Materials 9 - Is overshooting endemic to constant gain learning?

Laura Gáti

November 15, 2019

Overview

1 Model summary

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

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

$$i_t = \psi_\pi \pi_t + \psi_x x_t + \rho i_{t-1} + \bar{i}_t \tag{3}$$

I consider two variations of the learning rule. The first is a "mean-only" rule:

$$\hat{\mathbb{E}}_t z_{t+h} = \begin{bmatrix} \bar{\pi}_{t-1} \\ 0 \\ 0 \end{bmatrix} + b h_x^{h-1} s_t \quad \forall h \geq 1 \quad b = g_x\ h_x, \qquad \text{PLM1} \tag{4}$$

$$\text{but the first row of } b \text{ is} \quad b_1 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \tag{5}$$

$$\bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1} \underbrace{\big(\pi_t - \bar{\pi}_{t-1}\big)}_{\text{fcst error using (4)}} \tag{6}$$

The second is a "learning the slope too" rule:

$$\hat{\mathbb{E}}_t z_{t+h} = \begin{bmatrix} \bar{\pi}_{t-1} \\ 0 \\ 0 \end{bmatrix} + b_{t-1} h_x^{h-1} s_t \quad \forall h \geq 1 \quad b = g_x\ h_x, \qquad \text{PLM2} \tag{7}$$

$$\text{but the first row of } b \text{ is } b_{1,t} \text{ and is also learned. Let } \phi_t = \begin{bmatrix} \bar{\pi}_t & b_{1,t} \end{bmatrix} \tag{8}$$

$$\phi_t = \left(\phi_{t-1}' + k_t^{-1} \underbrace{\left(\pi_t - \phi_{t-1} \begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix}\right)}_{\text{fcst error using (7)}}\right)' \tag{9}$$

2 Compact notation

$$z_t = A_p^{RE}\,\mathbb{E}_t\, z_{t+1} + A_s^{RE} s_t \tag{10}$$

$$z_t = A_a^{LH} f_a(t) + A_b^{LH} f_b(t) + A_s^{LH} s_t \tag{11}$$

$$s_t = P s_{t-1} + \epsilon_t \qquad \to \quad s'_t = hx\; s'_{t-1} + \epsilon'_t \tag{12}$$

$$\text{where}\quad s'_t \equiv \begin{pmatrix} r_t^n \\ \bar{i}_t \\ u_t \\ i_{t-1} \end{pmatrix} \quad hx \equiv \begin{pmatrix} \rho_r & 0 & 0 & 0 \\ 0 & \rho_i & 0 & 0 \\ 0 & 0 & \rho_u & 0 \\ gx_{3,1} & gx_{3,2} & gx_{3,3} & gx_{3,4} \end{pmatrix} \quad \epsilon'_t \equiv \begin{pmatrix} \varepsilon_t^r \\ \varepsilon_t^i \\ \varepsilon_t^u \\ 0 \end{pmatrix} \quad \text{and} \quad \Sigma' = \begin{pmatrix} \sigma_r & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & \sigma_u & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{13}$$

And the A_s^{RE} and A_s^{LH} are given by:

$$A_s^{RE} = \begin{pmatrix} \frac{\kappa\sigma}{w} & -\frac{\kappa\sigma}{w} & 1 - \frac{\kappa\sigma\psi_\pi}{w} & 0 \\ \frac{\sigma}{w} & -\frac{\sigma}{w} & -\frac{\sigma\psi_\pi}{w} & 0 \\ \psi_x(\frac{\sigma}{w}) + \psi_\pi(\frac{\kappa\sigma}{w}) & \psi_x(-\frac{\sigma}{w}) + \psi_\pi(-\frac{\kappa\sigma}{w}) + 1 & \psi_x(-\frac{\sigma\psi_\pi}{w}) + \psi_\pi(1 - \frac{\kappa\sigma\psi_\pi}{w}) & \rho \end{pmatrix} \tag{14}$$

$$A_s^{LH} = \begin{pmatrix} g_{\pi s} \\ g_{xs} \\ \psi_\pi g_{\pi s} + \psi_x g_{xs} + \begin{bmatrix} 0 & 1 & 0 & \rho \end{bmatrix} \end{pmatrix} \tag{15}$$

$$g_{\pi s} = (1 - \frac{\kappa\sigma\psi_\pi}{w}) \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} (I_4 - \alpha\beta hx)^{-1} - \frac{\kappa\sigma}{w} \begin{bmatrix} -1 & 1 & 0 & \rho \end{bmatrix} (I_4 - \beta hx)^{-1} \tag{16}$$

$$g_{xs} = \frac{-\sigma\psi_\pi}{w} \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} (I_4 - \alpha\beta hx)^{-1} - \frac{\sigma}{w} \begin{bmatrix} -1 & 1 & 0 & \rho \end{bmatrix} (I_4 - \beta hx)^{-1} \tag{17}$$

