

Materials 38 - Bias in the neighborhood of zero forecast errors

Laura Gáti

July 20, 2020

Overview

1 Simulated “true” data	2
1.1 3 potential causes to lack of identification in the zero neighborhood	2
1.2 Reference for comparison: Fig 1. of Materials 37	2
1.3 Potential cause # 1: skewness \rightarrow take median instead of mean	3
1.4 Issue # +1: do N simulations instead of N estimations	3
A Model summary	4
B Target criterion	4
C Impulse responses to iid monopol shocks across a wide range of learning models	5

1 Simulated “true” data

1.1 3 potential causes to lack of identification in the zero neighborhood

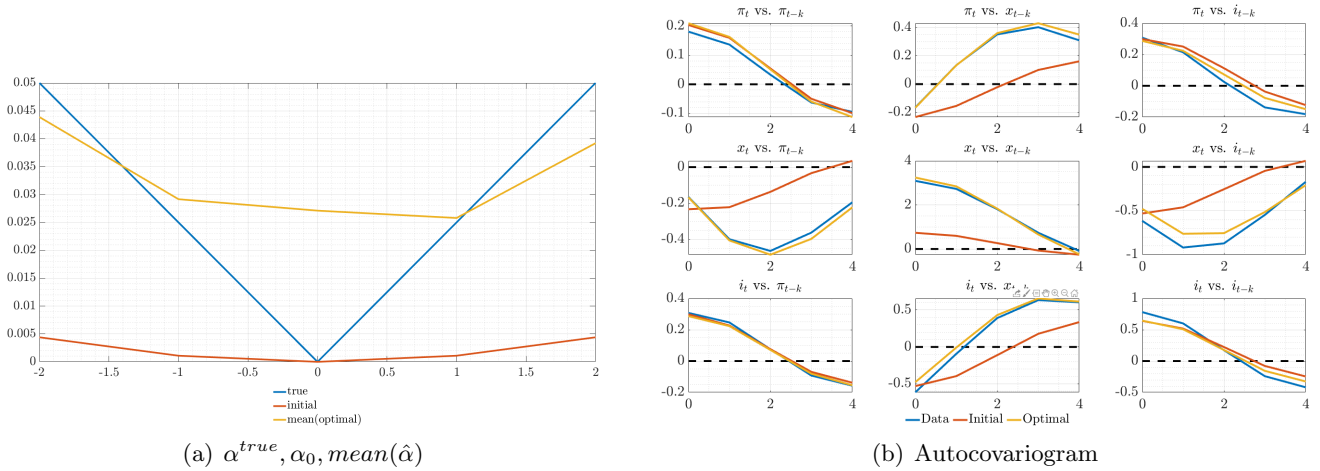
1. The distribution of estimates is skewed \rightarrow take $median(\hat{\alpha})$ instead of the mean.
2. The truth is based on a simulation that doesn't favor the zero neighborhood \rightarrow do 100 simulations from the “true” parameters and take the mean moments of those.
3. The gain doesn't matter if the forecast error is 0, or very close to it \rightarrow introduce a distinction between the forecast error that's used to choose the gain and the one used to update the coefficients of the learning rule.

+1 Taking mean moments across N histories is more natural than performing the estimation N times.

+2 Introduce expectation series (SPF)

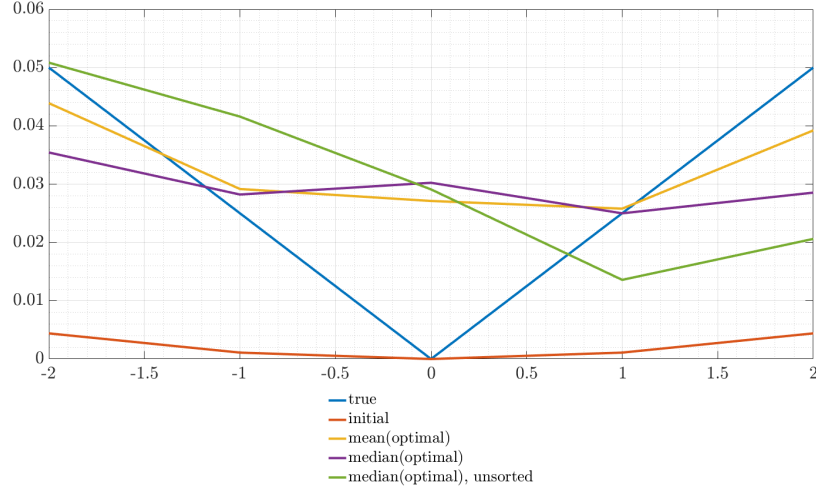
1.2 Reference for comparison: Fig 1. of Materials 37

