Online Appendix for "Diagnostic Business Cycles"

Francesco Bianchi, Cosmin Ilut and Hikaru Saijo

A An Investment Model and Endogenous Predictability

In this Appendix we study a two-period investment model that illustrates the point we made in footnote 5 regarding the endogenous predictability.

We start with the problem under RE.

$$\max_{I_1} D_1 + \frac{1}{1+r}\mathbb{E}\left[D_2\right]$$

$$\begin{aligned} D_1 &= A_1 K_1^{\nu} - I_1 \\ D_2 &= A_2 K_2^{\nu} + (1-\delta) K_2 \\ K_2 &= (1-\delta) K_1 + I_1 \end{aligned}$$

where $r > 0$, $0 < \delta < 1$, $0 < \nu < 1$, and K_1 are taken as given. A_t is an i.i.d. Normal process. We can replace the constraints in the object function and derive the FOC:

$$\frac{1}{1+r}\mathbb{E}\left[\nu A_2 \left((1-\delta) K_1 + I_1^*\right)^{\nu-1} + (1-\delta)\right] = 1$$

We obtain:

$$I_1^* = \left[\frac{R+\delta}{\nu \mathbb{E}\left[A_2\right]}\right]^{\frac{1}{\nu-1}} - (1-\delta) K_1$$

Thus, A_1 is irrelevant as long as dividends are allowed to be negative.

In turn, under DE:

$$\max_{I_1} D_1 + \frac{1}{1+r}\mathbb{E}^{\theta}\left[D_2\right]$$

$$\begin{aligned} D_1 &= A_1 K_1^{\nu} - I_1 \\ D_2 &= A_2 K_2^{\nu} + (1-\delta) K_2 \\ K_2 &= (1-\delta) K_1 + I_1 \end{aligned}$$

We can replace the constraints and compute the FOC:

$$\frac{1}{1+r}\mathbb{E}^{\theta}\left[\nu A_2 \left((1-\delta) K_1 + I_1^*\right)^{\nu-1} + (1-\delta)\right] = 1$$

In equilibrium, under the assumption of the model, the optimal investment choice is not stochastic. Thus, despite the non-linearity of the problem, normality is preserved. Since the product of a normal times a constant is still a normal, we get:

$$(1+\theta)\,\nu\left((1-\delta) K_1 + I_1^*\right)^{\nu-1}\mathbb{E}\left[A_2\right] - \theta\nu\left((1-\delta) K_1 + I_1^*\right)^{\nu-1}\mathbb{E}_{-1}\left[A_2\right] = (1+r) - (1-\delta)$$

Given that TFP is i.i.d., we have $\mathbb{E}\left[A_2\right] = \mathbb{E}_{-1}\left[A_2\right]$:

$$\nu\left((1-\delta) K_1 + I_1\right)^{\nu-1}\mathbb{E}\left[A_2\right] = (1+r) - (1-\delta)$$

Then

$$I_1^* = \left[\frac{r+\delta}{\nu\mathbb{E}\left[A_2\right]}\right]^{\frac{1}{\nu-1}} - (1-\delta)\, K_1$$

Thus, the solution is identical to the RE solution. This is because there is no revision in expectations coming from what happens at time 1.

B Omitted Proofs

B.1 Proof of Lemma 1

Proof.

$$\begin{aligned}
\mathbb{E}_t^\theta\left[\mathbb{E}_{t+1}^\theta\left[C_{t+1+n}\right]\right] &= \mathbb{E}_t^\theta\left[\mathbb{E}_{t+1}\left[C_{t+1+n}\right] + \theta\left(\mathbb{E}_{t+1}\left[C_{t+1+n}\right] - \mathbb{E}_{t+1-J}\left[C_{t+1+n}\right]\right)\right] \\
&= \mathbb{E}_t\left[\mathbb{E}_{t+1}\left[C_{t+1+n}\right] + \theta\left(\mathbb{E}_{t+1}\left[C_{t+1+n}\right] - \mathbb{E}_{t+1-J}\left[C_{t+1+n}\right]\right)\right] \\
&+ \theta\left\{\mathbb{E}_t\left[\mathbb{E}_{t+1}\left[C_{t+1+n}\right] + \theta\left(\mathbb{E}_{t+1}\left[C_{t+1+n}\right] - \mathbb{E}_{t+1-J}\left[C_{t+1+n}\right]\right)\right]\right. \\
&\left.-\mathbb{E}_{t-J}\left[\mathbb{E}_{t+1}\left[C_{t+1+n}\right] + \theta\left(\mathbb{E}_{t+1}\left[C_{t+1+n}\right] - \mathbb{E}_{t+1-J}\left[C_{t+1+n}\right]\right)\right]\right\} \\
&= \mathbb{E}_t\left[C_{t+1+n}\right] + \theta\left(\mathbb{E}_t\left[C_{t+1+n}\right] - \mathbb{E}_{t-J}\left[C_{t+1+n}\right]\right) \\
&+ \theta(1+\theta)\left(\mathbb{E}_t\left[C_{t+1+n}\right] - \mathbb{E}_t\mathbb{E}_{t+1-J}\left[C_{t+1+n}\right]\right)
\end{aligned}$$

The term $\left(\mathbb{E}_t\left[C_{t+1+n}\right] - \mathbb{E}_t\mathbb{E}_{t+1-J}\left[C_{t+1+n}\right]\right)$ in the last line is generically zero if and only if $J=1$. Thus, $\mathbb{E}_t^\theta\left[\mathbb{E}_{t+1}^\theta\left[C_{t+1+n}\right]\right] = \mathbb{E}_t^\theta\left[C_{t+1+n}\right]$ if and only if $J=1$. □

B.2 Proof of Proposition 1

Proof. The first order conditions at time 1 are:

$$\begin{aligned}
C_1 &= \mathbb{E}_1\left[C_2\right] \\
\mathbb{E}_1\left[C_2\right] &= \mathbb{E}_1\left[C_3\right]
\end{aligned}$$

or, equivalently

$$\begin{aligned}
Y_1 + K_0 - K_1 &= \mathbb{E}_1\left[Y_2 + K_1 - K_2\right] \\
\mathbb{E}_1\left[Y_2 + K_1 - K_2\right] &= \mathbb{E}_1\left[Y_3 + K_2\right]
\end{aligned}$$

The solution at time 1 and 2 can be obtained with backward induction or a guess-and-verify approach. We opt for the guess-and-verify approach because since it is easy to generalize for the infinite horizon case. We then guess that the solution assumes the following form:

$$K_1 = \alpha_1^{RE}\left(K_0 + \varepsilon_1\right); K_2 = \alpha_2^{RE}\left(K_1 + \varepsilon_2\right)$$

We then have:

$$\begin{aligned}
K_0 + \varepsilon_1(1-\alpha_1^{RE}) &= \mathbb{E}_1\left[\varepsilon_2 + \alpha_1^{RE}\left(K_0+\varepsilon_1\right) - \alpha_2^{RE}\left(\alpha_1^{RE}\left(K_0+\varepsilon_1\right)+\varepsilon_2\right)\right] \\
\mathbb{E}_1\left[\varepsilon_2 + \alpha_1^{RE}\left(K_0+\varepsilon_1\right) - \alpha_2^{RE}\left(\alpha_1^{RE}\left(K_0+\varepsilon_1\right)+\varepsilon_2\right)\right] &= \mathbb{E}_1\left[\varepsilon_3 + \alpha_2^{RE}\left(\alpha_1^{RE}\left(K_0+\varepsilon_1\right)+\varepsilon_2\right)\right]
\end{aligned}$$

Equating coefficients, we get

$$\alpha_1^{RE} = \frac{2}{3}, \quad \alpha_2^{RE} = \frac{1}{2}$$

It is immediate to verify that the solution is time-consistent. The agent at time 2 solves the problem (3.9). The first order condition at time 2 is:

$$C_2 = \mathbb{E}_2\left[C_3\right],$$

or, equivalently

$$\varepsilon_2 + K_1 - K_2 = \mathbb{E}_2\left[\varepsilon_3 + K_2\right].$$

We obtain:

$$K_2 = \frac{1}{2}\left[K_1 + \varepsilon_2\right] = \alpha_2^{RE}\left[K_1 + \varepsilon_2\right].$$

$\square$

B.3 Proof of Proposition 2 and 3

Proof. For the time 1 problem, we conjecture the planned policy:

$$K_1^{\theta,p} = \alpha_{K_0}^{\theta,p} K_0 + \alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1;\ K_2^{\theta,p} = \alpha_{K_1}^{\theta,p} K_1^{\theta,p} + \alpha_{\varepsilon_2}^{\theta,p}\varepsilon_2.$$

We have two first-order conditions:

$$Y_1 + K_0 - K_1^{\theta,p} = \mathbb{E}_1^{\theta}\left[Y_2 + K_1^{\theta,p} - K_2^{\theta,p}\right] \tag{1}$$

$$\mathbb{E}_1^{\theta}\left[Y_2 + K_1^{\theta,p} - K_2^{\theta,p}\right] = \mathbb{E}_1^{\theta}\left[Y_3 + K_2^{\theta,p}\right] \tag{2}$$

We first solve for the planned policy for period 2 by plugging in the conjecture into (2):

$$\begin{aligned}
&\mathbb{E}_1^{\theta}\left[K_1^{\theta,p} - \left(\alpha_{K_1}^{\theta,p} K_1^{\theta,p} + \alpha_{\varepsilon_2}^{\theta,p}\varepsilon_2\right)\right] = \mathbb{E}_1^{\theta}\left[\alpha_{K_1}^{\theta,p} K_1^{\theta,p} + \alpha_{\varepsilon_2}^{\theta,p}\varepsilon_2\right]\\
&\mathbb{E}_1^{\theta}\left[\left(1 - 2\alpha_{K_1}^{\theta,p}\right) K_1^{\theta,p} - 2\alpha_{\varepsilon_2}^{\theta,p}\varepsilon_2\right] = 0
\end{aligned}$$

Then:

$$\begin{aligned}
(1+\theta)\left(1 - 2\alpha_{K_1}^{\theta,p}\right) K_1^{\theta,p} - \theta\mathbb{E}_0\left[\left(1 - 2\alpha_{K_1}^{\theta,p}\right) K_1^{\theta,p}\right] &= 0\\
\left(1 - 2\alpha_{K_1}^{\theta,p}\right)\left[(1+\theta)\,\alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1 + \alpha_{K_0}^{\theta,p} K_0\right] &= 0
\end{aligned}$$

Hence we have:

$$\alpha_{K_1}^{\theta,p} = \frac{1}{2}$$

Endowed with the contingent plan for time 2, we can then solve for the time 1 problem:

$$\begin{aligned}
Y_1 + K_0 - K_1^{\theta,p} &= \mathbb{E}_1^{\theta}\left[Y_2 + K_1^{\theta,p} - K_2^{\theta,p}\right]\\
\varepsilon_1 + K_0 - K_1^{\theta,p} &= (1+\theta)\,\mathbb{E}_1\left[K_1^{\theta,p} - \frac{1}{2}K_1^{\theta,p}\right] - \theta\mathbb{E}_0\left[K_1^{\theta,p} - \frac{1}{2}K_1^{\theta,p}\right]\\
\varepsilon_1 + K_0 - K_1^{\theta,p} &= (1+\theta)\,\frac{1}{2}\left[\alpha_{K_0}^{\theta,p} K_0 + \alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1\right] - \theta\frac{1}{2}\alpha_{K_0}^{\theta,p} K_0
\end{aligned}$$

We get:

$$K_1^{\theta,p} = -\left(1+\theta\right)\frac{1}{2}\left[\alpha_{K_0}^{\theta,p}K_0 + \alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1\right] + \theta\frac{1}{2}\alpha_{K_0}^{\theta,p}K_0 + \varepsilon_1 + K_0$$
$$K_1^{\theta,p} = -\frac{1}{2}\left[\left(1+\theta\right)\alpha_{\varepsilon_1}^{\theta,p} - 2\right]\varepsilon_1 + \left[1 - \frac{1}{2}\alpha_{K_0}^{\theta,p}\right]K_0$$

Matching coefficients:

$$\alpha_{K_0}^{\theta,p} = \frac{2}{3}, \quad \alpha_{\varepsilon_1}^{\theta,p} = \frac{2}{3+\theta}$$

