Materials 18 - Optimal monpol under learning is an infinitely repeated prisoner's dilemma without the grim trigger

Laura Gáti

February 24, 2020

Take a very simple optimal policy problem where the planner chooses $\{\pi_t\}_{t=t_0}^{\infty}$ to minimize

$$\mathcal{L} = \mathbb{E}_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ \pi_t^2 + \varphi_t(\pi_t - \beta \pi_{t+1}) \right\} \tag{1}$$

Consider this same problem under three cases:

1. **RE and commitment**

   FOC:

$$2\pi_t + \varphi_t - \varphi_{t-1} = 0 \tag{2}$$

2. **RE and discretion**

   Write expected inflation as a variable f_t that the authority takes as given:

$$\mathcal{L} = \mathbb{E}_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ \pi_t^2 + \varphi_t(\pi_t - \beta f_t) \right\} \tag{3}$$

   FOC:

$$2\pi_t + \varphi_t = 0 \tag{4}$$

3. **Learning and commitment**

$$\mathcal{L} = \mathbb{E}_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left\{ \pi_t^2 + \varphi_{1,t}(\pi_t - \beta f_t) + \varphi_{2,t}(f_t - f_{t-1} - k^{-1}(\pi_t - f_{t-1})) \right\} \tag{5}$$

   FOCs:

$$2\pi_t + \varphi_{1,t} - \varphi_{2,t} k^{-1} = 0 \tag{6}$$

$$-\beta\varphi_{1,t} + \varphi_{2,t} + \mathbb{E}_t\, \varphi_{2,t+1}(-1 + k^{-1}) = 0 \tag{7}$$

   $\rightarrow$ no lagged multiplier despite taking formation of expectations into account!

Mele, Molnár & Santoro (2019): optimal monetary policy w/ learning does not involve commitment! In a sense, this was anticipated in the “Ramsey policy is indeterminate under RE” literature:

- Ramsey policy is time-inconsistent (Kydland & Prescott, 1977, Barro & Gordon, 1983). So invent RE to endow agents with threats for deviating $\rightarrow$ a commitment device.
- Now the problem is that the commitment solution is a Nash, but it's indeterminate (b/c purely forward-looking).
- Invent learning to introduce backward-lookingness: allows you to select the RE with commitment solution (focus of Evans & Honkapohja 2001 on E-stability).
- Problem: that solution is no longer optimal b/c the PS cannot enforce it! (“machine”)
- Circumvent that by introducing backward-looking expectation formation that satisfies some sense of optimality (Cho and Matsui, 1995, “inductive expectations”) so that the PS regains some of its strategic nature.

Here's a step of bold interpretation:

- Maybe Ramsey under RE is indeterminate because of the folk theorem: any feasible and individually rational payoff profile can be a Nash or a subgame perfect equilibrium of the infinitely repeated prisoner's dilemma.
- Learning breaks the folk theorem b/c the assumption of credible threats is gone when one of the players is an automaton and in particular, not an optimal one.

1 Optimal policy problem under learning w/ a smooth gain function

A simpler version of the problem is: postulate an endogenous gain choice as some differentiable function of the forecast error: $k_t^{-1} = f(\pi_t - \bar{\pi}_{t-1} - bs_{t-1})$. I have the CUSUM-mechanism in mind, but am agnostic about it to simplify the math. (Treat k_t^{-1} as a variable.)

$$\mathcal{L} = \mathbb{E}_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} \Big\{ (\pi_t^2 + \lambda x_t^2) \tag{8}$$

$$+ \varphi_{1,t} \Big(\pi_t - \kappa x_t - (1-\alpha)\beta f_a(t) - \kappa\alpha\beta b_2 (I_3 - \alpha\beta h_x)^{-1} s_t - e_2 (I_3 - \alpha\beta h_x)^{-1} s_t \Big) \tag{9}$$

$$+ \varphi_{2,t} \Big(x_t + \sigma i_t - \sigma f_b(t) - (1-\beta) b_2 (I_3 - \beta h_x)^{-1} s_t + \sigma\beta b_2 (I_3 - \beta h_x)^{-1} s_t - \sigma e_1 (I_3 - \alpha\beta h_x)^{-1} s_t) \Big) \tag{10}$$

$$+ \varphi_{3,t} \Big(f_a(t) - \frac{1}{1-\alpha\beta} \bar{\pi}_{t-1} - b_1 (I_3 - \alpha\beta h_x)^{-1} s_t \Big) \tag{11}$$

$$+ \varphi_{4,t} \Big(f_b(t) - \frac{1}{1-\beta} \bar{\pi}_{t-1} - b_1 (I_3 - \beta h_x)^{-1} s_t \Big) \tag{12}$$

$$+ \varphi_{5,t} \Big(\bar{\pi}_t - \bar{\pi}_{t-1} - k_t^{-1} \big(\pi_t - (\bar{\pi}_{t-1} + bs_{t-1}) \big) \Big) \tag{13}$$

$$+ \varphi_{6,t} \Big(k_t^{-1} - f(\pi_t - \bar{\pi}_{t-1} - bs_{t-1}) \Big) \Big\} \tag{14}$$

where I denote by $f_i(t) \in (0,1)$, $i = \pi, \bar{\pi}$, the potentially time-varying derivatives of the anchoring function $f(t)$. After a little bit of simplifying, the FOCs boil down to the following three equations:

$$2\pi_t + 2\frac{\lambda}{\kappa} x_t - \varphi_{5,t} k_t^{-1} - \varphi_{6,t} f_\pi(t) = 0 \tag{15}$$

$$-\frac{2(1-\alpha)\beta}{1-\alpha\beta} \frac{\lambda}{\kappa} x_{t+1} + \varphi_{5,t} - (1 - k_t^{-1}) \varphi_{5,t+1} + f_{\bar{\pi}}(t) \varphi_{6,t+1} = 0 \tag{16}$$

$$\varphi_{6,t} = (\pi_t - \bar{\pi}_{t-1} - bs_{t-1}) \varphi_{5,t} \tag{17}$$

Equation (15) is the analogue of Gaspar et al's Equation (22) (= Molnár & Santoro's (16)), except that there's an additional multiplier, φ_6, reflecting the fact that learning involves the gain equation as an additional constraint here. Note that when learning has converged ($\varphi_{5,t} = \varphi_{6,t} = 0$), this boils down to the discretionary RE case.

Equation (16) is the analogue of Gaspar et al's Equation (21). The novelty in the system is (17) which reflects the relationship between the constraints imposed on the system by learning. Combining the above three equations and solving for $\varphi_{5,t}$, using the notation that $\prod_{j=1}^{0} = 1$, I get what I will call the anchoring analogue of Gaspar et al's (24), the target criterion (ignoring $\mathbb{E}_t$):

$$\pi_t = -\frac{\lambda}{\kappa}\left\{x_t - \frac{(1-\alpha)\beta}{1-\alpha\beta}\left(k_t^{-1} + ((\pi_t - \bar{\pi}_{t-1} - bs_{t-1}))f_\pi(t)\right)\left(\sum_{i=1}^{\infty} x_{t+i}\prod_{j=1}^{i-1}(1 - k_{t+j}^{-1}(\pi_{t+1+j} - \bar{\pi}_{t+j} - bs_{t+j}))\right)\right\} \tag{18}$$