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



1.3 Potential cause # 1: skewness \rightarrow take median instead of mean

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

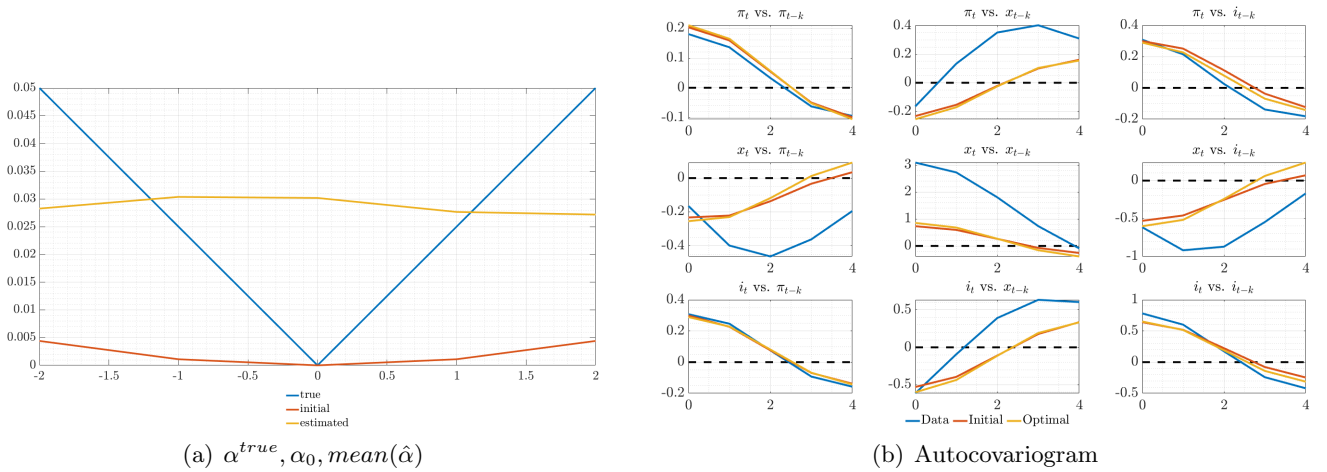


(a) $\alpha^{true}, \alpha_0, \text{mean}(\hat{\alpha}), \text{median}(\hat{\alpha}), \text{unsorted median}$

I understand what's happening! Half the estimates are L's, the other half are “inverted L's”, which is why taking a mean or a classical, sorted median has the tendency to produce these nonmonotonic zigzags.

1.4 Issue # +1: do N simulations instead of N estimations

Figure 3: Estimates for $N = 100$, targeting mean moments in a single estimation instead of N estimations of individual moments, imposing convexity with weight 100K, truth with $nfe = 5, fe \in (-2, 2)$



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

(b) Autocovariogram

The difference is striking!