Note that when $J = 1$, there is no contingent plan formed at time 1 on how to react to ε_2, given that ε_2 does not impact utility at time 1. It is then immediate to verify that the plan is time consistent. The agent at time 2 inherits the capital K_1^θ and solves (3.11). The first order condition at time 2 is:

$$C_2^\theta = \mathbb{E}_2^\theta\left[C_3^\theta\right].$$

We conjecture the solution $K_2^\theta(K_1^\theta, \varepsilon_2) = \alpha_{K_1}^\theta K_1^\theta + \alpha_{\varepsilon_2}^\theta\varepsilon_2$. Then:

$$Y_2 + K_1^\theta - K_2^\theta(K_1^\theta, \varepsilon_2) = \mathbb{E}_2\left[Y_3 + K_2^\theta(K_1^\theta, \varepsilon_2)\right] + \theta\left[\mathbb{E}_2\left(Y_3 + K_2^\theta(K_1^\theta, \varepsilon_2)\right) - \mathbb{E}_1\left(Y_3 + K_2^\theta(K_1^\theta, \varepsilon_2)\right)\right]$$
$$\varepsilon_2 + K_1^\theta - K_2^\theta(K_1^\theta, \varepsilon_2) = \alpha_{K_1}^\theta K_1^\theta + \alpha_{\varepsilon_2}^\theta\varepsilon_2 + \theta\alpha_{\varepsilon_2}^\theta\varepsilon_2$$

We have:

$$K_2^\theta(K_1^\theta, \varepsilon_2) = (1 - \alpha_{K_1}^\theta)K_1^\theta + (1 - \alpha_{\varepsilon_2}^\theta - \theta\alpha_{\varepsilon_2}^\theta)\varepsilon_2$$

Matching coefficients, we obtain:

$$\alpha_{K_1}^\theta = \frac{1}{2} = \alpha_{K_1}^{\theta,p}, \quad \alpha_{\varepsilon_2}^\theta = \frac{1}{2+\theta}$$

□

B.4 Proof of Proposition 4

Proof. We first solve the planning problem at time 1. We conjecture the solution:

$$K_1^{\theta,p} = \alpha_{\mathbb{N}_{-1,0}[K_0]}^{\theta,p}\mathbb{N}_{-1,0}\left[K_0\right] + \alpha_{K_0}^{\theta,p}K_0 + \alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1;\ K_2^{\theta,p} = \alpha_{K_1}^{\theta,p}K_1^{\theta,p} + \alpha_{\varepsilon_2}^{\theta,p}\varepsilon_2.$$

where $\mathbb{N}_{-1,0}\left[K_0\right] \equiv K_0 - \mathbb{E}_{-1}\left[K_0\right]$ represents the surprise in the stock of capital with respect to the expectations formed in the past. Note that $\mathbb{E}_{-1}\left[\mathbb{N}_{-1,0}\left[K_0\right]\right] = 0$. We have two first-order conditions:

$$Y_1 + K_0 - K_1^{\theta,p} = \mathbb{E}_1^\theta\left[Y_2 + K_1^{\theta,p} - K_2^{\theta,p}\right] \tag{3}$$
$$\mathbb{E}_1^\theta\left[Y_2 + K_1^{\theta,p} - K_2^{\theta,p}\right] = \mathbb{E}_1^\theta\left[Y_3 + K_2^{\theta,p}\right] \tag{4}$$

We first solve for the planned policy for period 2 by plugging in the conjecture into (4):

$$\mathbb{E}_1^\theta\left[K_1^{\theta,p} - K_2^{\theta,p}\right] = \mathbb{E}_1^\theta\left[K_2^{\theta,p}\right]$$
$$\mathbb{E}_1^\theta\left[\left(1 - 2\alpha_{K_1}^{\theta,p}\right)K_1^{\theta,p} - 2\alpha_{\varepsilon_2}^{\theta,p}\varepsilon_2\right] = 0$$

Then:

$$\left(1+\theta\right)\left(1-2\alpha_{K_1}^{\theta,p}\right)K_1^{\theta,p}-\theta\mathbb{E}_{-1}\left[\left(1-2\alpha_{K_1}^{\theta,p}\right)K_1^{\theta,p}\right]=0$$
$$\left(1-2\alpha_{K_1}^{\theta,p}\right)\left[\left(1+\theta\right)\left(\alpha_{K_0}^{\theta,p}K_0+\alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1\right)-\theta\mathbb{E}_{-1}\alpha_{K_0}^{\theta,p}K_0\right]=0$$

Hence we have:

$$\alpha_{K_1}^{\theta,p}=\frac{1}{2}$$

Note that the *planned* solution for time 2 is identical to the case in which $J=1$. This is because distant memory does not affect how the agent evaluates the trade-off between consumption at time 2 and consumption at time 3 from the point of view of time 1.

Endowed with the contingent plan for time 2, we can then solve for the time 1 problem:

$$\varepsilon_1+K_0-K_1^{\theta,p}=\left(1+\theta\right)\mathbb{E}_1\left[K_1^{\theta,p}-\frac{1}{2}K_1^{\theta,p}\right]-\theta\mathbb{E}_{-1}\left[K_1^{\theta,p}-\frac{1}{2}K_1^{\theta,p}\right]$$
$$\varepsilon_1+K_0-K_1^{\theta,p}=\left(1+\theta\right)\frac{1}{2}\left[\alpha_{\mathbb{N}_{-1,0}[K_0]}^{\theta,p}\mathbb{N}_{-1,0}\left[K_0\right]+\alpha_{K_0}^{\theta,p}K_0+\alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1\right]-\theta\frac{1}{2}\alpha_{K_0}^{\theta,p}\mathbb{E}_{-1}\left[K_0\right]$$
$$\varepsilon_1+K_0-K_1^{\theta,p}=\left(1+\theta\right)\frac{1}{2}\left[\alpha_{\mathbb{N}_{-1,0}[K_0]}^{\theta,p}\mathbb{N}_{-1,0}\left[K_0\right]+\alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1\right]+\frac{1}{2}\alpha_{K_0}^{\theta,p}K_0+\theta\frac{1}{2}\alpha_{K_0}^{\theta,p}\mathbb{N}_{-1,0}\left[K_0\right]$$
$$\varepsilon_1+K_0-K_1^{\theta,p}=\frac{1}{2}\left[\left(1+\theta\right)\alpha_{\mathbb{N}_{-1,0}[K_0]}^{\theta,p}+\theta\alpha_{K_0}^{\theta,p}\right]\mathbb{N}_{-1,0}\left[K_0\right]+\left(1+\theta\right)\frac{1}{2}\alpha_{\varepsilon_1}^{\theta,p}\varepsilon_1+\frac{1}{2}\alpha_{K_0}^{\theta,p}K_0$$

Then:

$$K_1^{\theta,p}=-\frac{1}{2}\left[\left(1+\theta\right)\alpha_{\mathbb{N}_{-1,0}[K_0]}^{\theta,p}+\theta\alpha_{K_0}^{\theta,p}\right]\mathbb{N}_{-1,0}\left[K_0\right]+\left[1-\frac{1}{2}\alpha_{K_0}^{\theta,p}\right]K_0-\frac{1}{2}\left[\left(1+\theta\right)\alpha_{\varepsilon_1}^{\theta,p}-2\right]\varepsilon_1$$

Hence we have

$$\alpha_{K_0}^{\theta,p}=\frac{2}{3},\quad \alpha_{\varepsilon_1}^{\theta,p}=\frac{2}{3+\theta}$$

and

$$\alpha_{\mathbb{N}_{-1,0}[K_0]}^{\theta,p}=-\frac{\theta}{3+\theta}\alpha_{K_0}^{\theta,p}=-\frac{2\theta}{3\left(3+\theta\right)}$$

Note that even when $J>1$, there is no contingent plan formed at time 1 on how to react to ε_2, given that this does not impact utility at time 1. However, with respect to the case of recent memory, now we have an additional state variable that depends on the news component of the inherited capital, $\mathbb{N}_{-1,0}\left[K_0\right]$. For a given level of inherited capital K_0, the larger the surprise, the lower the amount saved at time 1. This distortion increases with θ.

We now verify that the plan made at time 1 for time 2 is time inconsistent. The agent at time 2 inherits the capital K_1^{θ} and solves (3.11). The first order condition at time 2 is:

$$C_2^{\theta}=\mathbb{E}_2^{\theta}\left[C_3^{\theta}\right].$$

We conjecture the solution $K_2^\theta = \alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]} \mathbb{N}_{0,1}\left[K_1^\theta\right] + \alpha^\theta_{K_1} K_1^\theta + \alpha^\theta_{\varepsilon_2}\varepsilon_2$. Then:

$$\begin{aligned}
Y_2 + K_1^\theta - K_2^\theta &= \mathbb{E}_2\left[Y_3 + K_2^\theta\right] + \theta\left[\mathbb{E}_2\left[Y_3 + K_2^\theta\right] - \mathbb{E}_0\left[Y_3 + K_2^\theta\right]\right] \\
\varepsilon_2 + K_1^\theta - K_2^\theta &= \mathbb{E}_2\left[\alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]} \mathbb{N}_{0,1}\left[K_1^\theta\right] + \alpha^\theta_{K_1} K_1^\theta + \alpha^\theta_{\varepsilon_2}\varepsilon_2\right] \\
&\quad + \theta\left[\begin{array}{c} \mathbb{E}_2\left[\alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]} \mathbb{N}_{0,1}\left[K_1^\theta\right] + \alpha^\theta_{K_1} K_1^\theta + \alpha^\theta_{\varepsilon_2}\varepsilon_2\right] \\ -\mathbb{E}_0\left[\alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]} \mathbb{N}_{0,1}\left[K_1^\theta\right] + \alpha^\theta_{K_1} K_1^\theta + \alpha^\theta_{\varepsilon_2}\varepsilon_2\right] \end{array}\right]
\end{aligned}$$

We get:

$$\begin{aligned}
\varepsilon_2 + K_1^\theta - K_2^\theta &= \alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^{\theta,p}\right]} \mathbb{N}_{0,1}\left[K_1^{\theta,p}\right] + \alpha^\theta_{K_1} K_1^{\theta,p} + \alpha^\theta_{\varepsilon_2}\varepsilon_2 \\
&\quad + \theta\left[\alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^{\theta,p}\right]} \mathbb{N}_{0,1}\left[K_1^{\theta,p}\right] + \alpha^\theta_{K_1} K_1^{\theta,p} + \alpha^\theta_{\varepsilon_2}\varepsilon_2 - \mathbb{E}_0 \alpha^\theta_{K_1} K_1^{\theta,p}\right]
\end{aligned}$$

Then:

$$\varepsilon_2 + K_1^\theta - K_2^\theta = \left[(1+\theta)\,\alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]} + \theta\alpha^\theta_{K_1}\right] \mathbb{N}_{0,1}\left[K_1^\theta\right] + \alpha^\theta_{K_1} K_1^\theta + (1+\theta)\,\alpha^\theta_{\varepsilon_2}\varepsilon_2$$

Rearrange:

$$K_2^\theta = -\left[(1+\theta)\,\alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]} + \theta\alpha^\theta_{K_1}\right] \mathbb{N}_{0,1}\left[K_1^\theta\right] + \left(1 - \alpha^\theta_{K_1}\right) K_1^\theta - \left[(1+\theta)\,\alpha^\theta_{\varepsilon_2} - 1\right]\varepsilon_2$$

Matching coefficients, we obtain:

$$\alpha^\theta_{K_1} = \frac{1}{2}, \quad \alpha^\theta_{\varepsilon_2} = \frac{1}{2+\theta}$$

and

$$\alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]} = -\frac{\theta}{2+\theta}\alpha^\theta_{K_1} = -\frac{\theta}{2\,(2+\theta)}.$$

The revised time 2 policy can then be rewritten as

$$K_2^\theta = \frac{\theta}{2(2+\theta)}\mathbb{E}_0\left[K_1^\theta\right] + \frac{1}{2+\theta}K_1^\theta + \frac{1}{2+\theta}\varepsilon_2,$$

and so the coefficient on K_1^θ is not equal to that of the time 1 plan (which is 0.5). □

The time inconsistency arises because of the information content of $K_1^{\theta,p}$ with respect to the capital expected at time zero. Between when reference expectations were formed, at time 0, and when a new decision is made, at time 2, an income shock occurred and agents reacted to the shock. As a result, capital is not what the agent expected it to be. Agents do not take into account this surprise in capital when they solve the planning problem at time 1.

