A. Appendix for "online" publication

A.1. Demand for good i, Aggregate Price Index and Optimality Conditions

The representative household derives utility from consumption of different goods, indexed $i \in I = [0, 1]$, according to the consumption index. Let $\mathcal{C} = \{C_t \in \mathcal{L}^1 : C_t : I \to \mathbb{R}$ is quasi-concave and Borel measurable, $t \in \mathbb{Z}_+\}$ be the set of consumption choice functions over the set of goods I in the economy at a given period t.

Given the price function $P_t : I \to \mathbb{R}_+$ with $\|P_t\|_\infty < \infty$, and for a fixed endowment $Z_t \in \mathbb{R}_+$, the representative household's maximization problem at period t is $\widetilde{C}_t = \max_{C_t \in \mathcal{C}} \left[\int_0^1 C_t(i)^{\frac{\epsilon_t - 1}{\epsilon_t}} di \right]^{\frac{\epsilon_t}{\epsilon_t - 1}}$, subject to the budget constraint:

$$\int_0^1 P_t(i) C_t(i) \; di \leq Z_t \tag{22}$$

which will be satisfied with equality in the optimum. The derivative of the Lagrangian with respect to $C_t(i)$, the consumption level of good i, yields $\left[\int_0^1 C_t(i)^{\frac{\epsilon_t - 1}{\epsilon_t}} di \right]^{\frac{1}{\epsilon_t - 1}} C_t(i)^{-\frac{1}{\epsilon_t}} - \lambda_t P_t(i) = 0 \implies \widetilde{C}_t^{\frac{1}{\epsilon_t}} C_t(i)^{-\frac{1}{\epsilon_t}} = \lambda_t P_t(i)$ where λ_t is the sequence of Lagrange multipliers attached to the sequence of restrictions (22). By dividing the last expression for two different goods $i, j \in I$, we find relation between the optimal consumption levels of two different goods:

$$C_t(i) = \left[\frac{P_t(j)}{P_t(i)} \right]^{\epsilon_t} C_t(j) \tag{23}$$

and inserting (23) into (22),

$$Z_t = \int_0^1 P_t(i) \left[\frac{P_t(j)}{P_t(i)} \right]^{\epsilon_t} C_t(j) \; di \implies C_t(j) = \frac{Z_t P_t(j)^{-\epsilon_t}}{\int_0^1 P_t(i)^{1-\epsilon_t} di} \tag{24}$$

we obtain an expression for the optimal consumption levels of almost all goods in terms of prices and the initial endowment. Integrating the last equation over all goods gives the optimal aggregate consumption level $\widetilde{C}_t = \left[\int_0^1 \left(\frac{Z_t P_t(i)^{-\epsilon_t}}{\int_0^1 P_t(i)^{1-\epsilon_t} di} \right)^{\frac{\epsilon_t - 1}{\epsilon_t}} di \right]^{\frac{\epsilon_t}{\epsilon_t - 1}} = Z_t \left[\int_0^1 P_t(i)^{1-\epsilon_t} di \right]^{\frac{1}{\epsilon_t - 1}}$.

Now, let's define $\widetilde{P}_t$ as the unit cost of the aggregate consumption level $\widetilde{C}_t$ at endowment level Z, $\widetilde{P}_t\widetilde{C}_t = Z_t$. Hence,

$$\widetilde{P}_t Z_t \left[\int_0^1 P_t(i)^{1-\epsilon_t} di\right]^{\frac{1}{\epsilon_t - 1}} = Z_t \implies \widetilde{P}_t = \left[\int_0^1 P_t(i)^{1-\epsilon_t} di\right]^{\frac{1}{1-\epsilon_t}} \tag{25}$$

where (25) is the price index. Inserting (25) into (24), $C_t(j) = \frac{Z_t P_t(j)^{-\epsilon_t}}{\widetilde{P}_t^{1-\epsilon_t}} = \frac{Z_t}{\widetilde{P}_t}\left[\frac{\widetilde{P}_t}{P_t(i)}\right]^{\epsilon_t}$. And finally, replacing Z_t we find the desired optimal consumption for good i in terms of the aggregate good and the aggregate price:

$$C_t(i) = \left[\frac{P_t(i)}{\widetilde{P}_t}\right]^{-\epsilon_t} \widetilde{C}_t \tag{26}$$

With market clearing and a representative household setting, $C_t(i) = Y_t(i)$ and $\widetilde{C}_t = \widetilde{Y}_t$, and we obtain expression (3). Since we deal with the aggregate quantities in the rest of the paper, with a slight abuse of notation we drop the tilde from the aggregate terms.

Finally, in order to obtain the optimality conditions we form the Lagrangian, $\mathcal{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{(C_t - h\overline{C}_{t-1})^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} + \lambda_t \left[B_{t-1} + W_t N_t + T_t - P_t C_t - Q_t B_t\right]\right]$. The FOCs with respect to C_t, B_t and N_t yield,

$$\begin{aligned} C_t:&\quad \lambda_t P_t = (C_t - h\overline{C}_{t-1})^{-\sigma} \\ N_t:&\quad \lambda_t W_t = N_t^{\varphi} \\ B_t:&\quad \lambda_t Q_t = \lambda_{t+1} \end{aligned}$$

Combining them and cancelling the lagrange multiplier λ_t we obtain the optimality conditions $\frac{W_t}{P_t} = \frac{N_t^{\varphi}}{(C_t - h\overline{C}_{t-1})^{-\sigma}}$ and $Q_t = \beta \mathbb{E}_t \left[\left(\frac{C_{t+1} - h\overline{C}_t}{C_t - h\overline{C}_{t-1}}\right)^{-\sigma} \frac{P_t}{P_{t+1}}\right]$.