3 Recap of timing

Define some objects: *(I usually let t denote the time in which the variable is formed.)*

$$f_t^j = \hat{\mathbb{E}}_t(z_{t+1}) \quad \text{one-period-ahead forecast formed at time } t,\ j = m, e \text{ (morning or evening)} \tag{18}$$

$$FE_t = z_{t+1} - f_t \quad \text{one-period-ahead forecast error realized at time } t+1 \tag{19}$$

$$= ALM(t+1) - PLM(t) \tag{20}$$

$$\theta_t = \hat{\mathbb{E}}_{t-1}(z_t) - \mathbb{E}_{t-1}(z_t) \quad \text{CEMP's criterion} \tag{21}$$

$$= PLM(t-1) - \mathbb{E}_{t-1}\, ALM(t) \tag{22}$$

$PLM(t) : \hat{\mathbb{E}}_t z_{t+1} = \bar{z}_{t-1} + bs_t$

Morning: morning of time t available: $\mathcal{I}_t^m = \{\bar{z}_{t-1}, s_t, k_{t-1}, FE_{t-2}\}$

1. Form all future expectations using $PLM(t)$(morning forecast) $\to z_t$ realized, $\to FE_{t-1}$ realized

2. Form $\theta_t \to k_t$ realized

3. **Evening**: Update $\bar{z}_t = \bar{z}_{t-1} + k_t^{-1}(FE_{t-1}^e)$

   where $FE_{t-1}^e = z_t - f_{t-1}^e = z_t - (\bar{z}_{t-1} + bs_{t-1})$ is the most recent realized FE, so:

   $\bar{z}_t = \bar{z}_{t-1} + k_t^{-1}(z_t - (\bar{z}_{t-1} + bs_{t-1}))$

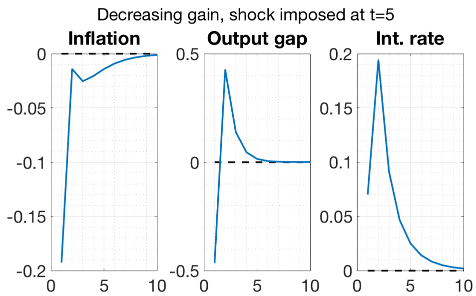
   $\to$ evening of time t available: $\mathcal{I}_t^e = \{\bar{z}_t, s_t, k_t, FE_{t-1}\}$

4 Current set of baseline parameters

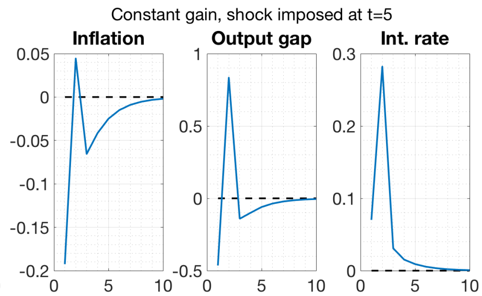
β	0.99	stochastic discount factor	standard (Woodford 2003/2011)
σ	1	IES	consistent with balanced growth
α	0.5	Calvo probability of not adjusting	match 6-month duration of prices (can increase to 0.75)
ψ_π	1.5	coefficient of inflation in Taylor rule	Taylor
ψ_x	0	coefficient of output gap in Taylor rule	focus on π
$\bar{g}$	0.145	value of the constant gain	CEMP
$\bar{\theta}$	1	threshold deviation between $\hat{\mathbb{E}}$ & $\mathbb{E}$	CEMP: 0.029
ρ_r	0	persistence of natural rate shock	n.a.
ρ_i	0.6	persistence of monetary policy shock	CEMP: 0.877 (can increase to 0.78 if $\alpha = 0.75$)
ρ_u	0	persistence of cost-push shock	CEMP
σ_r	0.1	standard deviation of natural rate shock	n.a.
σ_i	0.359	standard deviation of mon. policy shock	CEMP
σ_u	0.277	standard deviation of cost-push shock	CEMP
θ	10	price elasticity of demand	Woodford 2003/2011, Chari, Kehoe & McGrattan 2000
ω	1.25	elasticity of marginal cost to output	Woodford 2003/2011, Chari, Kehoe & McGrattan 2000

5 Cross-sectional IRFs, mon. pol shock only, cgain & dgain only ◀

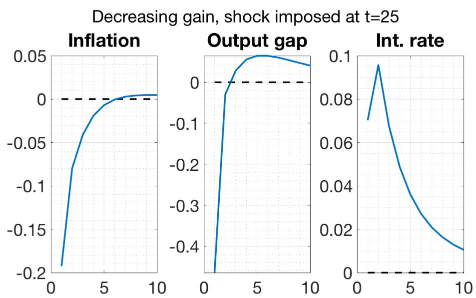
Figure 1: IRF for observables, shock imposed at t