B.5 Proof of Proposition 5

Proof. To obtain the time 1 policy function, we consider the conjecture

$$K_1^\theta = \alpha^\theta_{\mathbb{N}_{-1,0}[K_0]} \mathbb{N}_{-1,0}\left[K_0\right] + \alpha^\theta_{K_0} K_0 + \alpha^\theta_{\varepsilon_1}\varepsilon_1.$$

The time 1 trade-off is given by

$$C_1^\theta = \mathbb{E}_1^\theta\left[C_2^{RE}\right].$$

The right hand side equals

$$\begin{aligned}
\mathbb{E}_1^\theta\left[C_2^{RE}\right] &= (1+\theta)\mathbb{E}_1\left[Y_2 + K_1^\theta - K_2^{RE}\right] - \theta\mathbb{E}_{-1}\left[Y_2 + K_1^{RE} - K_2^{RE}\right]\\
&= (1+\theta)\mathbb{E}_1\left[\overline{Y} + \varepsilon_2(1-\alpha^{RE}_{\varepsilon_2}) + K_1^\theta(1-\alpha^{RE}_{K_1})\right] - \theta\mathbb{E}_{-1}\left[\overline{Y} + \varepsilon_2(1-\alpha^{RE}_{\varepsilon_2}) + K_1^{RE}(1-\alpha^{RE}_{K_1})\right]\\
&= \overline{Y} + (1-\alpha^{RE}_{K_1})\left[(1+\theta)K_1^\theta - \theta\mathbb{E}_{-1}\left[K_1^{RE}\right]\right]\\
&= \overline{Y} + \frac{1}{2}\left[(1+\theta)K_1^\theta - \frac{2}{3}\theta\mathbb{E}_{-1}\left[K_0\right]\right]
\end{aligned}$$

where we have substituted in the RE policy $K_2^{RE} = \alpha^{RE}_{K_1}K_1 + \alpha^{RE}_{\varepsilon_2}\varepsilon_2$ in the second line and substituted in $\alpha^{RE}_{K_1} = 1/2$ and $\alpha^{RE}_{K_0} = 2/3$ in the fourth line. Connecting this with the left hand side, we have

$$\varepsilon_1 + K_0 - K_1^\theta = \frac{1}{2}\left[(1+\theta)K_1^\theta - \frac{2}{3}\theta\mathbb{E}_{-1}\left[K_0\right]\right].$$

Plugging in the conjectured solution $K_1^\theta = \alpha^\theta_{\mathbb{N}_{-1,0}[K_0]}\mathbb{N}_{-1,0}\left[K_0\right] + \alpha^\theta_{K_0}K_0 + \alpha^\theta_{\varepsilon_1}\varepsilon_1$ and equating coefficients give us

$$\alpha^\theta_{\mathbb{N}_{-1,0}[K_0]} = -\frac{2\theta}{3(3+\theta)}, \quad \alpha^\theta_{K_0} = \frac{2}{3}, \quad \alpha^\theta_{\varepsilon_1} = \frac{2}{3+\theta}.$$

To obtain the time 2 policy function, we consider the conjecture

$$K_2^\theta = \alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]}\mathbb{N}_{0,1}\left[K_1^\theta\right] + \alpha^\theta_{K_1}K_1^\theta + \alpha^\theta_{\varepsilon_2}\varepsilon_2.$$

The time 2 trade-off is given by

$$C_2^\theta = \mathbb{E}_2^\theta\left[C_3^{RE}\right].$$

The right hand side equals

$$\begin{aligned}
\mathbb{E}_2^\theta\left[C_3^{RE}\right] &= (1+\theta)\mathbb{E}_2\left[Y_3 + K_2^\theta\right] - \theta\mathbb{E}_0\left[Y_3 + K_2^{RE}\right]\\
&= \overline{Y} + (1+\theta)K_2^\theta - \theta\mathbb{E}_0\left[K_2^{RE}\right]\\
&= \overline{Y} + (1+\theta)K_2^\theta - \frac{1}{2}\theta\mathbb{E}_0\left[K_1^{RE}\right]
\end{aligned}$$

where we substituted in $\alpha^{RE}_{K_1} = 1/2$. Connecting this with the left hand side, we have

$$\varepsilon_2 + K_1 - K_2^\theta = (1+\theta)K_2^\theta - \frac{1}{2}\theta\mathbb{E}_0 K_1^{RE}.$$

Plugging in the conjectured solution $K_2^\theta = \alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]}\mathbb{N}_{0,1}\left[K_1^\theta\right] + \alpha^\theta_{K_1}K_1^\theta + \alpha^\theta_{\varepsilon_2}\varepsilon_2$ and equating coefficients give us

$$\alpha^\theta_{\mathbb{N}_{0,1}\left[K_1^\theta\right]} = -\frac{\theta}{2(2+\theta)}, \quad \alpha^\theta_{K_1} = \frac{1}{2}, \quad \alpha^\theta_{\varepsilon_2} = \frac{1}{2+\theta}.$$

□

B.6 Proof of Proposition B1

The proposition below considers naïveté and sophistication under $J = 1$. The naïveté problem was described in the main text. Entering period 2, the sophisticated agent's problem is

$$\max_{K_2^\theta} \left[u(C_2^\theta) + \mathbb{E}_2^\theta u(C_3^\theta) \right] \tag{5}$$

where now

$$C_2^\theta = Y_2 + K_1^\theta - K_2^\theta ; C_3^\theta = Y_3 + K_2^\theta - K_3^\theta . \tag{6}$$

Sophistication means that at time 1 the agent understands that her future action is dictated by equation (5) (as well as $K_3^\theta = 0$). Thus, the sophisticated agent solves

$$\max_{K_1^\theta} \left\{ u(C_1^\theta) + \mathbb{E}_1^\theta \left[u(C_2^\theta) + u(C_3^\theta) \right] \right\}, \tag{7}$$

where $C_1^\theta = Y_1 + K_0 - K_1^\theta$, while C_2^θ and C_3^θ are determined as in (6). We assume that the comparison group for K_2^θ is $\mathbb{E}_{2-J} K_2^\theta$, i.e. the conditional expectation of the DE savings choice at time 2 made J periods ago by the former sophisticated self, under the true density.

Proposition B1. *When $J = 1$, the naïveté and sophistication policy functions are the same and recover the DE optimal choices based on time-consistency.*

Proof. **Policies under naïveté.** Conjecture

$$K_1^\theta = \alpha_{K_0}^\theta K_0 + \alpha_{\varepsilon_1}^\theta \varepsilon_1; \quad K_2^\theta = \alpha_{K_1}^\theta K_1^\theta + \alpha_{\varepsilon_2}^\theta \varepsilon_2 .$$

The time 2 trade-off is given by

$$C_2^\theta = \mathbb{E}_2^\theta \left[C_3^{RE} \right]$$

The right hand side equals

$$\begin{aligned} \mathbb{E}_2^\theta \left[C_3^{RE} \right] &= (1+\theta) \mathbb{E}_2 \left[Y_3 + K_2^\theta \right] - \theta \mathbb{E}_1 \left[Y_3 + K_2^{RE} \right] \\ &= \overline{Y} + (1+\theta) K_2^\theta - \theta \mathbb{E}_1 \left[K_2^{RE} \right] \\ &= \overline{Y} + (1+\theta) K_2^\theta - \frac{1}{2} \theta K_1 , \end{aligned}$$

where we substituted in $\alpha_{K_1}^{RE} = 1/2$ in the third line. Connecting this with the left hand side, we have

$$\varepsilon_2 + K_1^\theta - K_2^\theta = (1+\theta) K_2^\theta - \frac{1}{2} \theta K_1^\theta .$$

Plugging in the conjectured solution $K_2^\theta = \alpha_{K_1}^\theta K_1^\theta + \alpha_{\varepsilon_2}^\theta \varepsilon_2$ and equating coefficients give us

$$\alpha_{K_1}^\theta = \frac{1}{2}, \quad \alpha_{\varepsilon_2}^\theta = \frac{1}{2+\theta} .$$

By Lemma 2 the time 1 trade-off is given by

$$C_1^\theta = \mathbb{E}_1^\theta \left[C_2^{RE} \right] .$$

The right hand side equals

$$\begin{aligned}\mathbb{E}_1^\theta\left[C_2^{RE}\right] &= (1+\theta)\mathbb{E}_1\left[Y_2 + K_1^\theta - K_2^{RE}\right] - \theta\mathbb{E}_0\left[Y_2 + K_1^{RE} - K_2^{RE}\right] \\ &= (1+\theta)\mathbb{E}_1\left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^{RE}) + K_1^\theta(1-\alpha_{K_1}^{RE})\right] - \theta\mathbb{E}_0\left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^{RE}) + K_1^{RE}(1-\alpha_{K_1}^{RE})\right] \\ &= \overline{Y} + (1-\alpha_{K_1}^{RE})\left[(1+\theta)K_1^\theta - \theta\mathbb{E}_0\left[K_1^{RE}\right]\right] \\ &= \overline{Y} + \frac{1}{2}\left[(1+\theta)K_1^\theta - \frac{2}{3}\theta K_0\right]\end{aligned}$$

where we have substituted in the RE policy $K_2^{RE} = \alpha_{K_1}^{RE}K_1^\theta + \alpha_{\varepsilon_2}^{RE}\varepsilon_2$ in the second line and substituted in $\alpha_{K_1}^{RE} = 1/2$ and $\alpha_{K_0}^{RE} = 2/3$ in the fourth line. Connecting this with the left hand side, we have

$$\varepsilon_1 + K_0 - K_1^\theta = \frac{1}{2}\left[(1+\theta)K_1^\theta - \frac{2}{3}\theta K_0\right].$$

Plugging in the conjectured solution $K_1^\theta = \alpha_{K_0}^\theta K_0 + \alpha_{\varepsilon_1}^\theta\varepsilon_1$ and equating coefficients give us

$$\alpha_{K_0}^\theta = \frac{2}{3}, \quad \alpha_{\varepsilon_1}^\theta = \frac{2}{3+\theta}.$$

Policies under sophistication. Conjecture

$$K_1^\theta = \alpha_{K_0}^{\theta,s}K_0 + \alpha_{\varepsilon_1}^{\theta,s}\varepsilon_1; \quad K_2^\theta = \alpha_{K_1}^{\theta,s}K_1^\theta + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2.$$

The time 2 trade-off is given by

$$C_2^\theta = \mathbb{E}_2^\theta\left[C_3^\theta\right]$$

The right hand side equals

$$\begin{aligned}\mathbb{E}_2^\theta C_3^\theta &= (1+\theta)\mathbb{E}_2\left[Y_3 + K_2^\theta\right] - \theta\mathbb{E}_1\left[Y_3 + K_2^\theta\right] \\ &= \overline{Y} + (1+\theta)K_2^\theta - \theta\mathbb{E}_1\left[K_2^\theta\right] \\ &= \overline{Y} + (1+\theta)K_2^\theta - \theta\alpha_{K_1}^{\theta,s}K_1^\theta.\end{aligned}$$

Connecting this with the left hand side, we have

$$\varepsilon_2 + K_1^\theta - K_2^\theta = (1+\theta)K_2^\theta - \theta\alpha_{K_1}^{\theta,s}K_1^\theta.$$

Plugging in the conjectured solution $K_2^\theta = \alpha_{K_1}^{\theta,s}K_1^\theta + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2$ and equating coefficients give us

$$\alpha_{K_1}^{\theta,s} = \frac{1}{2}, \quad \alpha_{\varepsilon_2}^{\theta,s} = \frac{1}{2+\theta}.$$

The time 1 trade-off is given by

$$C_1^\theta = \mathbb{E}_1^\theta\left[C_2^\theta + \frac{\partial K_2^\theta}{\partial K_1^\theta}\left(C_3^\theta - C_2^\theta\right)\right].$$

but the indirect effect of current choice captured by the last term disappears under $J = 1$, so

$$C_1^\theta = \mathbb{E}_1^\theta\left[C_2^\theta\right].$$

The right hand side equals

$$
\begin{aligned}
\mathbb{E}_1^\theta\left[C_2^\theta\right] &= (1+\theta)\mathbb{E}_1\left[Y_2 + K_1^\theta - K_2^\theta\right] - \theta\mathbb{E}_0\left[Y_2 + K_1^\theta - K_2^\theta\right] \\
&= (1+\theta)\mathbb{E}_1\left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^{\theta,s}) + K_1^\theta(1-\alpha_{K_1}^{\theta,s})\right] - \theta\mathbb{E}_0\left[\overline{Y} + \varepsilon_2(1-\alpha_{\varepsilon_2}^{\theta,s}) + K_1^\theta(1-\alpha_{K_1}^{\theta,s})\right] \\
&= \overline{Y} + (1-\alpha_{K_1}^{\theta,s})\left[(1+\theta)K_1^\theta - \theta\mathbb{E}_0\left[K_1^\theta\right]\right] \\
&= \overline{Y} + \frac{1}{2}\left[(1+\theta)K_1^\theta - \alpha_{K_0}^{\theta,s}\theta K_0\right]
\end{aligned}
$$

where we have substituted in the DE policy $K_2^\theta = \alpha_{K_1}^{\theta,s}K_1 + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2$ in the second line and substituted in $\alpha_{K_1}^{\theta,s} = 1/2$ in the fourth line. Connecting this with the left hand side, we have

$$
\varepsilon_1 + K_0 - K_1^{\theta,s} = \frac{1}{2}\left[(1+\theta)K_1^\theta - \alpha_{K_0}^{\theta,s}\theta K_0\right].
$$

Plugging in the conjectured solution $K_1^\theta = \alpha_{K_0}^{\theta,s}K_0 + \alpha_{\varepsilon_1}^{\theta,s}\varepsilon_1$ and equating coefficients give us

$$
\alpha_{K_0}^{\theta,s} = \frac{2}{3}, \quad \alpha_{\varepsilon_1}^{\theta,s} = \frac{2}{3+\theta}.
$$

□

B.7 Proof of Proposition B2

The Proposition below considers the solution to the three-period model under distant memory ($J = 2$) and sophistication.