A.2. Log-linearization of Behavioural Household's Optimality Conditions

We now proceed to log-linearize (4) and (6). Starting with (4), taking a first order Taylor approximation around a zero-inflation steady state, we obtain, $\widehat{w}_t - \widehat{p}_t = \varphi \widehat{n}_t + \frac{\sigma}{1-h}\widehat{c}_t - \frac{\sigma h}{1-h}\widehat{c}_{t-1}$. Turning to (6), we first take logs, $\log Q_t = \log\beta + \mathbb{E}_t^B\{-\sigma\log(C_{t+1} - hC_t) + \sigma\log(C_t - hC_{t-1}) + \log P_t - \log P_{t+1}\}$. Since $Q_t = 1/(1 + i_t)$, one can show that $i_t \approx -\log Q_t$. We then write $\rho = -\log\beta$ and $\pi_{t+1} = \log\frac{P_{t+1}}{P_t}$. Let us now log-linearize the terms that include consumption, $\log[C_{t+1} - hC_t] \approx \log[(1-h)C] + \frac{1}{(1-h)C}[C_{t+1} - C] - \frac{h}{(1-h)C}[C_t - C] =$

$\log[(1-h)C] + \frac{1}{1-h}\widehat{c}_{t+1} - \frac{h}{1-h}\widehat{c}_t$. Proceeding in a similar manner with the other consumption term, and plugging into the above expression leads to

$$0 = \mathbb{E}_t^B \left\{ i_t - \rho - \frac{\sigma}{1-h}\left[\widehat{c}_{t+1} - (1+h)\widehat{c}_t + h\widehat{c}_{t-1}\right] - \pi_{t+1} \right\} \tag{27}$$

Under cognitive discounting, $\mathbb{E}_t^B x_{t+k} = \overline{m}^k \mathbb{E}_t x_{t+k}$ for any variable x. Hence, $0 = \widehat{i}_t - \frac{\sigma}{1-h}\overline{m}\mathbb{E}_t\widehat{c}_{t+1} - \frac{(1+h)\sigma}{1-h}c_t - \frac{h\sigma}{1-h}c_{t-1} - \overline{m}\mathbb{E}_t\pi_{t+1}$, where we have defined $\widehat{i}_t = i_t - i = i_t - \rho$. Rewritting this last expression leads to (8). Written in natural terms and denoting the real interest rate as $r_t = \widehat{i}_t - \overline{m}\mathbb{E}_t\pi_{t+1}$, the previous equation yields $\widehat{c}_t^n = \frac{h}{1+h}\widehat{c}_{t-1}^n + \frac{1}{1+h}\overline{m}\mathbb{E}_t\widehat{c}_{t+1}^n - \frac{1-h}{\sigma(1+h)}r_t^n$. Since $\widehat{c}_t = \widehat{y}_t = \widehat{c}_t^n = \widehat{y}_t^n$, we can rewrite it in terms of the output gap $\widetilde{y}_t = y_t - y_t^n$ and it yields (9).

A.3. Solving the Firm Problem

We can rewrite condition (13) as

$$\begin{aligned} P_t^*(i) &= \frac{\mathbb{E}_t \sum_{k=0}^{\infty} (\theta\beta)^k (C_{t+k} - hC_{t+k-1})^{-\sigma} C_{t+k} P_{t+k}^{\epsilon_{t+k}-1} P_{t+k-1}^{-\omega\epsilon_{t+k}} \frac{W_{t+k}}{A_{t+k}} \mathcal{M}_{t+k}}{\mathbb{E}_t \sum_{k=0}^{\infty} (\theta\beta)^k (C_{t+k} - hC_{t+k-1})^{-\sigma} C_{t+k} P_{t+k}^{\epsilon_{t+k}-1} P_{t+k-1}^{\omega(1-\epsilon_{t+k})}} P_{t-1}^{\omega} = \\ &= \frac{\mathbb{E}_t \sum_{k=0}^{\infty} (\theta\beta)^k (C_{t+k} - hC_{t+k-1})^{-\sigma} C_{t+k} P_{t+k}^{\epsilon_{t+k}} P_{t+k-1}^{-\omega\epsilon_{t+k}} MC_{t+k} \mathcal{M}_{t+k}}{\mathbb{E}_t \sum_{k=0}^{\infty} (\theta\beta)^k (C_{t+k} - hC_{t+k-1})^{-\sigma} C_{t+k} P_{t+k}^{\epsilon_t-1} P_{t+k-1}^{\omega(1-\epsilon_{t+k})}} P_{t-1}^{\omega} \end{aligned} \tag{28}$$

where we have used $MC_{t+k} = \frac{W_{t+k}}{A_{t+k}P_{t+k}}$. With flexible prices, (28) collapses to

$$P_t^*(i) = \mathcal{M}_t \frac{(C_t - hC_{t-1})^{-\sigma} C_t P_t^{\epsilon_t} P_{t-1}^{-\omega\epsilon_t} MC_t}{(C_t - hC_{t-1})^{-\sigma} C_t P_t^{\epsilon_t-1} P_{t-1}^{\omega(1-\epsilon_t)}} P_{t-1}^{\omega} = \mathcal{M}_t P_t MC_t \tag{29}$$

where (29) is the frictionless mark-up. To simplify computation, we now log-linearize (28). Separating both sides,

$$\begin{aligned} &P_t^* \mathbb{E}_t \sum_{k=0}^{\infty} (\theta\beta)^k (C_{t+k} - hC_{t+k-1})^{-\sigma} C_{t+k} P_{t+k}^{\epsilon_{t+k}-1} P_{t+k-1}^{\omega(1-\epsilon_{t+k})} = \\ &\qquad = \mathbb{E}_t \sum_{k=0}^{\infty} (\theta\beta)^k (C_{t+k} - hC_{t+k-1})^{-\sigma} C_{t+k} P_{t+k}^{\epsilon_{t+k}} P_{t+k-1}^{-\omega\epsilon_{t+k}} MC_{t+k} P_{t-1}^{\omega} \mathcal{M}_{t+k} \end{aligned} \tag{30}$$

We know that, in steady-state, $P_t^* = P_t = P_{t-1} = P$, $\Pi_t = \Pi = 1$, $C_t = C$, $Q_{t,t+k} = \beta^k$ and $MC_t = MC$. It lasts to find MC. To obtain it, we can write (28) in steady-state and solve for MC, $P = \mathcal{M}P\text{MC}$. Hence, $MC = \frac{1}{\mathcal{M}}$. Before log-linearizing, divide (30) by P_{t-1},

$$\frac{P_t^*}{P_{t-1}}\mathbb{E}_t\sum_{k=0}^{\infty}(\theta\beta)^k(C_{t+k}-hC_{t+k-1})^{-\sigma}C_{t+k}P_{t+k}^{\epsilon_{t+k}-1}P_{t+k-1}^{\omega(1-\epsilon_{t+k})} =$$

(31)

$$= \mathbb{E}_t\sum_{k=0}^{\infty}(\theta\beta)^k(C_{t+k}-hC_{t+k-1})^{-\sigma}C_{t+k}P_{t+k}^{\epsilon_{t+k}}P_{t+k-1}^{-\omega\epsilon_{t+k}}MC_{t+k}P_{t-1}^{\omega-1}\mathcal{M}_{t+k}$$

