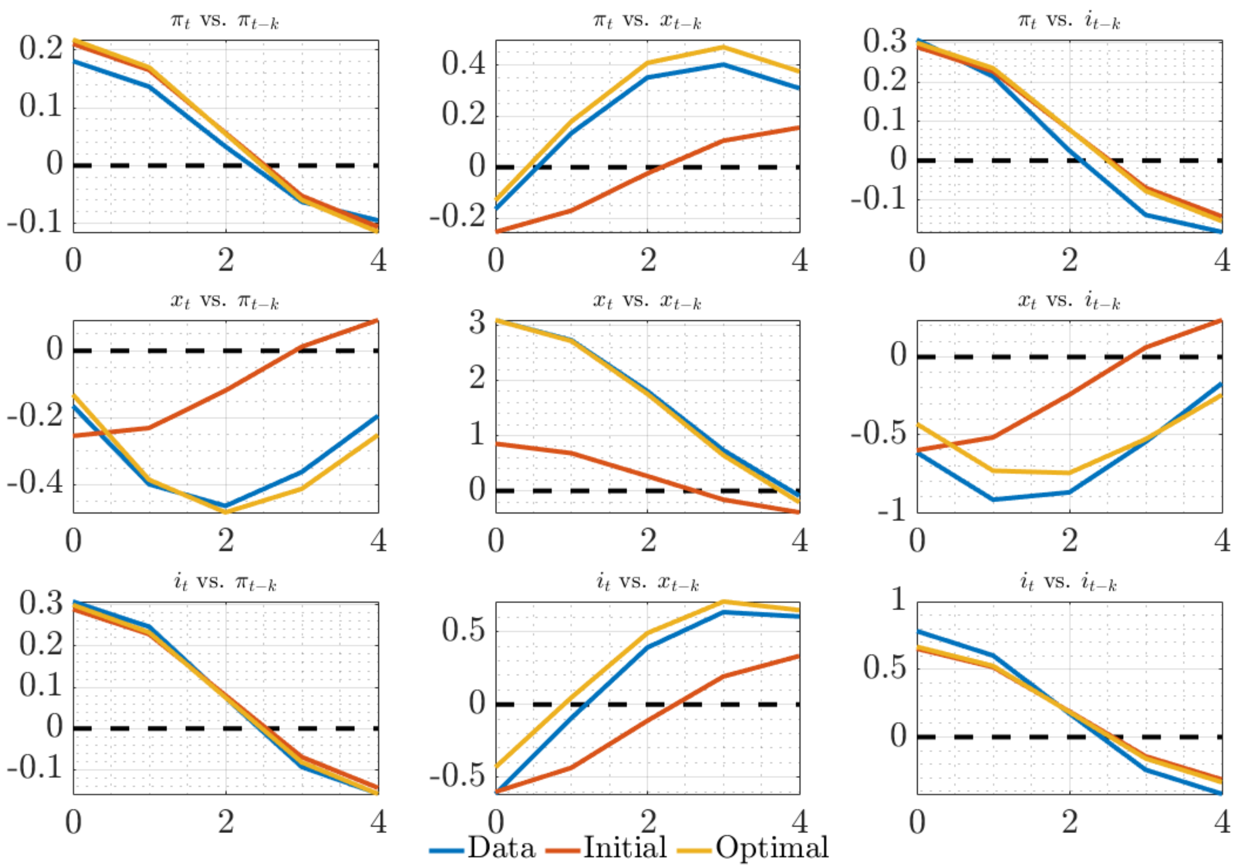
(b) Autocovariogram

7.7 Identity weighting matrix

Figure 22: identity weighting matrix, not using 1-step ahead forecasts of inflation, not rescaling W, estimate mean moments once, imposing convexity with weight 1000K, w/o measurement error, truth with $nfe = 5, fe \in (-2, 2)$



(a) Estimated parameters



(b) Autocovariogram

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, $k_t^{-1} = \mathbf{g}(f e_{t-1})$):