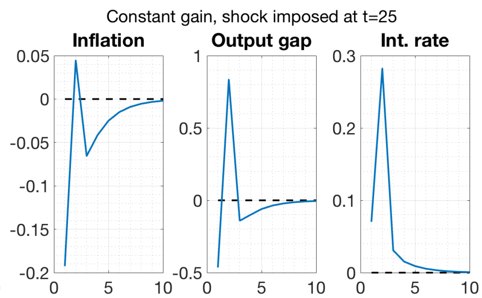
(a) Decreasing gain, $t = 5$



(b) Constant gain, $t = 5$

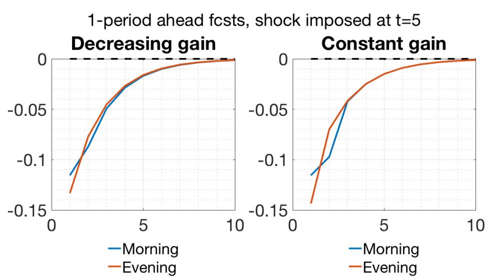


(c) Decreasing gain, $t = 25$

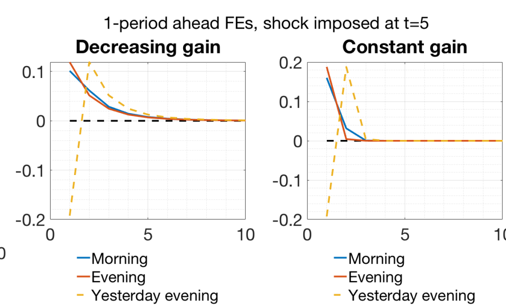


(d) Constant gain, $t = 25$

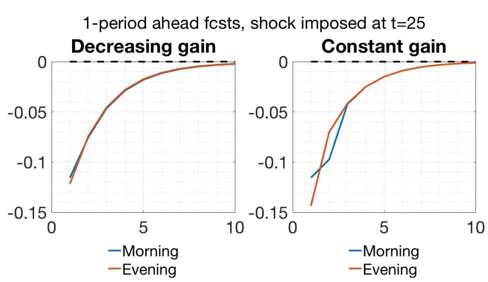
Figure 2: IRF for 1-period ahead forecasts and FEs, together, morning and evening, shock imposed at t



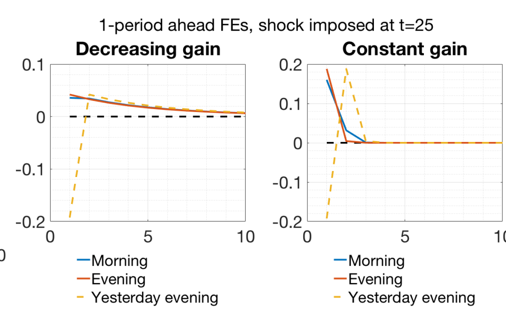
(a) Forecasts, $t = 5$



(b) Forecast errors, $t = 5$

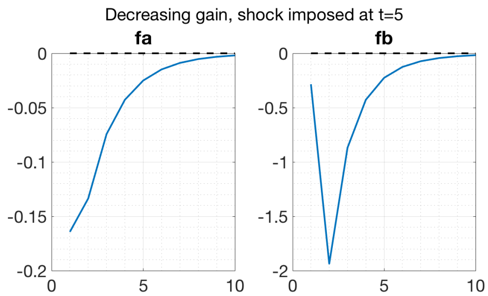


(c) Forecasts, $t = 25$

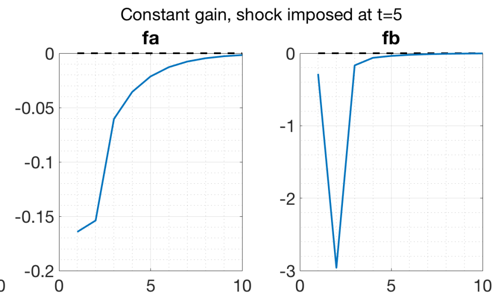


(d) Forecast errors, $t = 25$

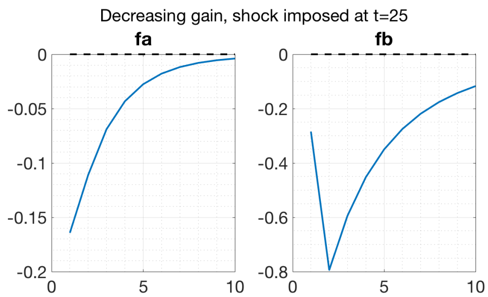
Figure 3: IRF for LH forecasts, shock imposed at t



