

Materials 41 - Need large forecast errors for identification

Laura Gáti

August 13, 2020

Overview

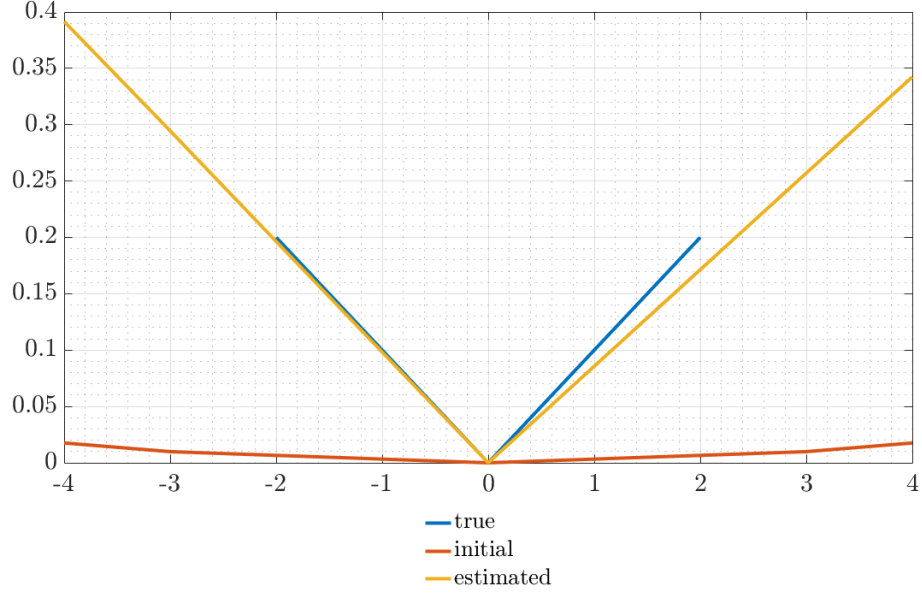
1	What I need for identification	2
2	Only α corresponding to large forecast errors are identified - Figure 15 from Materials 40	3
3	GMM weighting matrix mystery solved: I didn't scale the weight on the convexity moments	4
3.1	Taking square root of elements of W	4
A	Model summary	6
B	Target criterion	6
C	Impulse responses to iid monopol shocks across a wide range of learning models	7

1 What I need for identification

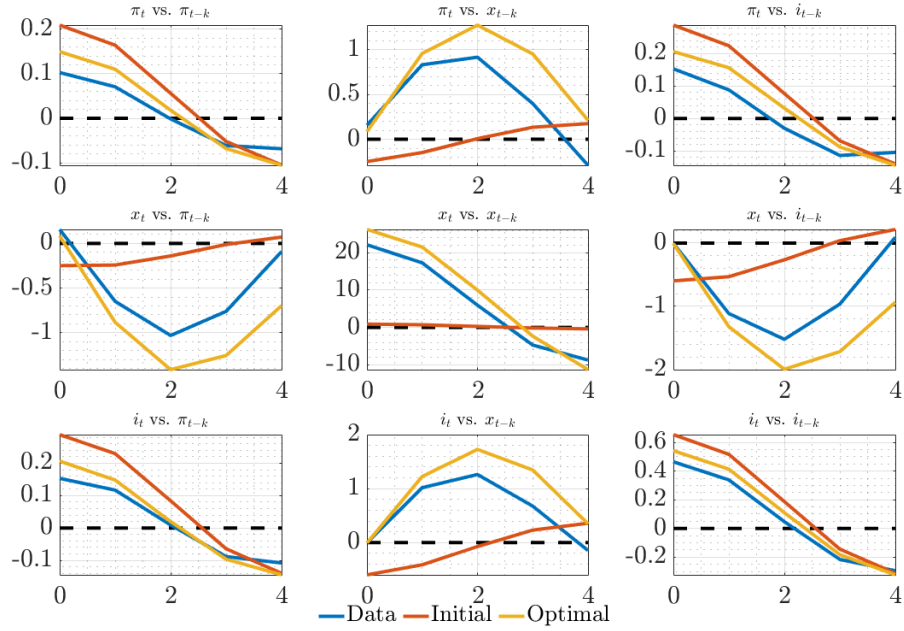
- Need approximating coefficients to pertain to large forecast errors (α s out in the edges)
- Need those large forecast errors to occur in the sample
 - scaled up the “true” α s
 - could play around with variance of shocks, σ_i^2