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

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

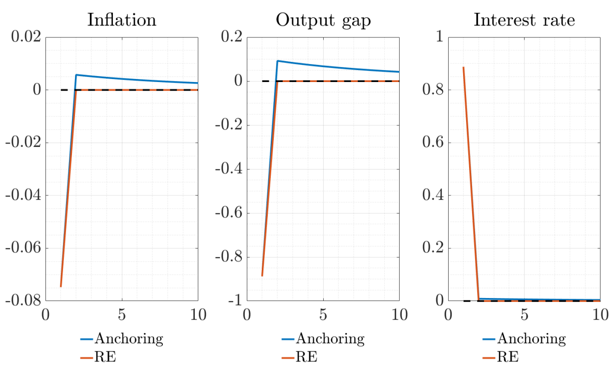
$$\pi_t = -\frac{\lambda_x}{\kappa} x_t + \frac{\lambda_x}{\kappa}\frac{(1-\alpha)\beta}{1-\alpha\beta}\Big(k_t^{-1} + f e_{t|t-1}^{eve}\mathbf{g}_\pi(t)\Big)\mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i}$$
$$-\frac{\lambda_x}{\kappa}\frac{(1-\alpha)\beta}{1-\alpha\beta}\Big(k_t^{-1} + f e_{t|t-1}^{eve}\mathbf{g}_\pi(t)\Big)\Big(\mathbb{E}_t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

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

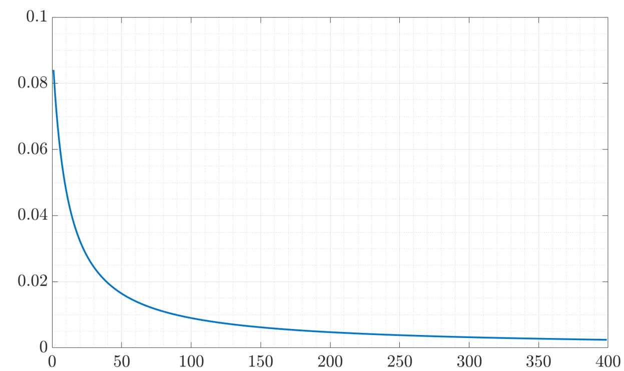
C Impulse responses to iid monpol shocks across a wide range of learning models

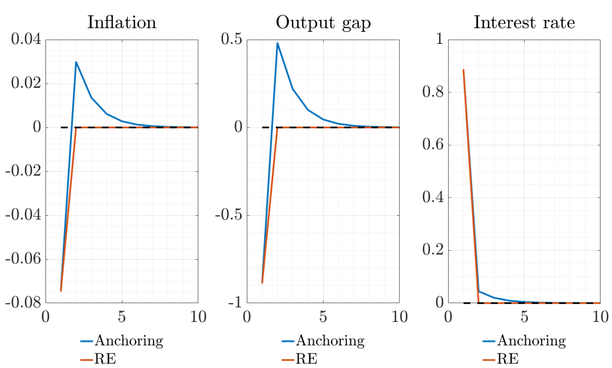
$T = 400, N = 100, n_{drop} = 5$, shock imposed at $t = 25$, calibration as above, Taylor rule assumed to be known, PLM = learn constant only, of inflation only.

Figure 23: IRFs and gain history (sample means)

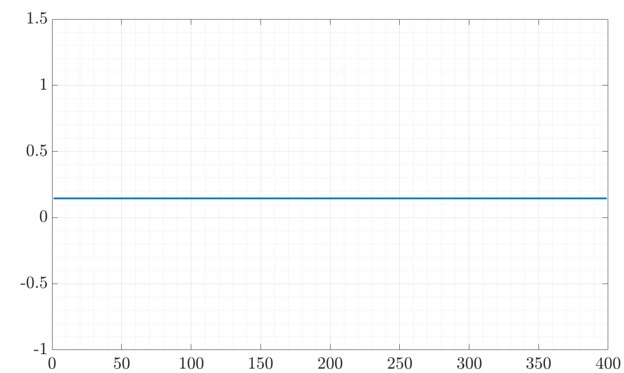


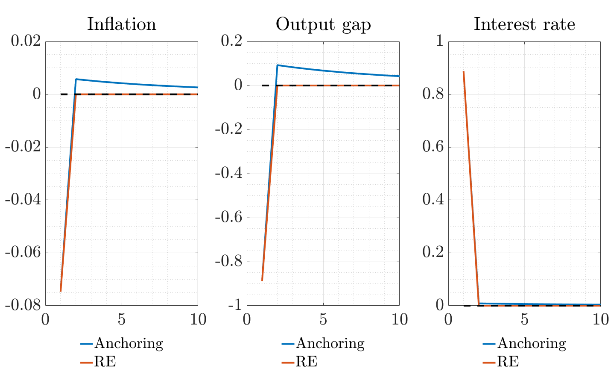
(a) Decreasing gain learning


(b) Mean gain



(c) Constant gain learning


(d) Mean gain



(e) CEMP criterion (vector)