Log-linearizing the LHS,

$$\frac{P_t^*}{P_{t-1}}\mathbb{E}_t\sum_{k=0}^{\infty}(\theta\beta)^k(C_{t+k}-hC_{t+k-1})^{-\sigma}C_{t+k}P_{t+k}^{\epsilon-1}P_{t+k-1}^{\omega(1-\epsilon)} \simeq$$

$$\begin{aligned}
&\simeq\sum_{k=0}^{\infty}(\theta\beta)^kC^{1-\sigma}(1-h)^{-\sigma}P^{-(1-\epsilon)(1-\omega)}+\\
&+\frac{1}{P}\sum_{k=0}^{\infty}(\theta\beta)^kC^{1-\sigma}(1-h)^{-\sigma}P^{-(1-\epsilon)(1-\omega)}(P_t^*-P)-\\
&-\frac{P}{P^2}\sum_{k=0}^{\infty}(\theta\beta)^kC^{1-\sigma}(1-h)^{-\sigma}P^{-(1-\epsilon)(1-\omega)}(P_{t-1}-P)+\\
&+\sum_{k=0}^{\infty}(\theta\beta)^kC^{1-\sigma}(1-h)^{-\sigma}(\epsilon-1)P^{\epsilon-2}P^{\omega(1-\epsilon)}(P_{t+k}-P)+\\
&+\sum_{k=0}^{\infty}(\theta\beta)^kC^{1-\sigma}(1-h)^{-\sigma}P^{\epsilon-1}\omega(1-\epsilon)P^{\omega(1-\epsilon)-1}(P_{t+k-1}-P)+\\
&+\sum_{k=0}^{\infty}(\theta\beta)^k\{\underbrace{(-\sigma)[C(1-h)]^{-\sigma-1}C+[C(1-h)]^{-\sigma}}_{C^{-\sigma}(1-h)^{-\sigma-1}(1-h-\sigma)}\}P^{-(1-\epsilon)(1-\omega)}(C_{t+k}-C)+\\
&+\sum_{k=0}^{\infty}(\theta\beta)^k\underbrace{(-\sigma)[C(1-h)]^{-\sigma-1}(-h)C}_{\sigma hC^{-\sigma}(1-h)^{-\sigma-1}}P^{-(1-\epsilon)(1-\omega)}(C_{t+k-1}-C)+\\
&+\sum_{k=0}^{\infty}(\theta\beta)^kC^{1-\sigma}(1-h)^{-\sigma}P^{-(1-\epsilon)(1-\omega)}\ln(P^{1-\omega})(\epsilon_{t+k}-\epsilon)=\\
&=\sum_{k=0}^{\infty}(\theta\beta)^kC^{1-\sigma}(1-h)^{-\sigma}P^{-(1-\epsilon)(1-\omega)}\Bigg\{1+p_t^*-p-p_{t-1}+p-(1-\epsilon)(p_{t+k}-p)+\\
&+\omega(1-\epsilon)(p_{t+k-1}-p)+\left(1-\frac{\sigma}{1-h}\right)(c_{t+k}-c)+\frac{\sigma h}{1-h}(c_{t+k-1}-c)+\ln(P^{1-\omega})(\epsilon_{t+k}-\epsilon)\Bigg\}
\end{aligned}$$

Log-linearizing the RHS,

$$
\begin{aligned}
&\mathbb{E}_t \sum_{k=0}^{\infty} (\theta\beta)^k (C_{t+k} - hC_{t+k-1})^{-\sigma} C_{t+k} P_{t+k}^{\epsilon_{t+k}} P_{t+k-1}^{-\omega\epsilon_{t+k}} MC_{t+k} P_{t-1}^{\omega-1} \mathcal{M}_{t+k} \simeq \\
&\qquad \simeq \mathcal{M} \sum_{k=0}^{\infty} (\theta\beta)^k C^{1-\sigma} (1-h)^{-\sigma} P^{-(1-\epsilon)(1-\omega)} MC + \\
&\qquad\quad + \mathcal{M} \sum_{k=0}^{\infty} (\theta\beta)^k C^{1-\sigma} (1-h)^{-\sigma} P^{\epsilon} P^{-\omega\epsilon} MC (\omega - 1) P^{\omega-2} (P_{t-1} - P) + \\
&\qquad\quad + \mathcal{M} \sum_{k=0}^{\infty} (\theta\beta)^k C^{1-\sigma} (1-h)^{-\sigma} \epsilon P^{\epsilon-1} P^{-\omega\epsilon} MC P^{\omega-1} (P_{t+k} - P) + \\
&\qquad\quad + \mathcal{M} \sum_{k=0}^{\infty} (\theta\beta)^k C^{1-\sigma} (1-h)^{-\sigma} P^{\epsilon} (-\omega\epsilon) P^{-\omega\epsilon-1} MC P^{\omega-1} (P_{t+k-1} - P) + \\
&\qquad\quad + \mathcal{M} \sum_{k=0}^{\infty} (\theta\beta)^k C^{-\sigma} (1-h)^{-\sigma-1} (1-h-\sigma) P^{-(1-\epsilon)(1-\omega)} MC (C_{t+k} - C) + \\
&\qquad\quad + \mathcal{M} \sum_{k=0}^{\infty} (\theta\beta)^k C^{-\sigma} (1-h)^{-\sigma-1} \sigma h P^{-(1-\epsilon)(1-\omega)} MC (C_{t+k-1} - C) + \\
&\qquad\quad + \mathcal{M} \sum_{k=0}^{\infty} (\theta\beta)^k C^{1-\sigma} (1-h)^{-\sigma} P^{-(1-\epsilon)(1-\omega)} (MC_{t+k} - MC) + \\
&\qquad\quad + \mathcal{M} \sum_{k=0}^{\infty} (\theta\beta)^k C^{1-\sigma} (1-h)^{-\sigma} P^{-(1-\epsilon)(1-\omega)} \ln(P^{1-\omega}) \mathrm{MC} (\epsilon_{t+k} - \epsilon) + \\
&\qquad\quad + \sum_{k=0}^{\infty} (\theta\beta)^k C^{1-\sigma} (1-h)^{-\sigma} P^{-(1-\epsilon)(1-\omega)} \mathrm{MC} (\mathcal{M}_{t+k} - \mathcal{M}) = \\
&\qquad = \sum_{k=0}^{\infty} (\theta\beta)^k C^{1-\sigma} (1-h)^{-\sigma} P^{-(1-\epsilon)(1-\omega)} \Big\{ 1 - (1-\omega)(p_{t-1} - p) + \epsilon(p_{t+k} - p) - \\
&\qquad\quad - \omega\epsilon(p_{t+k-1} - p) + \left(1 - \frac{\sigma}{1-h}\right)(c_{t+k} - c) + \frac{\sigma h}{1-h}(c_{t+k-1} - c) + mc_{t+k} - mc + \\
&\qquad\quad + \ln(P^{1-\omega})(\epsilon_{t+k} - \epsilon) + (\mu_{t+k} - \mu) \Big\}
\end{aligned}
$$