2 Only α corresponding to large forecast errors are identified - Figure 15 from Materials 40

Figure 1: 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

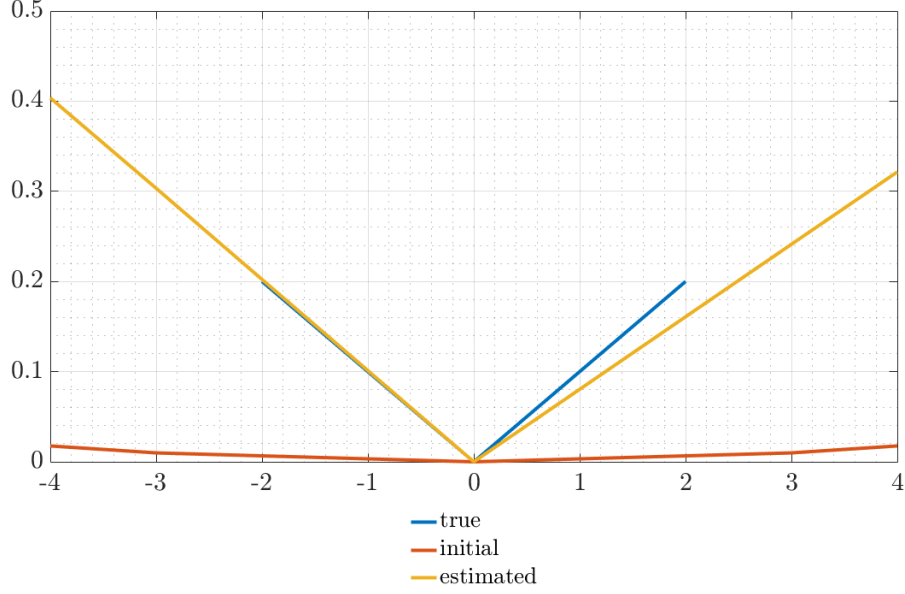


(b) Autocovariogram

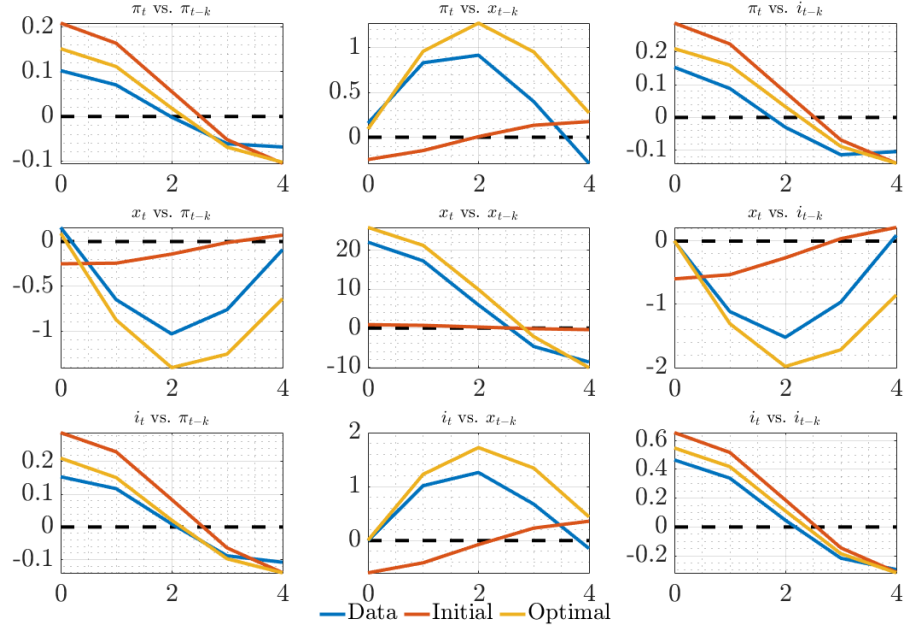
3 GMM weighting matrix mystery solved: I didn't scale the weight on the convexity moments

3.1 Taking square root of elements of W

Figure 2: 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, taking square root of elements of W