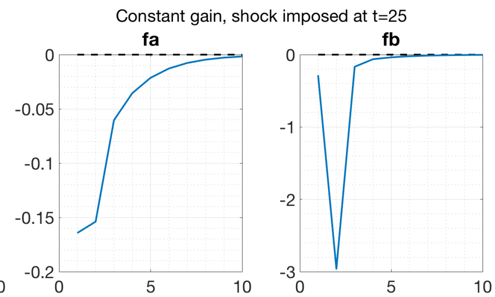
(a) Decreasing gain, $t = 5$



(b) Constant gain, $t = 5$



(c) Decreasing gain, $t = 25$



(d) Constant gain, $t = 25$

4 puzzling things, 4 attempts to explain them:

1. Criss-crossing of expectations

   If the gain is high enough, expectations overupdate, FE switch signs. (Can make extremely small with lower gain, e.g. $\bar{g} = 0.02$, the value of the dgain after 50 periods.)

   Note also that f_a, f_b are mainly driven by the intercept, not the slope. The intercept term is $\frac{\bar{\pi}}{1-\alpha\beta}$ and $\frac{\bar{\pi}}{1-\beta}$. So when α is far from 1, f_b is much bigger

   - both in absolute terms,
   - as well as relatively to f_a.

2. $i \uparrow$ at $t = 2$ after shock

   $i = \pi \downarrow + \rho_i^{t-1}\delta \uparrow$. So b/c of correction in π at $t = 2$, $|\rho_i^{t-1}\delta| - |\pi|$ becomes larger.

3. Overshooting later than $t = 2$

   IR decomposed into two effects pulling opposite ways: $IR \approx \mathbb{E}(\pi) + \delta$. Overshooting happens when the expectation effect is larger. As the shock recedes faster than the expectation, x, π

overshoot. Testable implication: if shock iid, overshooting should happen at $t=2$, and indeed it does.

4. Responses to expectations in RE vs learning (recursive vs. LH-horizon)

   Here's the stylized representation of how endogenous variables respond to expectations in the two formulations (once you've plugged in the interest rate):

$$\begin{aligned}
&\textit{RE} \\
&x_t = \overset{-}{\mathbb{E}}(\pi) + \overset{+}{\mathbb{E}}(x) \\
&\pi_t = \overset{+}{\mathbb{E}}(\pi) + \overset{+}{\mathbb{E}}(x) \\
&\textit{Learning} \\
&x_t = \overset{-}{\mathbb{E}}(\pi) + \overset{-}{\mathbb{E}}(x) \\
&\pi_t = \overset{+}{\mathbb{E}}(\pi) + \overset{+}{\mathbb{E}}(x)
\end{aligned}$$

   The difference, marked in red, comes from the fact in learning, expectations are more explicit. Ignoring shocks and setting $\psi_x = 0$, so the Taylor rule is just $i_t = \psi_\pi \pi_t$, the two systems are

$$\begin{aligned}
&\textit{RE} \\
&x_t = -\sigma i_t + \mathbb{E}_t\, x_{t+1} + \sigma\, \mathbb{E}_t\, \pi_{t+1} \\
&\pi_t = \kappa x_t + \beta\, \mathbb{E}_t\, \pi_{t+1} \\
&\textit{Learning} \\
&x_t = -\sigma i_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \beta^{T-t}\big((1-\beta)x_{T+1} - \sigma\beta i_{T+1} + \sigma\pi_{T+1})\big) \\
&\pi_t = \kappa x_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t}\big(\kappa\alpha\beta x_{T+1} + (1-\alpha)\beta\pi_{T+1}\big)
\end{aligned}$$

   When you plug in the interest rate, you see that the recursive representation hides the negative part of x_t's dependance on its own future values into current inflation and its dependance on future inflation:

$$
\begin{aligned}
&\textit{RE} \\
&x_t = -\sigma\psi_\pi\pi_t + \mathbb{E}_t\, \overset{+}{x}_{t+1} + \sigma\, \mathbb{E}_t\, \pi_{t+1} \\
&\pi_t = \kappa x_t + \beta\, \mathbb{E}_t\, \pi_{t+1} \\
\textit{Learning}& \\
&x_t = -\sigma\psi_\pi\overset{-}{\pi}_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \beta^{T-t}\big((1-\beta)\overset{+}{x}_{T+1} + \sigma(1-\beta\psi_\pi)\pi_{T+1}\big) \\
&\pi_t = \kappa x_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t}\big(\kappa\alpha\beta x_{T+1} + (1-\alpha)\beta\pi_{T+1}\big)
\end{aligned}
$$

To align with this interpretation, it must be the case that π_t's dependance on future inflation is stronger under RE than learning. Comparing coefficients for my baseline parameterization, this is true.