Noticing that in steady state $P = 1$, and setting LHS=RHS, eliminating $C^{1-\sigma}(1 -$

$h)^{-\sigma} P^{-(1-\epsilon)(1-\omega)}$ on both sides,

$$\sum_{k=0}^{\infty}(\theta\beta)^k\Big\{1+p_t^*-p-p_{t-1}+p-(1-\epsilon)(p_{t+k}-p)+\omega(1-\epsilon)(p_{t+k-1}-p)+$$
$$+\left(1-\frac{\sigma}{1-h}\right)(c_{t+k}-c)+\frac{\sigma h}{1-h}(c_{t+k-1}-c)+(\mu_{t+k}-\mu)\Big\}=$$
$$=\sum_{k=0}^{\infty}(\theta\beta)^k\Big\{1-(1-\omega)(p_{t-1}-p)+\epsilon(p_{t+k}-p)-\omega\epsilon(p_{t+k-1}-p)+$$
$$+\left(1-\frac{\sigma}{1-h}\right)(c_{t+k}-c)+\frac{\sigma h}{1-h}(c_{t+k-1}-c)+mc_{t+k}-mc+\mu_{t+k}-\mu\Big\}$$

Rearranging and cancelling terms, we end up with

$$\begin{aligned}p_t^* &= (1-\theta\beta)\sum_{k=0}^{\infty}(\theta\beta)^k\mathbb{E}_t\left[p_{t+k}-\omega(\omega\, p_{t+k-1}-p_{t-1})+mc_{t+k}-mc+\mu_{t+k}-\mu\right]\\ &= (1-\theta\beta)\sum_{k=0}^{\infty}(\theta\beta)^k\mathbb{E}_t\left[p_{t+k}-\omega(\omega\, p_{t+k-1}-p_{t-1})+\widehat{mc}_{t+k}+\widehat{\mu}_{t+k}\right]\\ &= p_t+(1-\theta\beta)\sum_{k=0}^{\infty}(\theta\beta)^k\mathbb{E}_t\left[(p_{t+k}-p_t)-\omega(\omega\, p_{t+k-1}-p_{t-1})+\widehat{mc}_{t+k}+\widehat{\mu}_{t+k}\right]\end{aligned}$$

which can be rewritten as (15).

A.4. Aggregate Price Dynamics

Let S_t denote the subset of firms not reoptimizing at time t,

$$P_t=\left[\int_0^1 P_t(i)^{1-\epsilon_t}di\right]^{\frac{1}{1-\epsilon_t}}=\Bigg\{\underbrace{\int_{S_t}[P_{t-1}(i)\Pi_{t-1}^{\omega}]^{1-\epsilon_t}di}_{\Pi_{t-1}^{\omega(1-\epsilon_t)}\int_{S_t}P_{t-1}(i)^{1-\epsilon_t}di}+\int_{S_t^C}(P_t^*)^{1-\epsilon_t}di\Bigg\}^{\frac{1}{1-\epsilon_t}}=$$

$$=\left[\Pi_{t-1}^{\omega(1-\epsilon_t)}\theta\int_0^1 P_{t-1}(i)^{1-\epsilon_t}di+(1-\theta)\int_0^1(P_t^*)^{1-\epsilon_t}di\right]^{\frac{1}{1-\epsilon_t}}=\left[\Pi_{t-1}^{\omega(1-\epsilon_t)}\theta P_{t-1}^{1-\epsilon_t}+(1-\theta)(P_t^*)^{1-\epsilon_t}\right]^{\frac{1}{1-\epsilon_t}}$$

Moving the exponent from the RHS to the LHS, and dividing in both sides by $P_{t-1}^{1-\epsilon_t}$,

(32)
$$\left(\frac{P_t}{P_{t-1}}\right)^{1-\epsilon_t} = \Pi_{t-1}^{\omega(1-\epsilon_t)}\theta + (1-\theta)\left(\frac{P_t^*}{P_{t-1}}\right)^{1-\epsilon_t} \implies \Pi_t^{1-\epsilon_t} = \left(\frac{P_{t-1}}{P_{t-2}}\right)^{\omega(1-\epsilon_t)}\theta + (1-\theta)\left(\frac{P_t^*}{P_{t-1}}\right)^{1-\epsilon_t}$$

To simplify computation, I now log-linearize the left-hand side of (32),

$$\Pi_t^{1-\epsilon_t} \simeq \Pi^{1-\epsilon} + (1-\epsilon)\Pi^{-\epsilon}\underbrace{(\Pi_t - \Pi)}_{\pi_t} = 1 + (1-\epsilon)\pi_t$$

since $\Pi = \frac{P}{P} = 1$. A log-linearization of the right-hand side around a zero-inflation steady-state yields

$$\begin{aligned}
\left(\frac{P_{t-1}}{P_{t-2}}\right)^{\omega(1-\epsilon_t)}\theta + (1-\theta)\left(\frac{P_t^*}{P_{t-1}}\right)^{1-\epsilon_t} \simeq & \left(\frac{P}{P}\right)^{\omega(1-\epsilon)}\theta + (1-\theta)\left(\frac{P^*}{P}\right)^{1-\epsilon} + \\
& + (1-\theta)(1-\epsilon)P^{-\epsilon}P^{\epsilon-1}(P_t^* - P) + \\
& + \left[\theta\omega(1-\epsilon)P^{\omega(1-\epsilon)-1}P^{-\omega(1-\epsilon)} - (1-\theta)(1-\epsilon)P^{1-\epsilon}P^{2-\epsilon}\right] \times \\
& \times (P_{t-1} - P) - \theta\omega(1-\epsilon)P^{\omega(1-\epsilon)}P^{-\omega(1-\epsilon)-1}(P_{t-2} - P) = \\
= & \ \theta + 1 - \theta + (1-\theta)(1-\epsilon)\widehat{p}_t^* + \\
& + [\theta\omega(1-\epsilon) - (1-\theta)(1-\epsilon)]\widehat{p}_{t-1} - \theta\omega(1-\epsilon)\widehat{p}_{t-2} = \\
= & \ 1 + (1-\theta)(1-\epsilon)\widehat{p}_t^* - (1-\epsilon)[1-\theta(1+\omega)]\widehat{p}_{t-1} - \\
& - \theta\omega(1-\epsilon)\widehat{p}_{t-2} = \\
= & \ 1 + (1-\theta)(1-\epsilon)\,p_t^* - (1-\epsilon)[1-\theta(1+\omega)]\,p_{t-1} - \\
& - \theta\omega(1-\epsilon)\,p_{t-2}
\end{aligned}$$