Proposition B2. *When $J = 2$, under sophistication, the time 2 policy function is given by*

$$
K_2^\theta = \alpha_{\mathbb{E}_0[K_1]}^{\theta,s}\mathbb{E}_0[K_1^\theta] + \alpha_{K_1}^{\theta,s}K_1^\theta + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2,
$$

where the coefficients are identical to the naïveté case (modified to expressed in terms of coefficients $\alpha_{\mathbb{E}_0[K_1]}^\theta$, $\alpha_{K_1}^\theta$ and $\alpha_{\varepsilon_2}^\theta$) and are given by

$$
\alpha_{\mathbb{E}_0[K_1]}^{\theta,s} = \frac{\theta}{2(2+\theta)}; \; \alpha_{K_1}^{\theta,s} = \frac{1}{2+\theta}; \; \alpha_{\varepsilon_2}^{\theta,s} = \frac{1}{2+\theta}.
$$

The time 1 policy function is given by

$$
K_1^\theta = \alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s}\mathbb{E}_{-1}\left[K_0\right] + \alpha_{K_0}^{\theta,s}K_0 + \alpha_{\varepsilon_1}^{\theta,s}\varepsilon_1,
$$

which compared to the naïveté policy function in Proposition 5 (modified to expressed in terms of coefficients $\alpha_{\mathbb{E}_{-1}[K_0]}^\theta$, $\alpha_{K_0}^\theta$ and $\alpha_{\varepsilon_1}^\theta$) is characterized by the following properties (1) $\alpha_{\varepsilon_1}^{\theta,s} < \alpha_{\varepsilon_1}^\theta$; (2) $\alpha_{K_0}^{\theta,s} < \alpha_{K_0}^\theta$ if $\theta < 1$, and $\alpha_{K_0}^{\theta,s} > \alpha_{K_0}^\theta$ if $\theta > 1$; (3), $\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s} > \alpha_{\mathbb{E}_{-1}[K_0]}^\theta$ if $\theta < 1$, and $\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s} < \alpha_{\mathbb{E}_{-1}[K_0]}^\theta$ if $\theta > 1$.

Proof. For the time 2 policy, consider the conjecture

$$
K_2^\theta = \alpha_{\mathbb{E}_0[K_1]}^{\theta,s}\mathbb{E}_0\left[K_1^\theta\right] + \alpha_{K_1}^\theta K_1^\theta + \alpha_{\varepsilon_2}^\theta\varepsilon_2.
$$

The time 2 trade-off is given by

$$C_2^\theta = \mathbb{E}_2^\theta \left[C_3^\theta\right]$$

The right hand side equals

$$\begin{aligned}
\mathbb{E}_2^\theta \left[C_3^\theta\right] &= (1+\theta)\mathbb{E}_2 \left[Y_3 + K_2^\theta\right] - \theta\mathbb{E}_0 \left[Y_3 + K_2^\theta\right] \\
&= \overline{Y} + K_2^\theta + \theta \left[K_2^\theta - \mathbb{E}_0 \left[K_2^\theta\right]\right] \\
&= \overline{Y} + \alpha_{\mathbb{E}_0[K_1]}^{\theta,s}\mathbb{E}_0 \left[K_1^\theta\right] + \alpha_{K_1}^{\theta,s} K_1^\theta + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2 + \theta \left[\alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2 + \alpha_{K_1}^{\theta,s}\left(K_1^\theta - \mathbb{E}_0\left[K_1^\theta\right]\right)\right].
\end{aligned}$$

Connecting this with the left hand side, we have

$$\varepsilon_2 + K_1^\theta - K_2^\theta = \alpha_{\mathbb{E}_0[K_1]}^{\theta,s}\mathbb{E}_0 \left[K_1^\theta\right] + \alpha_{K_1}^{\theta,s} K_1^\theta + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2 + \theta \left[\alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2 + \alpha_{K_1}^{\theta,s}\left(K_1^\theta - \mathbb{E}_0\left[K_1^\theta\right]\right)\right].$$

Plugging in the conjectured solution $K_2^\theta = \alpha_{\mathbb{E}_0[K_1]}^{\theta,s}\mathbb{E}_0 \left[K_1^\theta\right] + \alpha_{K_1}^{\theta,s} K_1^\theta + \alpha_{\varepsilon_2}^{\theta,s}\varepsilon_2$ and equating coefficients give us

$$\alpha_{\mathbb{E}_0[K_1]}^{\theta,s} = \frac{1}{2(2+\theta)}, \quad \alpha_{K_1}^{\theta,s} = \frac{1}{2+\theta}, \quad \alpha_{\varepsilon_2}^{\theta,s} = \frac{1}{2+\theta}.$$

For the time 1 policy function, conjecture

$$K_1^\theta = \alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s}\mathbb{E}_{-1}\left[K_0\right] + \alpha_{K_0}^{\theta,s} K_0 + \alpha_{\varepsilon_1}^{\theta,s}\varepsilon_1.$$

The time 1 tradeoff is given by

$$C_1^\theta = \mathbb{E}_1^\theta \left[C_2^\theta + \alpha_{K_1}^{\theta,s}(\mathbb{E}_2\left[C_3^\theta\right] - C_2^\theta)\right], \tag{8}$$

where the term $\alpha_{K_1}^{\theta,s}(\mathbb{E}_2 C_3^\theta - C_2^\theta)$ captures the fact that the sophisticated agent internalizes the fact that the current choice affects the future tradeoff. This term is zero when $J = 1$ because in that case the plan is time consistent. The right hand side equals

$$\begin{aligned}
&\mathbb{E}_1^\theta \left[C_2^\theta + \alpha_{K_1}^{\theta,s}(\mathbb{E}_2 C_3^\theta - C_2^\theta)\right] = (1-\alpha_{K_1}^{\theta,s})\mathbb{E}_1^\theta C_2^\theta + \alpha_{K_1}^{\theta,s}\mathbb{E}_1^\theta C_3^\theta \\
&= (1-\alpha_{K_1}^{\theta,s})\left\{(1+\theta)\mathbb{E}_1\left[Y_2 + K_1^\theta - K_2^\theta\right] - \theta\mathbb{E}_{-1}\left[Y_2 + K_1^\theta - K_2^\theta\right]\right\} \\
&+ \alpha_{K_1}^{\theta,s}\left\{(1+\theta)\mathbb{E}_1\left[Y_3 + K_2^\theta\right] - \theta\mathbb{E}_{-1}\left[Y_3 + K_2^\theta\right]\right\}
\end{aligned}$$

After some algebra, we find that this equals

$$\begin{aligned}
&= \overline{Y} + (1-\alpha_{K_1}^{\theta,s})(1+\theta)\left[\left(1 - \alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s} - \alpha_{K_1}^{\theta,s}\right)\left(\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s}\mathbb{E}_{-1}\left[K_0\right] + \alpha_{K_0}^{\theta,s}K_0\right) + (1-\alpha_{K_1}^{\theta,s})\alpha_{\varepsilon_1}^{\theta,s}\varepsilon_1\right] \\
&- (1-\alpha_{K_1}^{\theta,s})\theta\left(1 - \alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s} - \alpha_{K_1}^{\theta,s}\right)\left(\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s} + \alpha_{K_0}^{\theta,s}\right)\mathbb{E}_{-1}\left[K_0\right] \\
&+ \alpha_{K_1}^{\theta,s}(1+\theta)\left[\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s}\left(\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s}\mathbb{E}_{-1}\left[K_0\right] + \alpha_{K_0}^{\theta,s}K_0\right) + \alpha_{K_1}^{\theta,s}\left(\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s}\mathbb{E}_{-1}\left[K_0\right] + \alpha_{K_0}^{\theta,s}K_0 + \alpha_{\varepsilon_1}^{\theta,s}\varepsilon_1\right)\right] \\
&- \alpha_{K_1}^{\theta,s}\theta\left(\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s} + \alpha_{K_1}^{\theta}\right)\left(\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s} + \alpha_{K_0}^{\theta,s}\right)\mathbb{E}_{-1}\left[K_0\right]
\end{aligned}$$

The left hand side is given by

$$C_1^\theta = \overline{Y} + \varepsilon_1 + K_0 - K_1^\theta.$$

We then connect the left hand side to the right hand side and equate coefficients after substituting in the conjectured solution for K_1^θ. Equating coefficients, we have

$$\begin{aligned}
\alpha_{\varepsilon_1}^{\theta,s} &= \frac{1}{1+(1+\theta)\left[(1-\alpha_{K_1}^{\theta,s})^2+(\alpha_{K_1}^{\theta,s})^2\right]} = \frac{(2+\theta)^2}{(2+\theta)^2+(1+\theta)\left[(1+\theta)^2+1\right]} \\
\alpha_{K_0}^{\theta,s} &= \frac{1}{1+(1+\theta)\left[(1-\alpha_{K_1}^{\theta,s})(1-\alpha_{\mathbb{E}_0[K_1]}^{\theta,s}-\alpha_{K_1}^{\theta,s})+\alpha_{K_1}^{\theta,s}(\alpha_{\mathbb{E}_0[K_1]}^{\theta,s}+\alpha_{K_1}^{\theta,s})\right]} \\
&= \frac{2(2+\theta)^2}{2(2+\theta)^2+(1+\theta)\left[(1+\theta)(1+2\theta)+3\right]} \\
\alpha_{\mathbb{E}_{-1}[K_0]}^{\theta,s} &= \frac{\theta\left[(1-\alpha_{K_1}^{\theta,s})(1-\alpha_{\mathbb{E}_0[K_1]}^{\theta,s}-\alpha_{K_1}^{\theta,s})+\alpha_{K_1}^{\theta,s}(\alpha_{\mathbb{E}_0[K_1]}^{\theta,s}+\alpha_{K_1}^{\theta,s})\right]}{1+(1-\alpha_{K_1}^{\theta,s})(1-\alpha_{\mathbb{E}_0[K_1]}^{\theta,s}-\alpha_{K_1}^{\theta,s})+\alpha_{K_1}^{\theta,s}(\alpha_{\mathbb{E}_0[K_1]}^{\theta,s}+\alpha_{K_1}^{\theta,s})}\alpha_{K_0}^{\theta,s} \\
&= \frac{\theta\left[(1+2\theta)(1+\theta)+3\right]}{2(2+\theta)^2+(1+2\theta)(1+\theta)+3}\alpha_{K_0}^{\theta,s}
\end{aligned}$$

which give the specific coefficients in Proposition B2. When we compare this sophistication solution to the naïveté one, we find the patterns stated in Proposition B2.

□

The solution for the sophisticated choice K_2^θ follows the same logic as for naïveté choice, leading to the result in Proposition B2 that the optimal coefficients are the same. The subtle difference here is the comparison group formation. The naïveté solution can leverage the law of motion for K_1^{RE}, so that $\mathbb{E}_0\left[K_1^{RE}\right]$ can be immediately plugged in the determination of time 2 savings. In contrast, the corresponding $\mathbb{E}_0\left[K_1^\theta\right]$ is more difficult to transparently assess because it requires computing a feedback effect between the K_1^θ chosen by the time 1 sophisticated DE agent, which in turn is a function of expectations about K_2^θ.

There are three conceptual forces that affect the coefficients of sophisticated time 1 policy compared to their naïveté case. First, the agent now anticipates that she will over-consume (relative to her naive beliefs) at time 2 out of K_1, as the forecasted response of future savings out of capital entering period 2 is smaller than under naïveté, i.e. $\alpha_{K_1}^{\theta,s} < \alpha_{K_1}^{RE}$. This force alone, coming from the $\mathbb{E}_1^\theta\left[C_2^\theta\right]$ term in (8), leads the agent to consume more today out of ε_1 to achieve consumption smoothing between period 1 and 2. Second, the misalignment of her perceived tradeoffs means that following a positive innovation ε_1, from the viewpoint of current self, the time 2 self will under-consume in period $t = 3$ relative to $t = 2$. This constitutes an indirect effect, i.e. the second term in (8), that leads to more savings. The race between these two forces is dominated here by the former, direct effect, as $\alpha_{K_1}^{\theta,s} < 0.5$, and thus the agent ends up saving less out of ε_1 than under naïveté, i.e. $\alpha_{\varepsilon_1}^{\theta,s} < \alpha_{\varepsilon_1}^{\theta}$. Third, there is the conceptual difference of the comparison groups. With sophistication, the informational state $\mathbb{E}_0\left[K_1^\theta\right]$ (a) matters for the K_2^θ solution in Proposition B2 but also (b) needs to be itself based on K_1^θ, a choice that in turn is affected by $\mathbb{E}_1^\theta\left[K_2^\theta\right]$ in equation (8). The effect of this fixed point consideration is less transparent, as it turns out to amplify or dampen, through a non-monotonic relationship with θ, the optimal responses of sophisticated time 1 savings to K_0 and $\mathbb{E}_{-1}\left[K_0\right]$ compared to the naïveté case.