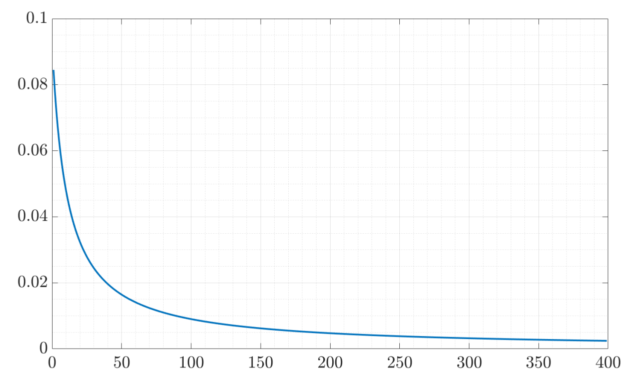
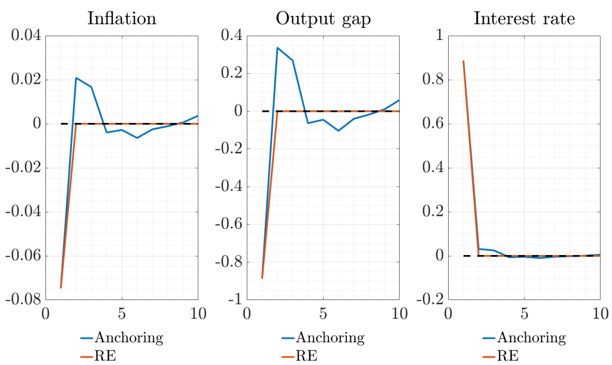
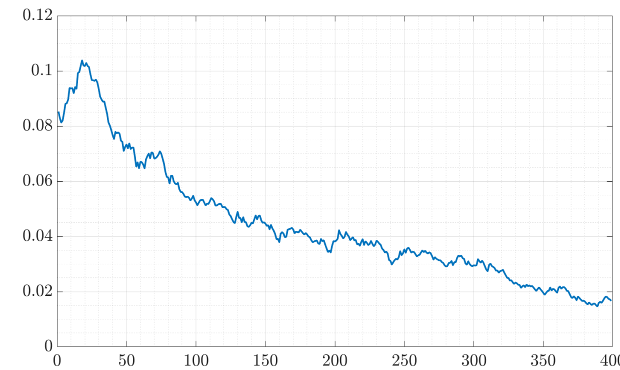

(f) Mean gain

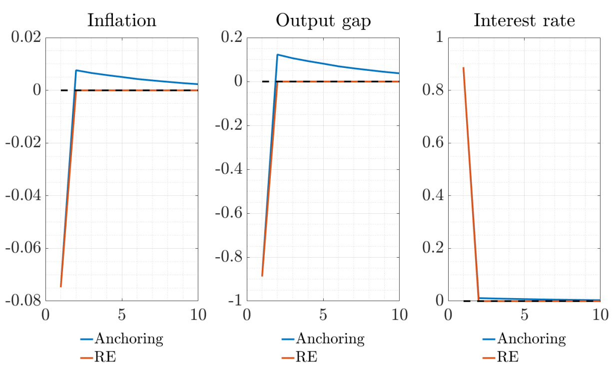
Figure 24: IRFs and gain history (sample means), continued



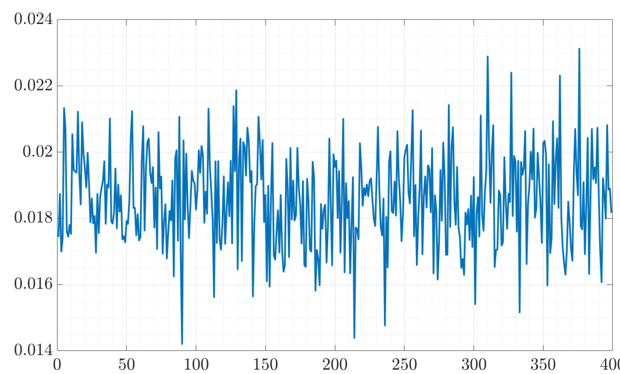
(a) CUSUM criterion (vector)



(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