Writing $\widehat{x}_t = x_t - x$, all log prices are cancelled out.

$$\begin{aligned}
\text{LHS=RHS:}\quad 1 + (1-\epsilon)\pi_t &= 1 + (1-\theta)(1-\epsilon)\,p_t^* - (1-\epsilon)[1-\theta(1+\omega)]\,p_{t-1} - \theta\omega(1-\epsilon)\,p_{t-2} \implies \\
\implies \pi_t &= (1-\theta)\,p_t^* - [1-\theta(1+\omega)]\,p_{t-1} - \theta\omega\,p_{t-2} \\
&= \theta\omega\pi_{t-1} + (1-\theta)(p_t^* - p_{t-1})
\end{aligned}$$

A.5. Deriving the Behavioural Hybrid New Keynesian Phillips Curve

Rewriting $\theta\beta\overline{m} = \delta$ and $\widetilde{\mathrm{mc}}_t = \widehat{\mathrm{mc}}_t + \widehat{\mu}_t$, the firm's problem optimality condition (16) reads

$$p_t^* = p_t + (1-\theta\beta)\sum_{k=0}^{\infty}\delta^k\mathbb{E}_t\left[\overline{m}(\pi_{t+1}+\ldots+\pi_{t+k}) - \omega\overline{m}(\pi_t+\ldots+\pi_{t+k-1}) + \overline{m}\widetilde{mc}_{t+k}\right]$$

(33)

$$= p_t + (1-\theta\beta)\mathbb{E}_t\left[\overline{m}\sum_{k=0}^{\infty}\delta^k(\pi_{t+1}+\ldots+\pi_{t+k}) - \omega\overline{m}\sum_{k=0}^{\infty}\delta^k(\pi_t+\ldots+\pi_{t+k-1}) + \overline{m}\sum_{k=0}^{\infty}\delta^k\widetilde{mc}_{t+k}\right]$$

We can calculate the following

$$H_t = \sum_{k=1}^{\infty}\delta^k(\pi_{t+1}+\ldots+\pi_{t+k}) = \sum_{j=1}^{\infty}\pi_{t+j}\sum_{k=j}^{\infty}\delta^k = \sum_{j=1}^{\infty}\pi_{t+j}\frac{\delta^j}{1-\delta} = \frac{1}{1-\delta}\sum_{j=1}^{\infty}\pi_{t+j}\delta^j =$$

$$= \frac{1}{1-\delta}\sum_{k=0}^{\infty}\pi_{t+k}\delta^k 1_{\{k>0\}}$$

$$\widetilde{H}_t = \sum_{k=1}^{\infty}\delta^k(\pi_t+\ldots+\pi_{t+k-1}) = \sum_{j=1}^{\infty}\pi_{t+j-1}\sum_{k=j}^{\infty}\delta^k = \sum_{j=1}^{\infty}\pi_{t+j-1}\frac{\delta^j}{1-\delta} = \frac{1}{1-\delta}\sum_{j=1}^{\infty}\pi_{t+j-1}\delta^j =$$

$$= \frac{1}{1-\delta}\sum_{k=0}^{\infty}\pi_{t+k-1}\delta^k 1_{\{k>0\}}$$

Rewriting (33),

$$p_t^* - p_t = (1-\theta\beta)\mathbb{E}_t\left[\overline{m}H_t - \omega\overline{m}\widetilde{H}_t + \overline{m}\sum_{k=0}^{\infty}\delta^k\widetilde{mc}_{t+k}\right]$$

$$= (1-\theta\beta)\mathbb{E}_t\left[\overline{m}\frac{1}{1-\delta}\sum_{k=0}^{\infty}\pi_{t+k}\delta^k 1_{\{k>0\}} - \omega\overline{m}\frac{1}{1-\delta}\sum_{k=0}^{\infty}\pi_{t+k-1}\delta^k 1_{\{k>0\}} + \overline{m}\sum_{k=0}^{\infty}\delta^k\widetilde{mc}_{t+k}\right]$$

$$= (1-\theta\beta)\mathbb{E}_t\sum_{k=0}^{\infty}\delta^k\left[\overline{m}\frac{1}{1-\delta}\pi_{t+k}1_{\{k>0\}} - \omega\overline{m}\frac{1}{1-\delta}\pi_{t+k-1}1_{\{k>0\}} + \overline{m}\widetilde{mc}_{t+k}\right]$$

$$= \mathbb{E}_t\sum_{k=0}^{\infty}\delta^k\left[\overline{m}\frac{1-\theta\beta}{1-\delta}\pi_{t+k}1_{\{k>0\}} - \omega\overline{m}\frac{1-\theta\beta}{1-\delta}\pi_{t+k-1}1_{\{k>0\}} + \overline{m}(1-\theta\beta)\widetilde{mc}_{t+k}\right]$$

(34)

$$= \mathbb{E}_t \sum_{k=0}^{\infty} \delta^k \left[\widetilde{m}_\pi \pi_{t+k} 1_{\{k>0\}} - \omega \widetilde{m}_\pi \pi_{t+k-1} 1_{\{k>0\}} + \widetilde{m}_\mu \widetilde{mc}_{t+k}\right]$$

where $\widetilde{m}_\pi = \overline{m}\frac{1-\theta\beta}{1-\delta}$ and $\widetilde{m}_\mu = \overline{m}(1-\theta\beta)$. Rewriting the price evolution expression (17),

$$p_t^* - p_{t-1} + p_t - p_t = \frac{\pi_t - \theta\omega\pi_{t-1}}{1-\theta} \implies p_t^* - p_t = \frac{\theta}{1-\theta}(\pi_t - \omega\pi_{t-1})$$