A Model summary

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

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

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

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

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

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

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

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

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

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

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

B Target criterion

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

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

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

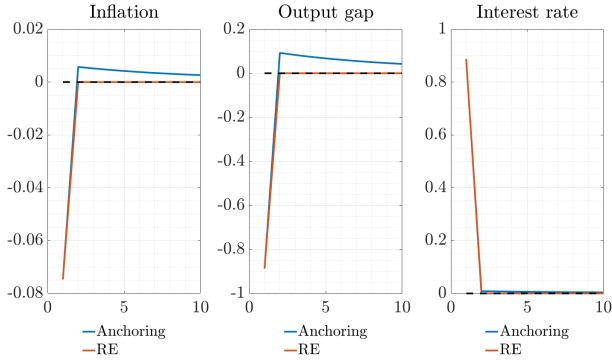
$$\pi_t = -\frac{\lambda_x}{\kappa} x_t + \frac{\lambda_x}{\kappa} \frac{(1-\alpha)\beta}{1-\alpha\beta} \left(k_t^{-1} + f e_{t|t-1}^{eve} \mathbf{g}_\pi(t) \right) \mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i} \\ - \frac{\lambda_x}{\kappa} \frac{(1-\alpha)\beta}{1-\alpha\beta} \left(k_t^{-1} + f e_{t|t-1}^{eve}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

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**

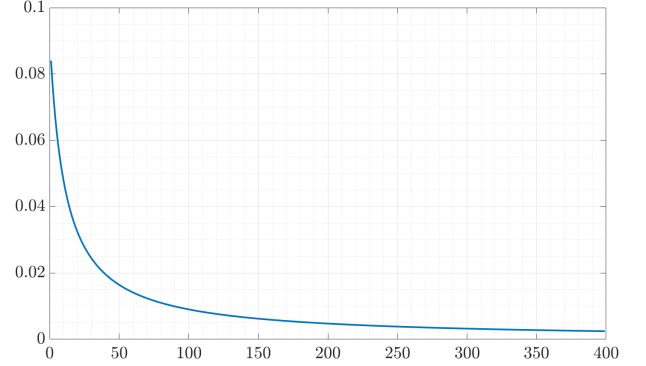
C Impulse responses to iid monpol shocks across a wide range of learning models

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

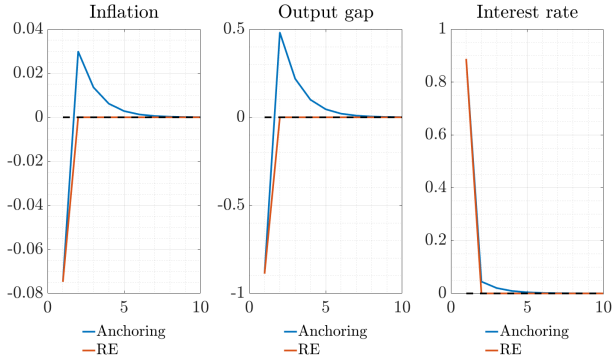
Figure 4: IRFs and gain history (sample means)