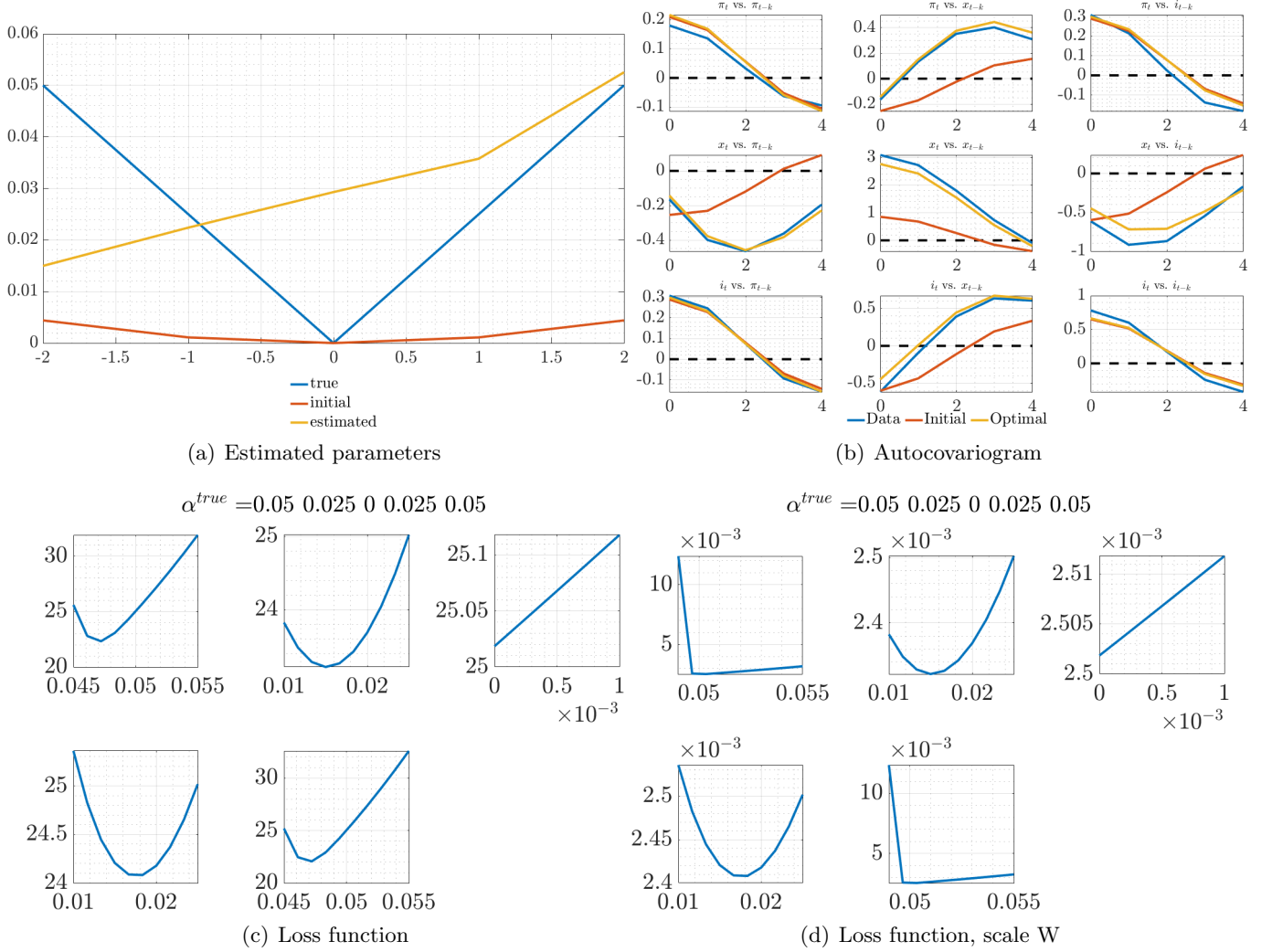


(a) Estimated parameters



(b) Autocovariogram

Figure 3: Not using 1-step ahead forecasts of inflation, estimate mean moments once, imposing convexity with weight 100K, w/o measurement error, truth with $nfe = 5$, $fe \in (-2, 2)$, **taking square root of elements of W** (This is to be compared with the default Nsimulations figure.)



- Rescaling is still doing the same thing to the loss as before: it pushes $\alpha_{1,5}$ up to the true value and makes the loss function nonsmooth. Why?

→ I've got why! It's the convexity restriction whose weight stays constant when I rescale, so effectively, the convexity moments become relatively more important when I rescale!

- How do I know? I've plotted the losses w/ and w/o rescaling when setting the weight on the convexity moment to zero and I see absolutely no change on the shape of the loss function!