B.8 Proof of Proposition 6

We first guess and verify the RE solution. Conjecture consumption policy

$$\begin{aligned} C_t^{RE} &= \frac{r}{1+r}\left(K_{t-1}^{RE} + \varepsilon_t + \frac{1+r}{r}\overline{Y}\right) \\ &= \frac{r}{1+r}\left(K_{t-1}^{RE} + \varepsilon_t\right) + \overline{Y} \end{aligned}$$

and the resulting savings

$$\begin{aligned} K_t^{RE} &= (1+r)\left[K_{t-1}^{RE} + \overline{Y} + \varepsilon_t - \frac{r}{1+r}\left(K_{t-1}^{RE} + \varepsilon_t\right) - \overline{Y}\right] \\ &= (1+r)\frac{1}{1+r}\left[K_{t-1}^{RE} + \varepsilon_t\right] = K_{t-1}^{RE} + \varepsilon_t \end{aligned}$$

Check the FOC by plugging in the above conjectures

$$\begin{aligned} C_t^{RE} &= \mathbb{E}_t(C_{t+1}^{RE}) \\ \frac{r}{1+r}\left[K_{t-1}^{RE} + \varepsilon_t\right] + \overline{Y} &= \mathbb{E}_t\left[\frac{r}{1+r}\left[K_t^{RE} + \varepsilon_{t+1}\right] + \overline{Y}\right] \\ \frac{r}{1+r}\left[K_{t-1}^{RE} + \varepsilon_t\right] &= \mathbb{E}_t\left[\frac{r}{1+r}\left[K_{t-1}^{RE} + \varepsilon_t + \varepsilon_{t+1}\right]\right] \\ \frac{r}{1+r}\left[K_{t-1}^{RE} + \varepsilon_t\right] &= \frac{r}{1+r}\left[K_{t-1}^{RE} + \varepsilon_t\right] \end{aligned}$$

so the both sides are indeed equal.

To solve for the DE poicy function, we take the FOC with respect to K_t^θ in (3.27):

$$u'(C_t^\theta) = \mathbb{E}_t^\theta\left[\mathcal{V}'(K_t^\theta)\right],$$

and use the envelope theorem in (3.28):

$$\mathcal{V}'(K_t^\theta) = u'(C_{t+1}^{RE}).$$

Combining the two, we have

$$u'(C_t^\theta) = \mathbb{E}_t^\theta\left(u'(C_{t+1}^{RE})\right).$$

We use the budget constraint to replace C_t^θ in the left hand side and obtain:

$$\begin{aligned} K_{t-1}^\theta + \overline{Y} + \varepsilon_t - \frac{K_t^\theta}{1+r} &= \mathbb{E}_t^\theta\left[\frac{r}{1+r}\left(K_t^\theta + \varepsilon_{t+1}\right) + \overline{Y}\right] \\ K_{t-1}^\theta + \varepsilon_t - \frac{K_t^\theta}{1+r} &= \frac{r}{1+r}\mathbb{E}_t^\theta\left(K_t^\theta + \varepsilon_{t+1}\right) \end{aligned}$$

Applying DE and using the fact that $\mathbb{E}_{t-J}K_t^{RE} = \mathbb{E}_{t-J}\left[K_{t-1}^{RE} + \varepsilon_t\right]$, we have that the expectation on the right hand side is equal to:

$$\mathbb{E}_t^\theta\left[K_t^\theta + \varepsilon_{t+1}\right] = K_t^\theta + \theta\left[K_t^\theta - \mathbb{E}_{t-J}K_t^{RE}\right] = K_t^\theta + \theta\left[K_t^\theta - \mathbb{E}_{t-J}\left(K_{t-1}^{RE}\right)\right]$$

Then

$$K_{t-1}^{\theta}+\varepsilon_t-\frac{K_t^{\theta}}{1+r}=\frac{r}{1+r}\left[K_t^{\theta}+\theta\left[K_t^{\theta}-\mathbb{E}_{t-J}\left(K_{t-1}^{RE}\right)\right]\right]$$

Rearrange:

$$\begin{aligned}(1+r)\,K_{t-1}^{\theta}+r\theta K_{t-1}^{\theta}+(1+r)\,\varepsilon_t &= K_t^{\theta}+rK_t^{\theta}+r\theta K_t^{\theta}+r\theta K_{t-1}^{\theta}-r\theta\mathbb{E}_{t-J}\left(K_{t-1}^{RE}\right)\\ \left[1+r\,(1+\theta)\right]K_{t-1}^{\theta}+(1+r)\,\varepsilon_t &= \left[1+r\,(1+\theta)\right]K_t^{\theta}-r\theta\mathbb{N}_{t-J,t-1}\left(K_{t-1}^{\theta}\right)\end{aligned}$$

Then:

$$K_t^{\theta}=K_{t-1}^{\theta}-\frac{r\theta}{1+r\,(1+\theta)}\mathbb{N}_{t-J,t-1}\left(K_{t-1}^{\theta}\right)+\frac{1+r}{1+r\,(1+\theta)}\varepsilon_t.$$

where $\mathbb{N}_{t-J,t-1}\left(K_{t-1}^{\theta}\right)=K_{t-1}^{\theta}-\mathbb{E}_{t-J}\left(K_{t-1}^{RE}\right)$.

Consistent with the discussion above and the naïveté assumption, we assume that memory is based on the rational expectation solution. This is how the agent perceives capital should have evolved from the point of view of time $t-J$. Thus, we have $\mathbb{E}_{t-J}\left(K_{t-1}^{RE}\right)=K_{t-J}^{\theta}$ and the solution becomes:

$$K_t^{\theta}=K_{t-1}^{\theta}-\frac{r\theta}{1+r\,(1+\theta)}\left[K_{t-1}^{\theta}-K_{t-J}^{\theta}\right]+\frac{1+r}{1+r\,(1+\theta)}\varepsilon_t.$$

We can also rewrite this as:

$$K_t^{\theta}=\frac{1}{1+r\,(1+\theta)}\left[K_{t-1}^{\theta}+r\theta K_{t-J}^{\theta}+(1+r)\,\varepsilon_t\right].$$

C Equilibrium Conditions of the New Keynesian Model

1. Capital Euler equation:

$$\mu_t^{\theta}=\beta\mathbb{E}_t^{\theta}\left[(C_{t+1}^{RE}-bC_t^{\theta})^{-1}(R_{t+1}^{k,RE}u_{t+1}^{RE}-a(u_{t+1}^{RE}))+\mu_{t+1}^{RE}(1-\delta)\right],$$

   where μ_t^{θ} is the Lagrangian multiplier on the capital accumulation equation.

2. Utilization choice:

$$R_t^{k,\theta}=R^k(u_t^{\theta})^{\tau}$$

3. Investment first-order condition:

$$(C_t^{\theta}-bC_{t-1}^{\theta})^{-1}=\mu_t^{\theta}\left\{1-\frac{\kappa}{2}\left(\Delta I_t^{\theta}-\gamma\right)^2-\kappa\left(\Delta I_t^{\theta}-\gamma\right)\Delta I_t^{\theta}\right\}+\beta\mathbb{E}_t^{\theta}\left[\mu_{t+1}^{RE}\kappa\left(\Delta I_{t+1}^{RE}-\gamma\right)\left(\Delta I_{t+1}^{RE}\right)^2\right]$$

4. Investment growth:

$$\Delta I_t^{\theta}=I_t^{\theta}/I_{t-1}^{\theta}$$

5. Consumption Euler equation:

$$Q_t^{\theta}=\frac{\beta R_t^{\theta}}{\Pi}\mathbb{E}_t^{\theta}\left[Q_{t+1}^{RE}\right]$$

6. Definition of Q_t^θ :
$$\frac{Q_t^\theta}{Q_{t-1}^\theta}=\frac{\Pi}{\Pi_t^\theta}\left(\frac{C_t^\theta-bC_{t-1}^\theta}{C_{t-1}^\theta-bC_{t-2}^\theta}\right)^{-1}$$

7. Capital accumulation:
$$K_t^\theta=(1-\delta)K_{t-1}^\theta+\left\{1-\frac{\kappa}{2}\left(\frac{I_t^\theta}{I_{t-1}^\theta}-\gamma\right)^2\right\}I_t^\theta$$

8. Real wage:
$$\widetilde{W}_t^\theta=MC_t^\theta(1-\alpha)\frac{Y_t^\theta}{N_t^\theta}$$
where $\widetilde{W}_t^\theta\equiv W_t^\theta/P_t^\theta$ is the real wage.

9. Capital rental rate:
$$R_t^{k,\theta}=MC_t^\theta\alpha\frac{Y_t^\theta}{K_{t-1}^\theta}$$

10. Production function:
$$Y_t^\theta=(u_t^\theta K_{t-1}^\theta)^\alpha(\gamma^tN_t^\theta)^{1-\alpha}$$

11. Optimal price setting:
$$Q_t^\theta\left\{-\frac{1}{\lambda_f-1}Y_t^\theta+\frac{\lambda_f}{\lambda_f-1}MC_t^\theta Y_t^\theta-\varphi_p\left(\Pi_t^\theta-\Pi\right)\Pi_t^\theta Y_t^\theta\right\}$$
$$+\frac{\beta\varphi_p}{\Pi}\mathbb{E}_t^\theta\left[Q_{t+1}^{RE}\left(\Pi_{t+1}^{RE}-\Pi\right)(\Pi_{t+1}^{RE})^2Y_{t+1}^{RE}\right]=0$$

12. Optimal wage setting:
$$Q_t^\theta\left[\left(-\frac{1}{\lambda_n-1}\right)N_t^\theta+(C_t^\theta-bC_{t-1}^\theta)\left(\frac{\lambda_n}{\lambda_n-1}\right)\left(N_t^\theta\right)^{1+\eta}\frac{1}{\widetilde{W}_t^\theta}-\varphi_w\left(\Pi_{w,t}^\theta-\gamma\Pi\right)\Pi_{w,t}^\theta\right]$$
$$+\frac{\beta\varphi_w}{\Pi}\mathbb{E}_t^\theta\left[Q_{t+1}^{RE}\left(\Pi_{w,t+1}^{RE}-\gamma\Pi\right)(\Pi_{w,t+1}^{RE})^2\right]=0$$

13. Nominal wage inflation:
$$\Pi_{w,t}^\theta=\Pi_t^\theta\frac{\widetilde{W}_t^\theta}{\widetilde{W}_{t-1}^\theta}$$

14. Resource constraint:
$$C_t^\theta+I_t^\theta+\frac{\varphi_p}{2}\left(\Pi_t^\theta-\Pi\right)^2Y_t^\theta+\frac{\varphi_w}{2}\left(\Pi_{w,t}^\theta-\gamma\Pi\right)^2\widetilde{W}_t^\theta+a(u_t^\theta)K_{t-1}^\theta=Y_t^\theta$$

15. GDP:
$$Y_t^{G,\theta}=Y_t^\theta-\frac{\varphi_p}{2}\left(\Pi_t^\theta-\Pi\right)^2Y_t^\theta-\frac{\varphi_w}{2}\left(\Pi_{w,t}^\theta-\gamma\Pi\right)^2\widetilde{W}_t^\theta-a(u_t^\theta)K_{t-1}^\theta$$

16. Taylor rule:

$$\frac{R_t^\theta}{R} = \left(\frac{R_{t-1}^\theta}{R}\right)^{\rho_R} \left\{ \left(\frac{\widetilde{\Pi}_t^\theta}{\Pi}\right)^{\phi_\pi} \left(\frac{Y_t^{G,\theta}}{\gamma Y_{t-1}^{G,\theta}}\right)^{\phi_Y} \right\}^{1-\rho_R} \exp(\varepsilon_t)$$