Hence, we can rewrite (34) as

$$\frac{\theta}{1-\theta}(\pi_t - \omega\pi_{t-1}) = \mathbb{E}_t \sum_{k=0}^{\infty} \delta^k \left[\widetilde{m}_\pi \pi_{t+k} 1_{\{k>0\}} - \omega \widetilde{m}_\pi \pi_{t+k-1} 1_{\{k>0\}} + \widetilde{m}_\mu \widetilde{mc}_{t+k}\right] \tag{35}$$

Let us now introduce the forward operator F such that $F^k x_t = x_{t+k}$. Using the forward operator, we can write

$$\sum_{k=0}^{\infty} \delta^k x_{t+k} = \sum_{k=0}^{\infty} \delta^k F^k x_t = \sum_{k=0}^{\infty} (\delta F)^k x_t = \frac{x_t}{1-\delta F} \tag{36}$$

Rewriting (35) using (36)

$$\begin{aligned}
\frac{\theta}{1-\theta}(\pi_t - \omega\pi_{t-1}) &= \widetilde{m}_\pi \mathbb{E}_t \left[\sum_{k=0}^{\infty} \delta^k \pi_{t+k} 1_{\{k>0\}}\right] - \omega \widetilde{m}_\pi \mathbb{E}_t \left[\sum_{k=0}^{\infty} \delta^k \pi_{t+k-1} 1_{\{k>0\}}\right] + \widetilde{m}_\mu \mathbb{E}_t \left[\sum_{k=0}^{\infty} \delta^k \widetilde{mc}_{t+k}\right] \\
&= \widetilde{m}_\pi \mathbb{E}_t \left[\sum_{k=0}^{\infty} (\delta F)^k \pi_t 1_{\{k>0\}}\right] - \omega \widetilde{m}_\pi \mathbb{E}_t \left[\sum_{k=0}^{\infty} (\delta F)^k \pi_{t-1} 1_{\{k>0\}}\right] + \widetilde{m}_\mu \mathbb{E}_t \left[\sum_{k=0}^{\infty} (\delta F)^k \widetilde{mc}_t\right] \\
&= \widetilde{m}_\pi \mathbb{E}_t \left[\sum_{k=0}^{\infty} (\delta F)^k \pi_t - \pi_t\right] - \omega \widetilde{m}_\pi \mathbb{E}_t \left[\sum_{k=0}^{\infty} (\delta F)^k \pi_{t-1} - \pi_{t-1}\right] + \widetilde{m}_\mu \mathbb{E}_t \left[\sum_{k=0}^{\infty} (\delta F)^k \widetilde{mc}_t\right] \\
&= \widetilde{m}_\pi \mathbb{E}_t \left[\frac{\pi_t}{1-\delta F} - \pi_t\right] - \omega \widetilde{m}_\pi \mathbb{E}_t \left[\frac{\pi_{t-1}}{1-\delta F} - \pi_{t-1}\right] + \widetilde{m}_\mu \mathbb{E}_t \left[\frac{\widetilde{mc}_t}{1-\delta F}\right] \\
&= \widetilde{m}_\pi \mathbb{E}_t \left[\frac{\delta F \pi_t}{1-\delta F}\right] - \omega \widetilde{m}_\pi \mathbb{E}_t \left[\frac{\delta F \pi_{t-1}}{1-\delta F}\right] + \widetilde{m}_\mu \mathbb{E}_t \left[\frac{\widetilde{mc}_t}{1-\delta F}\right]
\end{aligned}$$

Premultiplying by $(1-\delta F)$,

$$\frac{\theta}{1-\theta}(1-\delta F)(\pi_t - \omega\pi_{t-1}) = \widetilde{m}_\pi \mathbb{E}_t [\delta F \pi_t] - \omega \widetilde{m}_\pi \mathbb{E}_t [\delta F \pi_{t-1}] + \widetilde{m}_\mu \mathbb{E}_t \left[\widetilde{mc}_t\right]$$

which can be rearranged to (18). Let us now derive the Behavioural Hybrid New Keynesian Phillips curve. We have the following expressions

$$mc_t = w_t - p_t - a_t \tag{37}$$

$$y_t = a_t + n_t \tag{38}$$

$$w_t - p_t = \varphi n_t + \frac{\sigma}{1-h}c_t - \frac{\sigma h}{1-h}c_{t-1} \tag{39}$$

$$c_t = y_t \tag{40}$$

Hence, we can write

$$\begin{aligned} mc_t = w_t - p_t - a_t &= \varphi n_t + \frac{\sigma}{1-h}c_t - \frac{\sigma h}{1-h}c_{t-1} - a_t = \varphi(y_t - a_t) + \frac{\sigma}{1-h}c_t - \frac{\sigma h}{1-h}c_{t-1} - a_t \\ &= \varphi(y_t - a_t) + \frac{\sigma}{1-h}y_t - \frac{\sigma h}{1-h}y_{t-1} - a_t = \left(\varphi + \frac{\sigma}{1-h}\right) y_t - \frac{\sigma h}{1-h}y_{t-1} - (1+\varphi)a_t \end{aligned}$$

In the natural equilibrium (with no price frictions), the marginal cost is $mc_t^r = -\mu_t = \left(\varphi + \frac{\sigma}{1-h}\right) y_t^n - \frac{\sigma h}{1-h} y_{t-1}^n - (1+\varphi)a_t$, which we can rewrite as

$$y_t^n = \frac{\sigma h}{\varphi(1-h)+\sigma} y_{t-1}^n - \frac{1-h}{\varphi(1-h)+\sigma}\mu_t + \frac{(1+\varphi)(1-h)}{\varphi(1-h)+\sigma}a_t \tag{41}$$

hence, we can write $\widehat{mc}_t = mc_t - mc = \left(\varphi + \frac{\sigma}{1-h}\right)\widetilde{y}_t - \frac{\sigma h}{1-h}\widetilde{y}_{t-1}$ which, inserted into the (18), yields the Behavioural Hybrid New Keynesian Phillips curve (20).

A.6. Intrinsic Myopia

Consider instead a framework without habit formation nor price indexation, but with the following belief formation process. For any variable $\widehat{x}_t$, the BR forecast of such variable at horizon h is instead given by $\mathbb{E}_t^B \widehat{x}_{t+h} = \overline{m}^h \mathbb{E}_t \widehat{x}_{t+h} + (1-\overline{m}^h)\widehat{x}_{t-1}$ for $h > 0$, which implies that expectations of objects more in the future are more anchored to the past.