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

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

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

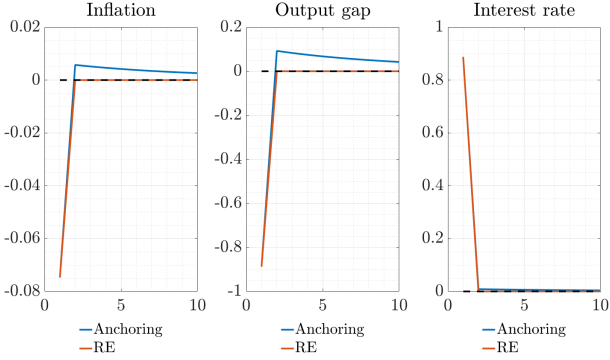
$$\begin{aligned} \pi_t = & -\frac{\lambda_x}{\kappa} x_t + \frac{\lambda_x}{\kappa} \frac{(1-\alpha)\beta}{1-\alpha\beta} \left(k_t^{-1} + f e_{t|t-1}^{eve} \mathbf{g}_\pi(t) \right) \mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i} \\ & - \frac{\lambda_x}{\kappa} \frac{(1-\alpha)\beta}{1-\alpha\beta} \left(k_t^{-1}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

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**

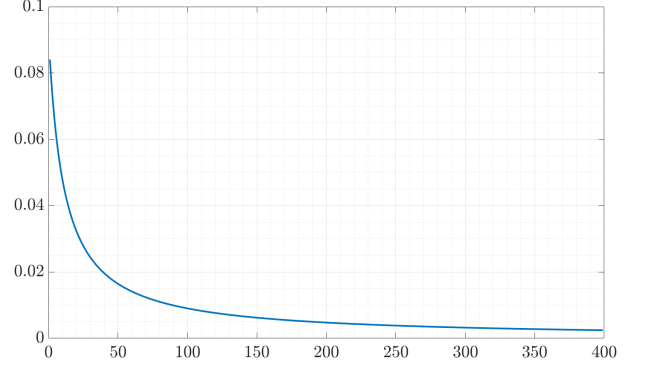
C Impulse responses to iid monpol shocks across a wide range of learning models

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

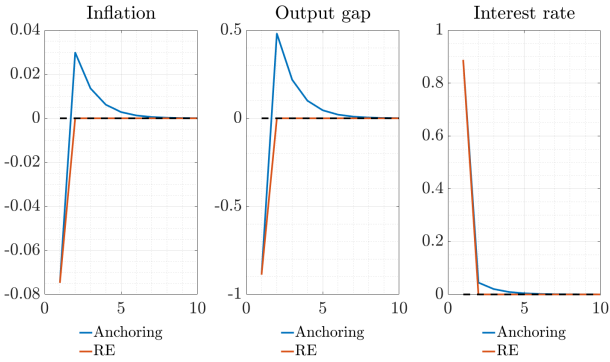
Figure 4: IRFs and gain history (sample means)