D Solution Algorithm

We start from a linear RE system

$$\underset{n\times n}{\boldsymbol{\Gamma}_0}\underset{n\times 1}{\mathbf{x}_t^{RE}} = \underset{n\times n}{\boldsymbol{\Gamma}_1}\underset{n\times 1}{\mathbf{x}_{t-1}^{RE}} + \underset{n\times n_s}{\boldsymbol{\Psi}}\underset{n_s\times 1}{\varepsilon_t} + \underset{n\times n_e}{\boldsymbol{\Pi}}\underset{n_e\times 1}{\eta_t^{RE}}$$

where $\mathbf{x}_t^{RE}$, ε_t and η_t^{RE} are vectors of endogenous variables, shocks, and expectation errors, respectively. A recursive law of motion can be obtained, using for example Sims (2000), as:

$$\mathbf{x}_t^{RE} = \mathbf{T}^{RE}\mathbf{x}_{t-1}^{RE} + \mathbf{R}^{RE}\varepsilon_t.$$

Note that the solution can be divided based on the non-expectation $\left(\widetilde{\mathbf{x}}_t^{RE}\right)$ and expectation terms $\left(\mathbb{E}_t\mathbf{y}_{t+1}^{RE}\right)$:

$$\begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_t^{RE}} \\ \underset{n_e\times 1}{\mathbb{E}_t\mathbf{y}_{t+1}^{RE}} \end{bmatrix} = \begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{T}_{11}^{RE}} & \underset{(n-n_e)\times n_e}{\mathbf{T}_{12}^{RE}} \\ \underset{n_e\times(n-n_e)}{\mathbf{T}_{21}^{RE}} & \underset{n_e\times n_e}{\mathbf{T}_{22}^{RE}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_{t-1}^{RE}} \\ \underset{n_e\times 1}{\mathbb{E}_{t-1}\mathbf{y}_t^{RE}} \end{bmatrix} + \begin{bmatrix} \underset{(n-n_e)\times n_s}{\mathbf{R}_1^{RE}} \\ \underset{n_e\times n_s}{\mathbf{R}_2^{RE}} \end{bmatrix} \varepsilon_t$$

where $\mathbf{y}_{t+1}^{RE}$ is a subset of $\widetilde{\mathbf{x}}_{t+1}^{RE}$.

Define:

$$\mathbf{x}_t^\theta = \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_t^\theta} \\ \underset{n_e\times 1}{\left(\mathbb{E}_t\mathbf{y}_{t+1}^{RE}\right)^\theta} \end{bmatrix}$$

Note that $\left(\mathbb{E}_t\mathbf{y}_{t+1}^{RE}\right)^\theta$ denotes the realized value for rational expectations, so it is different from $\mathbb{E}_t^\theta\mathbf{y}_{t+1}^{RE}$. We have:

$$\mathbb{E}_t\mathbf{y}_{t+1}^{RE} = \mathbf{M}\mathbf{T}^{RE}\mathbf{x}_t^\theta = \left(\mathbb{E}_t\mathbf{y}_{t+1}^{RE}\right)^\theta$$

where $\mathbf{M}$ is a matrix that extract the relevant elements from $\mathbf{T}^{RE}\mathbf{x}_t^\theta$. Note that the equation needs to be included to the system of equations for the DE model because it provides the law of motion for the realized expectations. To see this,

$$\begin{aligned}(\mathbb{E}_t\mathbf{y}_{t+1}^{RE})^\theta &= \underbrace{[\mathbf{M}_1 : \mathbf{0}]}_{\mathbf{M}} \begin{bmatrix} \mathbf{T}_{11}^{RE} & \mathbf{T}_{12}^{RE} \\ \mathbf{T}_{21}^{RE} & \mathbf{T}_{22}^{RE} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{x}}_t^\theta \\ (\mathbb{E}_t\mathbf{y}_{t+1}^{RE})^\theta \end{bmatrix} \\ &= \mathbf{M}_1\mathbf{T}_{11}^{RE}\widetilde{\mathbf{x}}_t^\theta + \mathbf{M}_1\mathbf{T}_{12}^{RE}(\mathbb{E}_t\mathbf{y}_{t+1}^{RE})^\theta\end{aligned}$$

so

$$-\mathbf{M}_1\mathbf{T}_{11}^{RE}\widetilde{\mathbf{x}}_t^\theta + (\mathbf{I} - \mathbf{M}_1\mathbf{T}_{12}^{RE})(\mathbb{E}_t\mathbf{y}_{t+1}^{RE})^\theta = 0.$$

It is useful to divide variables $\mathbf{x}_t^{RE}$ in the original gensys system into non-expectation terms and expectation terms:

$$\begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{\Gamma}_{0,11}} & \underset{(n-n_e)\times n_e}{\mathbf{\Gamma}_{0,12}} \\ \underset{n_e\times(n-n_e)}{\mathbf{\Gamma}_{0,21}} & \underset{n_e\times n_e}{\mathbf{\Gamma}_{0,22}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_t^{RE}} \\ \underset{n_e\times 1}{\mathbb{E}_t\mathbf{y}_{t+1}^{RE}} \end{bmatrix} = \begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{\Gamma}_{1,11}} & \underset{(n-n_e)\times n_e}{\mathbf{\Gamma}_{1,12}} \\ \underset{n_e\times(n-n_e)}{\mathbf{\Gamma}_{1,21}} & \underset{n_e\times n_e}{\mathbf{\Gamma}_{1,22}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_{t-1}} \\ \underset{n_e\times 1}{\mathbb{E}_{t-1}\mathbf{y}_t^{RE}} \end{bmatrix}$$
$$+ \begin{bmatrix} \underset{(n-n_e)\times n_s}{\mathbf{\Psi}_1} \\ \underset{n_e\times n_s}{\mathbf{\Psi}_2} \end{bmatrix} \underset{n_s\times 1}{\varepsilon_t} + \begin{bmatrix} \underset{(n-n_e)\times n_e}{\mathbf{\Pi}_1} \\ \underset{n_e\times n_e}{\mathbf{\Pi}_2} \end{bmatrix} \underset{n_e\times 1}{\eta_t^{RE}}$$

Then, the model under DE can be expressed using matrix notation as:

$$\mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta = \mathbf{\Gamma}_2^\theta \mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE} + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t \tag{9}$$

where $\mathbf{\Gamma}_0^\theta$ includes the RE restrictions:

$$\begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{\Gamma}_{0,11}} & \underset{(n-n_e)\times n_e}{\mathbf{0}} \\ \underset{n_e\times(n-n_e)}{-\mathbf{M}_1\mathbf{T}_{11}^{RE}} & \underset{n_e\times n_e}{\mathbf{I}-\mathbf{M}_1\mathbf{T}_{12}^{RE}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_t^\theta} \\ \underset{n_e\times 1}{\left(\mathbb{E}_t\mathbf{y}_{t+1}^{RE}\right)^\theta} \end{bmatrix} = \begin{bmatrix} \underset{(n-n_e)\times n_e}{-\mathbf{\Gamma}_{0,12}} \\ \underset{n_e\times n_e}{\mathbf{0}} \end{bmatrix} \mathbb{E}_t^\theta \mathbf{y}_{t+1}^{RE}$$
$$+ \begin{bmatrix} \underset{(n-n_e)\times(n-n_e)}{\mathbf{\Gamma}_{1,11}} & \underset{(n-n_e)\times n_e}{\mathbf{\Gamma}_{1,12}} \\ \underset{n_e\times(n-n_e)}{\mathbf{0}} & \underset{n_e\times n_e}{\mathbf{0}} \end{bmatrix} \begin{bmatrix} \underset{(n-n_e)\times 1}{\widetilde{\mathbf{x}}_{t-1}^\theta} \\ \underset{n_e\times 1}{\left(\mathbb{E}_{t-1}\mathbf{y}_t^{RE}\right)^\theta} \end{bmatrix} + \begin{bmatrix} \underset{(n-n_e)\times n_s}{\mathbf{\Psi}_1} \\ \underset{n_e\times n_s}{\mathbf{0}} \end{bmatrix} \varepsilon_t$$

Then:

$$\begin{aligned} \mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta &= \mathbf{\Gamma}_2^\theta E_t^\theta \mathbf{y}_{t+1}^{RE} + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t \\ \mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta &= \mathbf{\Gamma}_2^\theta \left[(1+\theta)\,\mathbb{E}_t \mathbf{y}_{t+1}^{RE} - \sum_{j=1}^{J} \theta\alpha_j \mathbb{E}_{t-j}\mathbf{y}_{t+1}^{RE} \right] + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t \end{aligned}$$

Suppose that we do not need all elements in $\mathbf{x}_t^\theta$ to form expectations about the future.[1] In particular, we have

$$\begin{aligned} \mathbf{y}_t^{RE} &= \mathbf{M}\mathbf{x}_t^{RE} \\ \mathbf{x}_t^{RE} &= \mathbf{T}^{RE}\mathbf{x}_{t-1}^{RE} + \mathbf{R}^{RE}\varepsilon_t \end{aligned}$$

but can be reduced to

$$\begin{aligned} \mathbf{y}_t^{RE} &= \widetilde{\mathbf{M}}\widetilde{\mathbf{x}}_t^{RE} \\ \widetilde{\mathbf{x}}_t^{RE} &= \widetilde{\mathbf{T}}^{RE}\widetilde{\mathbf{x}}_{t-1}^{RE} + \widetilde{\mathbf{R}}^{RE}\varepsilon_t \end{aligned}$$

[1]The method can easily allow for the case where we need full elements in $\mathbf{x}_t^\theta$ to form expectations. The advantage of the current method is that its state space is smaller and hence is useful for a DSGE estimation, among other things.

Then (9) becomes

$$\mathbf{\Gamma}_0^\theta \mathbf{x}_t^\theta = \mathbf{\Gamma}_2^\theta \left[(1+\theta)\, \mathbf{M}\mathbf{T}^{RE}\mathbf{x}_t^\theta - \sum_{j=1}^{J} \theta \alpha_j \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE} \right)^{j+1} \widetilde{\mathbf{x}}_{t-j}^\theta \right] + \mathbf{\Gamma}_1^\theta \mathbf{x}_{t-1}^\theta + \mathbf{\Psi}^\theta \varepsilon_t. \tag{10}$$

This becomes:

$$\begin{aligned}
\left[\mathbf{\Gamma}_0^\theta - \mathbf{\Gamma}_2^\theta (1+\theta)\, \mathbf{M}\mathbf{T}^{RE} \right] \mathbf{x}_t^\theta = & \left[\mathbf{\Gamma}_1^\theta - \mathbf{\Gamma}_2^\theta \theta \alpha_1 \mathbf{M} (\mathbf{T}^{RE})^2 \right] \mathbf{x}_{t-1}^\theta \\
& - \mathbf{\Gamma}_2^\theta \theta \alpha_2 \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE} \right)^3 \widetilde{\mathbf{x}}_{t-2}^\theta \\
& \dots \\
& - \mathbf{\Gamma}_2^\theta \theta \alpha_J \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE} \right)^{J+1} \widetilde{\mathbf{x}}_{t-J}^\theta \\
& + \mathbf{\Psi}^\theta \varepsilon_t.
\end{aligned}$$

The solution can be obtained inverting the left hand side matrix:

$$\begin{aligned}
\mathbf{x}_t^\theta = & (\mathbf{A}_0^\theta)^{-1} \left[\mathbf{\Gamma}_1^\theta - \mathbf{\Gamma}_2^\theta \theta \alpha_1 \mathbf{M} (\mathbf{T}^{RE})^2 \right] \mathbf{x}_{t-1}^\theta \\
& - (\mathbf{A}_0^\theta)^{-1} \mathbf{\Gamma}_2^\theta \theta \alpha_2 \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE} \right)^3 \widetilde{\mathbf{x}}_{t-2}^\theta \\
& \dots \\
& - (\mathbf{A}_0^\theta)^{-1} \mathbf{\Gamma}_2^\theta \theta \alpha_J \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE} \right)^{J+1} \widetilde{\mathbf{x}}_{t-J}^\theta \\
& + (\mathbf{A}_0^\theta)^{-1} \mathbf{\Psi}^\theta \varepsilon_t,
\end{aligned}$$

where $\mathbf{A}_0^\theta \equiv \left[\mathbf{\Gamma}_0^\theta - \mathbf{\Gamma}_2^\theta (1+\theta)\, \mathbf{M}\mathbf{T}^{RE} \right]$.