Let us first derive the DIS curve. Starting from (27) without habit formation, we can write $\widehat{c}_t = -\frac{1}{\sigma}(i_t - \rho - \mathbb{E}_t^B \pi_{t+1}) + \mathbb{E}_t^B \widehat{c}_{t+1} = -\frac{1}{\sigma}\left[i_t - \rho - \overline{m}\mathbb{E}_t \pi_{t+1} - (1-\overline{m})\pi_{t-1}\right] + \overline{m}\mathbb{E}_t \widehat{c}_{t+1} + (1-\overline{m})\widehat{c}_{t-1}$, and we can finally write the intrinsic myopia DIS curve as

$$\widetilde{y}_t = -\frac{1}{\sigma}\left[\widehat{i}_t - \overline{m}\mathbb{E}_t\pi_{t+1} - (1-\overline{m})\pi_{t-1} - r_t^n\right] + \overline{m}\mathbb{E}_t\widetilde{y}_{t+1} + (1-\overline{m})\widetilde{y}_{t-1}$$

Let us move to the supply side. Rewriting condition (15) without price indexation, but with intrinsic myopia,

$$
\begin{aligned}
p_t^* - p_{t-1} &= (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_t^B \widetilde{\mathrm{mc}}_{t+k} + \sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_t^B \pi_{t+k} + (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_t^B \widehat{\mu}_{t+k} \\
&= \frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\widetilde{\mathrm{mc}}_{t-1} + \frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\pi_{t-1} + (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta\overline{m})^k \mathbb{E}_t \widetilde{\mathrm{mc}}_{t+k} + \sum_{k=0}^{\infty}(\beta\theta\overline{m})^k \mathbb{E}_t \pi_{t+k}
\end{aligned}
$$

where we have used $\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_t^B \widehat{x}_{t+h} = \sum_{k=0}^{\infty}(\beta\theta\overline{m})^k \mathbb{E}_t \widehat{x}_{t+k} + \frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\widehat{x}_{t-1}$. If we take out past and present elements of each summation operator, the equation can be written more compactly as a difference equation,

$$
\begin{aligned}
p_t^* - p_{t-1} &= \frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\widetilde{\mathrm{mc}}_{t-1} + \frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\pi_{t-1} + (1-\beta\theta)\widetilde{\mathrm{mc}}_t + \pi_t \\
&+ (1-\beta\theta)\sum_{k=1}^{\infty}(\beta\theta\overline{m})^k \mathbb{E}_t \widetilde{\mathrm{mc}}_{t+k} + \sum_{k=1}^{\infty}(\beta\theta\overline{m})^k \mathbb{E}_t \pi_{t+k} \\
&= \frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\widetilde{\mathrm{mc}}_{t-1} + \frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\pi_{t-1} + (1-\beta\theta)\widetilde{\mathrm{mc}}_t + \pi_t \\
&+ \beta\theta\overline{m}\left[(1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta\overline{m})^k \mathbb{E}_t \widetilde{\mathrm{mc}}_{t+k+1} + \sum_{k=0}^{\infty}(\beta\theta\overline{m})^k \mathbb{E}_t \pi_{t+k+1}\right] \\
&= \frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\widetilde{\mathrm{mc}}_{t-1} + \frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\pi_{t-1} + (1-\beta\theta)\widetilde{\mathrm{mc}}_t + \pi_t \\
&+ \beta\theta\overline{m}\left[\mathbb{E}_t(p_{t+1}^* - p_t) - \frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\widetilde{\mathrm{mc}}_t - \frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\pi_t\right] \\
&= \beta\theta\overline{m}\mathbb{E}_t(p_{t+1}^* - p_t) + \left[1-\beta\theta-\beta\theta\overline{m}\frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\right]\widetilde{\mathrm{mc}}_t + \left[1-\beta\theta\overline{m}\frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\right]\pi_t \\
&+ \frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\widetilde{\mathrm{mc}}_{t-1} + \frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\pi_{t-1}
\end{aligned}
$$

Inserting the aggregate price dynamics (17), we can write

$$
\begin{aligned}
\frac{1}{1-\theta}\pi_t = \frac{\beta\theta\overline{m}}{1-\theta}\mathbb{E}_t\pi_{t+1} &+ \left[1-\beta\theta-\beta\theta\overline{m}\frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\right]\widetilde{\mathrm{mc}}_t + \left[1-\beta\theta\overline{m}\frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\right]\pi_t \\
&+ \frac{\beta\theta(1-\overline{m})}{1-\beta\theta\overline{m}}\widetilde{\mathrm{mc}}_{t-1} + \frac{\beta\theta(1-\overline{m})}{(1-\beta\theta)(1-\beta\theta\overline{m})}\pi_{t-1}
\end{aligned}
$$

rearranging elements,

$$\begin{aligned}
\pi_t = & \frac{\overline{m}\beta(1-\beta\theta)(1-\beta\theta\overline{m})}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}\mathbb{E}_t\pi_{t+1} + \frac{(1-\theta)(1-\beta\theta)\{1-\beta\theta[1+\overline{m}(1-\beta\theta\overline{m})]\}}{\theta\{1-\beta\theta[1+\overline{m}-\beta\overline{m}(1-\overline{m})]+\overline{m}^2\beta^2\theta^2\}}\widetilde{\mathrm{mc}}_t \\
& + \frac{(1-\overline{m})\beta(1-\beta\theta)(1-\theta)}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}\widetilde{\mathrm{mc}}_{t-1} + \frac{(1-\overline{m})\beta(1-\theta)}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}\pi_{t-1} \\
= & \frac{\overline{m}\beta(1-\beta\theta)(1-\beta\theta\overline{m})}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}\mathbb{E}_t\pi_{t+1} + \frac{(1-\theta)(1-\beta\theta)\{1-\beta\theta[1+\overline{m}(1-\beta\theta\overline{m})]\}}{\theta\{1-\beta\theta[1+\overline{m}-\beta\overline{m}(1-\overline{m})]+\overline{m}^2\beta^2\theta^2\}}(\widehat{\mathrm{mc}}_t+\widehat{\mu}_t) \\
& + \frac{(1-\overline{m})\beta(1-\beta\theta)(1-\theta)}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}(\widehat{\mathrm{mc}}_{t-1}+\widehat{\mu}_{t-1}) + \frac{(1-\overline{m})\beta(1-\theta)}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}\pi_{t-1}
\end{aligned}$$