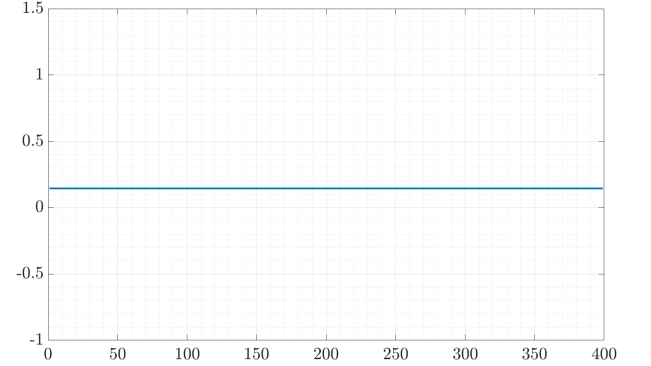
(a) Decreasing gain learning



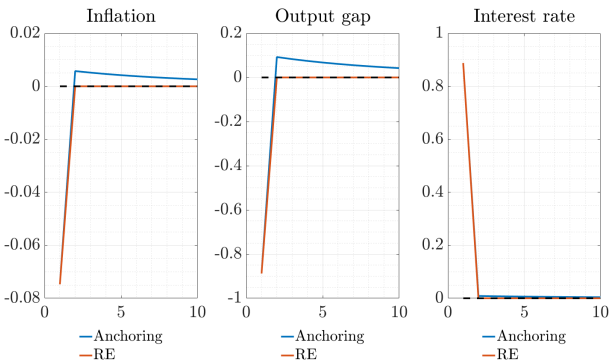
(b) Mean gain



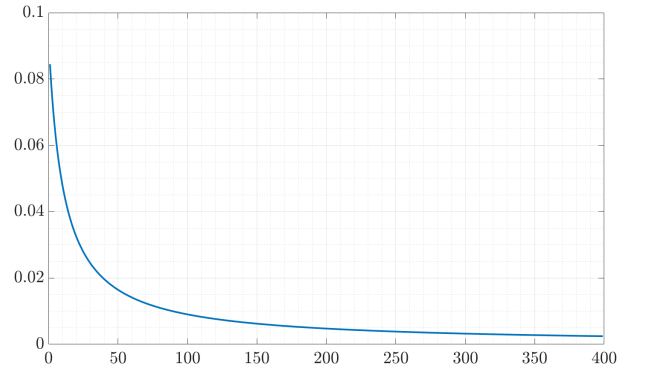
(c) Constant gain learning



(d) Mean gain

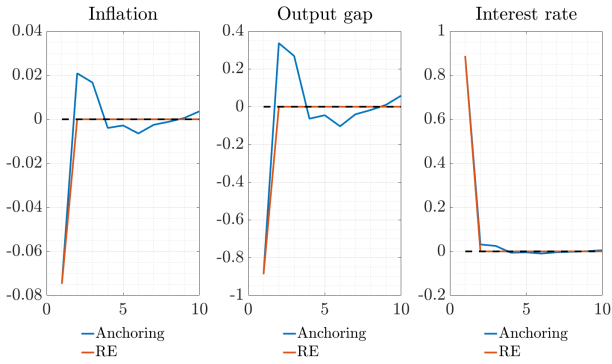


(e) CEMP criterion (vector)

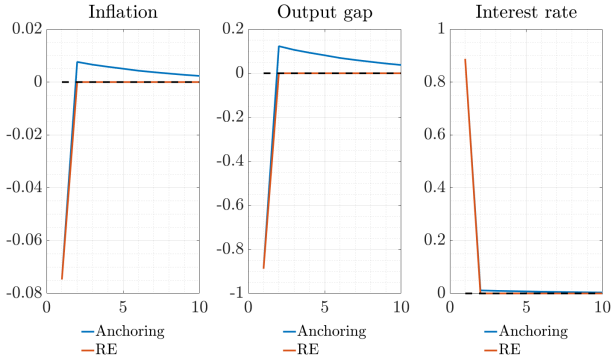


(f) Mean gain

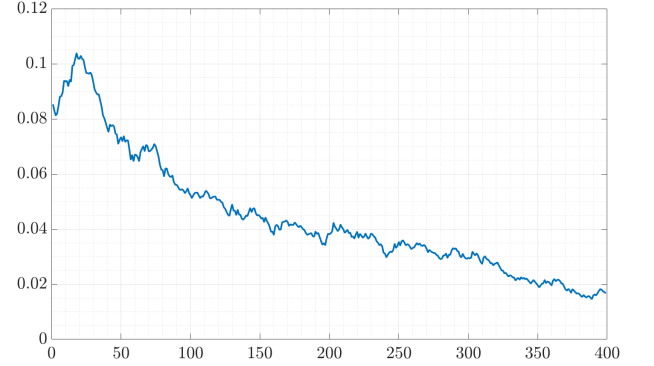
Figure 5: IRFs and gain history (sample means), continued



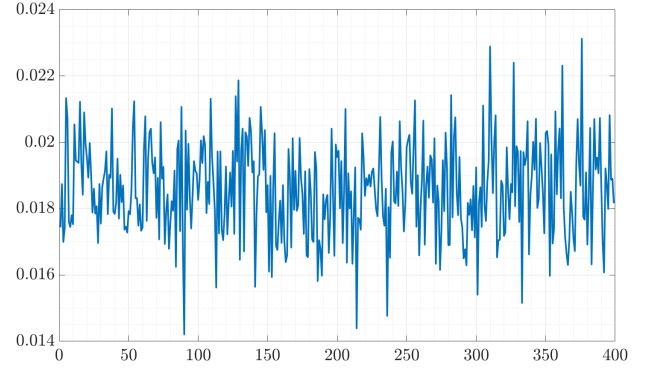
(a) CUSUM criterion (vector)



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



(b) Mean gain



(d) Mean gain