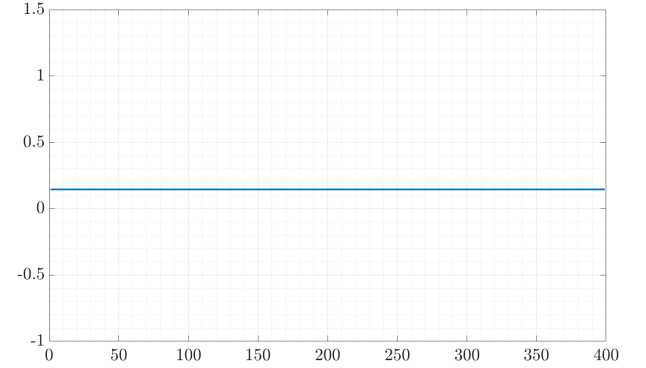
(a) Decreasing gain learning



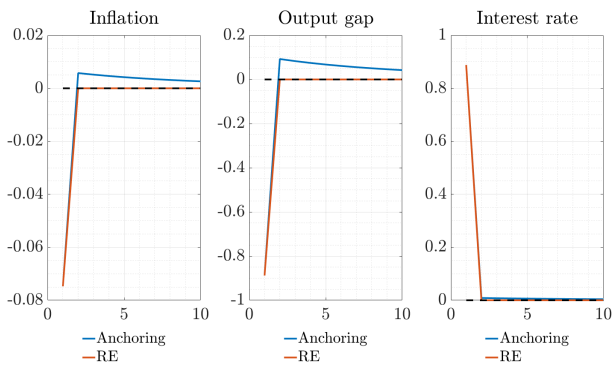
(b) Mean gain



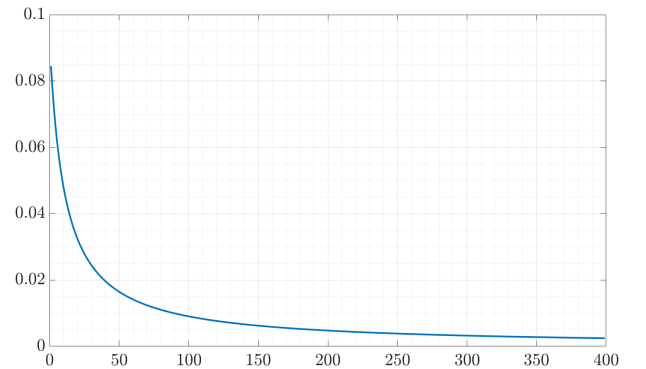
(c) Constant gain learning



(d) Mean gain

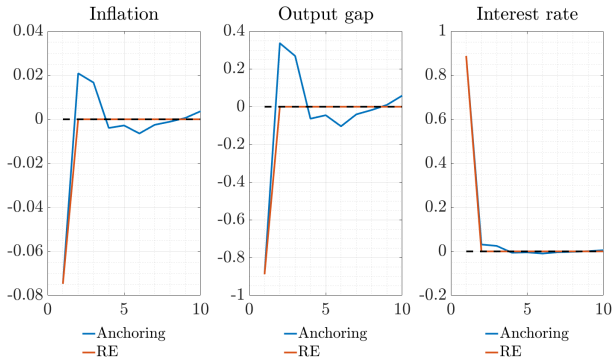


(e) CEMP criterion (vector)

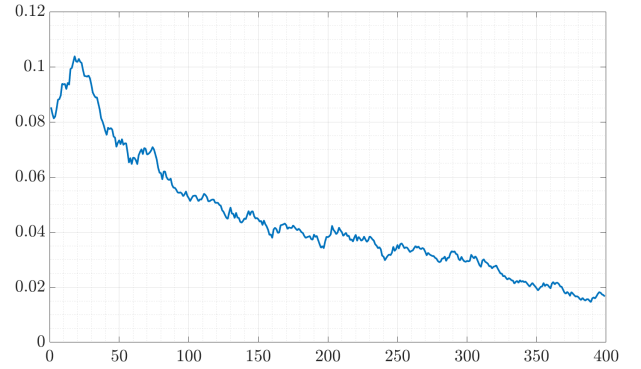


(f) Mean gain

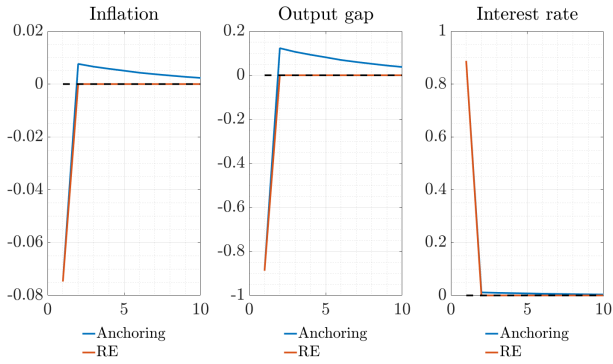
Figure 5: IRFs and gain history (sample means), continued



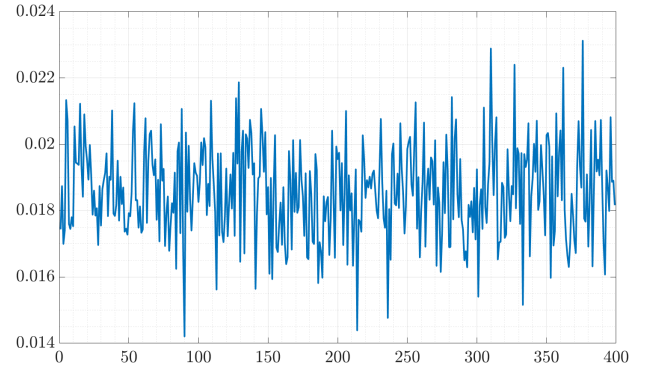
(a) CUSUM criterion (vector)



(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