Writing in a more compact form, we obtain

$$\underbrace{\begin{bmatrix} \mathbf{x}_t^\theta \\ \widetilde{\mathbf{x}}_{t-1}^\theta \\ \vdots \\ \widetilde{\mathbf{x}}_{t-J+1}^\theta \end{bmatrix}}_{\mathbf{z}_t^\theta}$$

$$= \underbrace{\begin{bmatrix} \left(\mathbf{A}_0^\theta\right)^{-1} \left[\mathbf{\Gamma}_1^\theta - \mathbf{\Gamma}_2^\theta \theta \alpha_1 \mathbf{M} (\mathbf{T}^{RE})^2 \right] & -\left(\mathbf{A}_0^\theta\right)^{-1} \mathbf{\Gamma}_2^\theta \theta \alpha_2 \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE} \right)^3 & \dots & -\left(\mathbf{A}_0^\theta\right)^{-1} \mathbf{\Gamma}_2^\theta \theta \alpha_J \widetilde{\mathbf{M}} \left(\widetilde{\mathbf{T}}^{RE} \right)^{J+1} \\ \mathbf{S} & \mathbf{0} & \dots & \mathbf{0} \\ & & \mathbf{I} & \\ \mathbf{0} & \dots & & \mathbf{0} \end{bmatrix}}_{\mathbf{T}^\theta}$$

$$\underbrace{\begin{bmatrix} \mathbf{x}_{t-1}^\theta \\ \widetilde{\mathbf{x}}_{t-2}^\theta \\ \vdots \\ \widetilde{\mathbf{x}}_{t-J}^\theta \end{bmatrix}}_{\mathbf{z}_{t-1}^\theta} + \underbrace{\begin{bmatrix} \left(\mathbf{A}_0^\theta\right)^{-1} \mathbf{\Psi}^\theta \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{\mathbf{R}^\theta} \varepsilon_t,$$

where $\mathbf{S}$ is a selection matrix that relates $\mathbf{x}_t^\theta$ to $\widetilde{\mathbf{x}}_t^\theta$:

$$\widetilde{\mathbf{x}}_t^\theta = \mathbf{S}\mathbf{x}_t^\theta.$$

Finally, we check that all variables over which we take DE present residual uncertainty. To do this, we define a vector $\mathbf{w}_t^{RE} = \mathbf{Q}\mathbf{x}_t^{RE}$ that extracts all relevant linear combinations from the vector $\mathbf{x}_t^{RE}$. This vector contains all and only the variables over which we compute DE. Then, for each element $w_{j,t}^{RE}$ of this vector we verify that the one-step-ahead conditional variance is positive:

$$Var_t\left(w_{j,t+1}^{RE}\right) = (\mathbf{Q}\mathbf{R}^{RE}\mathbf{\Sigma}(\mathbf{Q}\mathbf{R}^{RE})')_{j,j} > 0,$$

where $\mathbf{\Sigma} \equiv \mathbb{E}_t[\varepsilon_{t+1}\varepsilon_{t+1}']$ and $(\cdot)_{j,j}$ indicates the j-th diagonal element of the matrix.

E Estimation method

Our description of the methodology closely follows Christiano et al. (2010). The Bayesian estimation of impulse-response-matching method first computes the 'likelihood" of the data using approximation based on standard asymptotic distribution theory. Let $\hat{\psi}$ denote the impulse response function calculated from a local projection and let $\psi(\theta)$ denote the impulse response function from the DSGE model, which depend on the structural parameters θ. Suppose the DSGE model is correct and let θ_0 denote the true parameter vector; hence $\psi(\theta_0)$ is the true impulse response function. Then we have

$$\sqrt{T}(\hat{\psi} - \psi(\theta_0)) \xrightarrow{d} N(0, W(\theta_0)),$$

where T is the number of observations and $W(\theta_0)$ is the asymptotic sampling variance, which depends on θ_0. The asymptotic distribution of $\hat{\psi}$ can be rewritten as

$$\hat{\psi} \xrightarrow{d} N(\psi(\theta_0), V), \qquad V \equiv \frac{W(\theta_0)}{T}.$$

We use a consistent estimator of V, where the non-diagonal terms are set to zero and the main diagonal elements consist of the sample variance of $\hat{\psi}$.[2] As Christiano et al. (2011) explains, there are two advantages of this approach. First, it improves small sample efficiency and can be justified using a logic similar to the estimation of frequency-zero spectral densities in Newey and West (1987). Second, the interpretation of the estimator is graphically intuitive and transparent: it chooses parameters so that the model-implied impulse responses lie inside a confidence interval around the empirical responses.[3]

The method then calculates the likelihood

$$\mathcal{L}(\psi|\theta) = (2\pi)^{-\frac{N}{2}}|V|^{-\frac{1}{2}}\exp\{-0.5[\hat{\psi} - \psi(\theta)]'V^{-1}[\hat{\psi} - \psi(\theta)]\},$$

where N is the total number of elements in the impulse responses to be matched. Intuitively, the likelihood is higher when the model-based impulse response $\psi(\theta)$ is closer to the empirical

[2] Altig et al. (2011) and Christiano et al. (2005) use this approach in a frequentist context.

[3] In contrast, when the non-diagonal terms of V are non-zero, the estimator also takes into account the deviations of the model from data across different impulse responses in a non-intuitive manner.

counterpart $\hat{\psi}$, taking into account the precision of the estimated empirical responses. We use the Bayes law to obtain the posterior distribution $p(\theta|\psi)$:

$$p(\theta|\psi) = \frac{p(\theta)\mathcal{L}(\psi|\theta)}{p(\psi)},$$

where $p(\theta)$ is the prior and $p(\psi)$ is the marginal likelihood. We simulate the posterior distribution $p(\theta|\psi)$ using the random-walk Metropolis-Hasting algorithm.

To conduct model comparisons, we use marginal likelihoods, computed from the MCMC output using the Geweke (1999)'s modified harmonic mean estimator. Inoue and Shintani (2018) provide asymptotic justification for a such exercise. In particular, they show that a model with a higher marginal likelihood is either correct or a better approximation to true impulse responses as the sample size approaches infinity.

F Additional Results

In this Appendix we report some additional results for the estimated DSGE model.

Table F1 reports the priors and the posterior mode for the model parameters of the DE model and RE re-estimated model. Standard deviations are reported in parentheses. The priors are symmetric across the two models and diffuse.

Figure F1 reports the consumption impulse responses of DE and RE models without consumption habit. In the estimated DE model without habit, the diagnosticity parameter is estimated to be significantly smaller than the benchmark DE model at $\theta = 0.64$. The mean and standard deviation of the Beta distribution that control the weights for the comparison group are 0.35 and 0.17, respectively, resulting in a more distant and dispersed memory relative to the benchmark estimated DE model with habit. The left panel of Figure F1 shows that the DE model without habit is able to generate boom-bust cycle in consumption. In response to an unexpected Fed rate cut, consumption initially spikes. This is because without habit, according to the Euler equation (4.39), the lower-than-usual interest rate implies falling consumption (negative expected consumption growth).[4] In the medium run, consumption gradually increases because, as shown in the right panel, unexpectedly high consumption implies unexpectedly high RE consumption $\mathbb{E}_t\left[\widehat{c}_{t+1}^{RE}\right]$ (relative to the comparison group $\mathbb{E}_t^r\left[\widehat{c}_{t+1}^{RE}\right]$) and hence high DE consumption $\mathbb{E}_t^\theta\left[\widehat{c}_{t+1}^{RE}\right]$. As consumption drops, agents become overly pessimistic (low $\mathbb{E}_t^\theta\left[\widehat{c}_{t+1}^{RE}\right]$), feeding into significantly low consumption below trend. An interesting difference of the no habit model compared to the benchmark DE model with habit is that, because agents expect a quicker reversion of consumption to trend, movements in DE consumption $\mathbb{E}_t^\theta\widehat{c}_{t+1}^{RE}$ are smoother and, as a result, the surprise in cumulative inflation $\pi_{J,t}^*$ plays a smaller role in determining the dynamics. Finally, note that in contrast to the DE model, both counterfactual and re-estimated RE models generate transitory IRFs that understate the amplitude of empirical consumption response. As a result, the log marginal likelihood of the DE model is $(-388-(-437)=)$ 49 log points higher than the RE model. We conclude that the DE model delivers boom-bust dynamics irrespective of whether it features consumption habit or not.

Figure F2 reports the model impulse responses when we use an alterative prior for the diagnosticity parameter θ. Specifically, we consider a Normal prior with mean 1 and standard

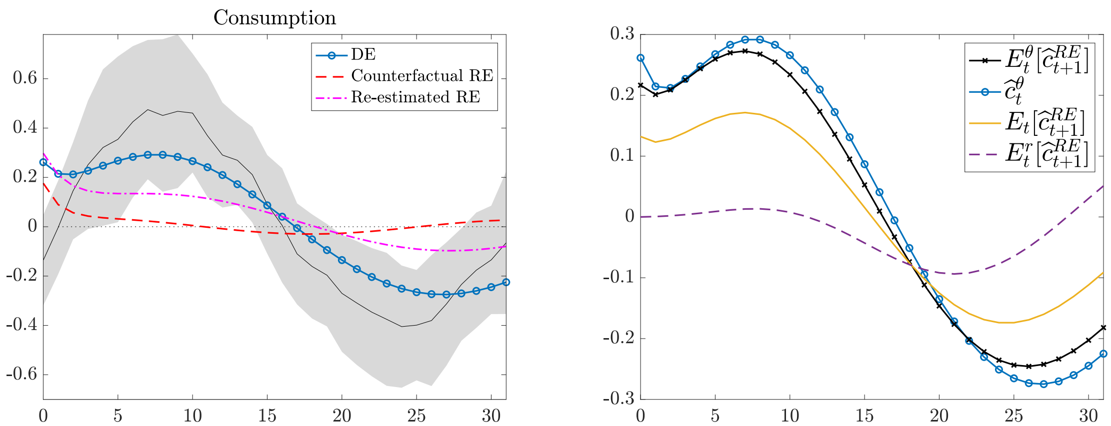
[4] In the benchmark model with habit, this initial spike is absent because habit suppresses the initial spike by breaking the tight link between low rate and negative consumption growth.

Table F1: Estimated parameters

		Prior			Posterior mode	
		Type	Mean	Std	DE	RE
η	Inverse Frisch elasticity	G	2	0.3	2.11	1.56
					(0.28)	(0.26)
b	Consumption habit	B	0.5	0.2	0.80	0.90
					(0.01)	(0.01)
τ	Utilization cost	IG	1	1	0.22	0.27
					(0.01)	(0.01)
κ	Investment adjustment cost	G	2	0.2	2.97	5.48
					(0.20)	(0.36)
φ_p	Price adjustment cost	G	100	20	195.4	291.28
					(21.2))	(31.6)
φ_w	Wage adjustment cost	G	100	20	88.6	78.8
					(20.3)	(18.0)
ρ_R	Taylor rule smoothing	B	0.5	0.2	0.009	0.82
					(0.008)	(0.018)
ϕ_π	Taylor rule inflation	N	1.5	0.4	1.000	1.000
					(0.01)	(0.025)
ϕ_Y	Taylor rule output	N	0.1	0.05	0.67	0.23
					(0.02)	(0.05)
$100\sigma_R$	Monetary policy shock	IG	1	1	0.15	0.21
					(0.01)	(0.02)
θ	Diagnosticity parameter*	N	0	0.2	1.97	–
					(0.10)	
μ	Memory distribution mean	B	0.5	0.2	0.17	–
					(0.01)	
σ	Memory distribution stdev	G	0.2	0.05	0.03	–
					(0.004)	
Log marginal likelihood					**-345**	**-369**

Notes: 'DE' corresponds to the model with diagnostic expectations and 'RE' corresponds to the rational expectations version. B refers to the Beta distribution, N to the Normal distribution, G to the Gamma distribution, IG to the Inverse-gamma distribution. (*For the prior for the diagnoscity parameter, we truncate the Normal distribution above $\theta \geq 0$.) Posterior standard deviations are in parentheses and are obtained from draws using the random-walk Metropolis-Hasting algorithm. The marginal likelihood is calculated using Geweke's modified harmonic mean estimator.

Figure F1: Consumption paths in a model without habit



Notes: The left panel shows the consumption IRFs in response to a monetary policy shock from the DE model without habit (blue circled line), counterfactual RE model without habit (red dashed line) and the re-estimated RE model without habit (magenta dashed line). The right panel plots DE expected consumption ($\mathbb{E}_t^\theta\left[\widehat{c}_{t+1}^{RE}\right]$), realized equilibrium consumption ($\widehat{c}_t^\theta$), RE expected consumption ($\mathbb{E}_t\left[\widehat{c}_{t+1}^{RE}\right]$) and reference expectation ($\mathbb{E}_t^r\left[\widehat{c}_{t+1}^{RE}\right]$).

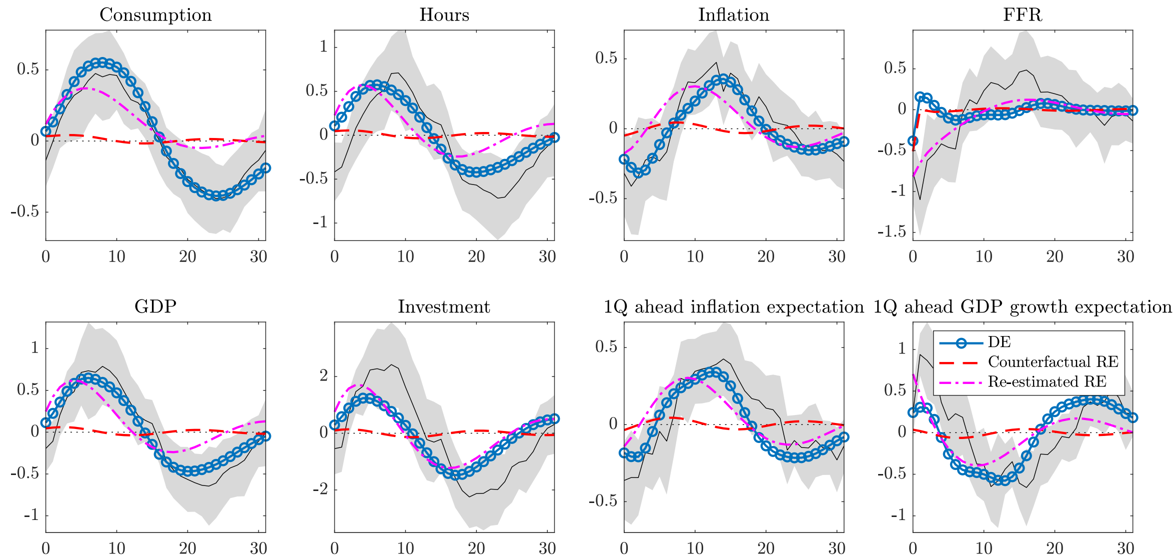
deviation 0.2. We find that the estimated $\theta = 2.16$, slightly higher than the benchmark estimate of $\theta = 1.97$. Nevertheless, the estimated DE IRF under the alternative prior is very similar to the benchmark IRF reported in the paper.

Figure F3 reports the model impulse responses when we target the inflation and output expectations. When expectations are targeted, the diagnosticity parameter is estimated to be slightly lower than the benchmark DE model at $\theta = 1.77$. The mean and standard deviation of the Beta distribution that control the weights for the comparison group are 0.19 and 0.04, respectively and thus are similar to the estimated values in the benchmark DE model. Figure F3 shows that the DE model is able to replicate the boom-bust cycles in macro variables as well as the responses of survey expectations, although it slightly overstates consumption during the decline after the peak. The counterfactual RE model where we set the diagnosticity parameter θ to 0 while holding fixed other estimated parameters generates transitory and negligible response. The re-estimated RE model misses the magnitude of the bust in consumption, hours and GDP. It also has difficulty matching realized and expected inflation. As a result, the log marginal likelihood of the DE model is $(-636 - (-645) =)$ 9 log points higher than the RE model.

Figure F4 reports the impulse response of the marginal utility to an expensionary monetary policy shock, given that the estimated Euler Equation features habits:

$$-\mathbb{E}_t^\theta[\widehat{\lambda}_{t+1}^{RE}] + \widehat{\lambda}_t^\theta = \widehat{r}_t^\theta - \mathbb{E}_t^\theta[\widehat{\pi}_{t+1}^{RE}] - \theta\pi_{J,t}^* \tag{11}$$

Figure F2: Impulse responses to a monetary policy shock: Alternative prior for θ



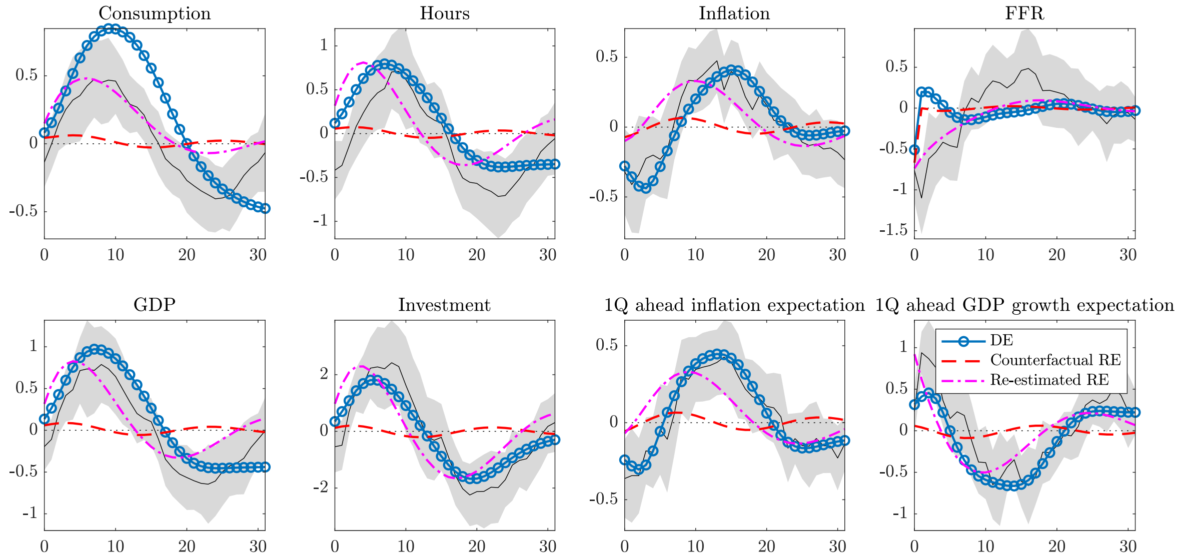
Notes: The interpretation of the plotted lines follow their description for Figure 2. The responses of consumption, hours, GDP and investment are in percentage deviations from the steady states while the inflation, FFR and inflation and output growth expectations are in annual percentage points.

where

$$\widehat{\lambda}_t^{\theta} = -\frac{\widehat{c}_t^{\theta} - b\gamma^{-1}\widehat{c}_{t-1}^{\theta}}{1 - b\gamma^{-1}}. \tag{12}$$

Marginal utility follows a symmetric pattern with respect to consumption, once controlling for habits. The initial increase in consumption is associated with low expected marginal utility that induces expectations of even lower marginal utility. Thus, agents expect consumption to increase even when controlling for the stock of habits. As the economy progresses in its response to the shock, consumption starts declining and marginal utility to increase. However, reference expectations for marginal utility also start increasing. This is because reference expectations were formed at a time of high consumption. Under RE, agents expect a fairly quick return to the steady state from above, implying consumption lower than the stock of habits, leading to a positive RE marginal utility. However, under DE, the return to the steady state is slower than expected as agents remain overly optimistic for a while. Agents are still surprised by the high consumption, leading to a negative surprise in marginal utility, amplified by DE. Thus, past decisions feed into current beliefs, affecting the duration and amplitude of the cycle. It is only around 15 quarters that reference expectations catch up with the current marginal utility. As consumption moves below trend, agents start expecting a return to the steady from below, generating a negative reference expectation for marginal utility as consumption is expected to be higher than the stock of habits. In the bust phase, agents are surprised by the fact that consumption is still well below trend, leading to a positive surprise in marginal utility, that induces magnified DE of high marginal utility in

Figure F3: Impulse responses to a monetary policy shock: Survey expectations as targets



Notes: The interpretation of the plotted lines follow their description for Figure 2. The responses of consumption, hours, GDP and investment are in percentage deviations from the steady states while the inflation, FFR and inflation and output growth expectations are in annual percentage points.

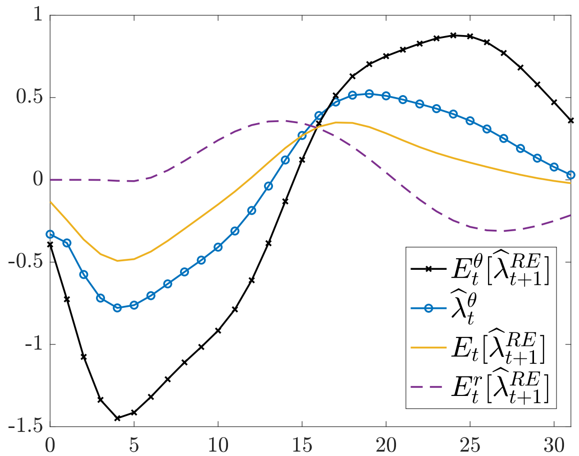
the future.

How can we rationalize this behavior from the perspective of the Euler equation under DE in (4.39)? As mentioned in the paper, a key role is played by the surprise in cumulative inflation $\pi^*_{J,t}$ with respect to the reference expectations formed in the past. On impact, because of an increase in utilization, inflation declines. This determines a negative surprise in the price level that induces a misperception in the model relevant real interest rate that starts increasing. This perceived high real interest is, in the eyes of the agent, justified in light of a perceived acceleration in consumption that more than compensates for the habit stock. In other words, not only agents expect consumption to be higher in the future, but they also expect the marginal utility to be lower: $-\mathbb{E}^\theta_t[\widehat{\lambda}^{RE}_{t+1}] + \widehat{\lambda}^\theta_t > 0$ implies $\mathbb{E}^\theta_t\left[\widehat{c}^{RE}_{t+1} - b\gamma^{-1}\widehat{c}^\theta_t\right] - \left[\widehat{c}^\theta_t - b\gamma^{-1}\widehat{c}^\theta_{t-1}\right] > 0$. Eventually, inflation starts picking up, leading first to a reduction in the negative surprises for the price level and then eventually to positive surprises. This determines a reversal in the model relevant real interest rate that moves into the negative territory during the bust part of the cycle, when agents find the perceived low real interest rate justified in light of their excessive pessimism. Now not only they expect consumption to decline, but also to do so in a way to increase the marginal utility.

G Alternative Expression of New Keynesian Phillips Curve

In this Appendix we derive an alternative expression of the New Keynesian Phillips Curve (NKPC) of our DE model that we use to discuss the connection of inflation and real activity

Figure F4: Impulse response of marginal utility



Notes: The Figure shows the DE marginal utility ($\mathbb{E}_t^\theta[\widehat{\lambda}_{t+1}^{RE}]$), realized equilibrium marginal utility ($\widehat{\lambda}_t^\theta$), RE marginal utility ($\mathbb{E}_t[\widehat{\lambda}_{t+1}^{RE}]$) and reference expectation of marginal utility ($\mathbb{E}_t^r[\widehat{\lambda}_{t+1}^{RE}]$).

in Section 4. Consider the NKPC:

$$\widehat{\pi}_t^\theta = \beta \mathbb{E}_t^\theta[\widehat{\pi}_{t+1}^{RE}] + \kappa_p \widehat{mc}_t^\theta,$$

where $\kappa_p \equiv (\varphi_p \Pi^2(\lambda_f - 1))^{-1}$. The shadow RE NKPC reads:

$$\widehat{\pi}_t^{RE} = \beta \mathbb{E}_t[\widehat{\pi}_{t+1}^{RE}] + \kappa_p \widehat{mc}_t^{RE}.$$

Iterating forward the RE version of the NKPC, RE inflation can be expressed as:

$$\widehat{\pi}_t^{RE} = \kappa_p \sum_{i=0}^{\infty} \beta^i \mathbb{E}_t \widehat{mc}_{t+i}^{RE}.$$

Thus, we have:

$$\widehat{\pi}_t^\theta = \kappa_p \sum_{i=1}^{\infty} \beta^i \mathbb{E}_t^\theta[\widehat{mc}_{t+i}^{RE}] + \kappa_p \widehat{mc}_t^\theta,$$

where we have used the fact that DE are additive as long as uncertainty is present. This expression makes clear that inflation depends on the DE of future marginal costs for a given starting value of current marginal costs.

References

Altig, D., L. J. Christiano, M. Eichenbaum, and J. Lindé (2011, April). Firm-specific capital, nominal rigidities and the business cycle. *Review of Economic Dynamics 14*(2), 225–247.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy 113*(1), 1–45.

Christiano, L. J., M. Trabandt, and K. Walentin (2010). Involuntary unemployment and the business cycle. NBER Working Paper.

Christiano, L. J., M. Trabandt, and K. Walentin (2011). Dsge models for monetary policy analysis. In B. M. Friedman and M. Woodford (Eds.), *Handbook of Monetary Economics*, Volume 3A, Chapter 7, pp. 285–367. Elsevier.

Geweke, J. (1999). Using simulation methods for bayesian econometric models: Inference, development, and communication. *Econometric Reviews 18*, 1–126.

Inoue, A. and M. Shintani (2018). Quasi-bayesian model selection. *Quantitative Economics 9*, 1265–1297.

Newey, W. K. and K. D. West (1987, May). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Sims, C. A. (2000). Solving linear rational expectations models. *Computational Economics 20*(1–2), 1–20.