Finally, introducing (19) without habit formation, we can write the intrinsic myopia NK Phillips curve,

$$\begin{aligned}
\pi_t = & \frac{\overline{m}\beta(1-\beta\theta)(1-\beta\theta\overline{m})}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}\mathbb{E}_t\pi_{t+1} + \frac{(1-\theta)(1-\beta\theta)\{1-\beta\theta[1+\overline{m}(1-\beta\theta\overline{m})]\}}{\theta\{1-\beta\theta[1+\overline{m}-\beta\overline{m}(1-\overline{m})]+\overline{m}^2\beta^2\theta^2\}}[(\sigma+\varphi)\widetilde{y}_t+\widehat{\mu}_t] \\
& + \frac{(1-\overline{m})\beta(1-\beta\theta)(1-\theta)}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}[(\sigma+\varphi)\widetilde{y}_{t-1}+\widehat{\mu}_{t-1}] + \frac{(1-\overline{m})\beta(1-\theta)}{1-\beta\theta\{1+\overline{m}[1-\beta+\beta\overline{m}(1-\theta)]\}}\pi_{t-1}
\end{aligned}$$

A.7. Robustness Checks

	Prior Distribution		Posterior Distribution			
		Mean (S.d)	**GDP Deflator**	**Estimating $\overline{m}$**	**1985:I-2007:III**	**Targeting Output**
β	*Beta*	0.99	0.990	0.990	0.990	0.990
		(0.001)	(0.989, 0.992)	(0.988, 0.992)	(0.988, 0.992)	(0.988, 0.992)
σ	*Normal*	1.50	1.280	1.239	1.261	1.353
		(0.37)	(0.703, 1.829)	(0.634, 1.777)	(0.663, 1.817)	(0.783, 1.917)
φ	*Normal*	2	1.425	1.446	1.409	1.446
		(0.75)	(0.500, 2.257)	(0.500, 2.306)	(0.500, 2.269)	(0.500, 2.298)
ϕ_π	*Normal*	1.50	1.430	1.346	1.388	1.352
		(0.15)	(1.219, 1.642)	(1.143, 1.547)	(1.134, 1.638)	(1.150, 1.547)
ϕ_y	*Normal*	0.15	0.340	0.333	0.288	0.336
		(0.10)	(0.231, 0.451)	(0.225, 0.444)	(0.134, 0.456)	(0.229, 0.445)
θ	*Beta*	0.50	0.930	0.904	0.873	0.922
		(0.10)	(0.912, 0.953)	(0.870, 0.939)	(0.828, 0.921)	(0.898, 0.948)
h	*Beta*	0.70	0.654	0.638	0.623	0.701
		(0.15)	(0.500, 0.813)	(0.473, 0.808)	(0.420, 0.816)	(0.559, 0.851)
ω	*Beta*	0.50	0.077	0.781	0.268	0.777
		(0.15)	(0.0243 0.129)	(0.707, 0.857)	(0.114, 0.419)	(0.705, 0.854)
$\overline{m}$	*Implied*	—	0.51	0.365	0.391	0.65
		—	(—)	(0.153, 0.568)	(0.167, 0.603)	(—)
ρ_i	*Beta*	0.50	0.853	0.861	0.934	0.859
		(0.20)	(0.815, 0.891)	(0.828, 0.895)	(0.904, 0.960)	(0.825, 0.893)
ρ_d	*Beta*	0.50	0.684	0.709	0.752	0.646
		(0.20)	(0.591, 0.783)	(0.610, 0.806)	(0.628, 0.874)	(0.546, 0.753)
ρ_s	*Beta*	0.50	0.830	0.069	0.070	0.068
		(0.20)	(0.770, 0.895)	(0.010, 0.124)	(0.041, 0.336)	(0.010, 0.125)
ρ_{e_i}	*Beta*	0.50	0.206	0.167	0.571	0.164
		(0.20)	(0.099, 0.313)	(0.066, 0.261)	(0.422, 0.723)	(0.065, 0.256)
σ_d	*Inv. gamma*	0.10	0.461	0.524	0.294	0.394
		(∞)	(0.416, 0.505)	(0.414, 0.630)	(0.217, 0.365)	(0.349, 0.439)
σ_s	*Inv. gamma*	0.10	0.175	0.339	0.329	0.285
		(∞)	(0.158, 0.193)	(0.284, 0.390)	(0.277, 0.381)	(0.259, 0.310)
σ_i	*Inv. gamma*	0.10	0.212	0.207	0.081	0.207
		(∞)	(0.194, 0.229)	(0.190, 0.224)	(0.071, 0.091)	(0.190, 0.223)
	Log data density		–283.853	–352.369	–24.153	–354.076

Note: Results are reported at the posterior mean. 90% confidence intervals in parenthesis. The model-implied forecast-underrevision coefficients are 1.2362 (column 4), 1.4432 (column 5), 0.5061 (column 6) and 0.7428 (column 7). In columns 5 and 6, which directly estimate $\overline{m}$, we assume a prior Beta distribution with mean (S.d.) of 0.50 (0.15). In column 7, the forecast-underrevision coefficient refers to the one implied by output gap (a value of 0.7523 in the data following Coibion and Gorodnichenko 2015).

TABLE 2. Estimated Structural Parameters: Robustness Checks

A.8. Smets and Wouters (2007) with Bounded Rationality

A.8.1. Theory

Throughout this section, we outline the differences between the large scale DSGE model in Smets and Wouters (2007) and its BR extension. We assume that all structural shocks follow the same stochastic processes as in Smets and Wouters (2007). We follow their notation for ease of comparison. For a detailed description of all parameters (except for $\overline{m}$), we refer the reader to Smets and Wouters (2007).

The aggregate resource constraint, [1] in their paper, is identical since it does not involve BR expectations. Their consumption Euler equation [2] is changed to

$$c_t = \frac{h/\overline{\gamma}}{1+h/\overline{\gamma}} c_{t-1} + \frac{\overline{m}}{1+h/\overline{\gamma}} \mathbb{E}_t c_{t+1} - \frac{(\sigma_c - 1) W_*^h L_*/C_*}{\sigma_c(1+h/\overline{\gamma})} (\overline{m}\mathbb{E}_t l_{t+1} - l_t) - \frac{1-h/\overline{\gamma}}{\sigma_c(1+h/\overline{\gamma})} (r_t - \overline{m}\mathbb{E}_t \widehat{\pi}_{t+1} + \varepsilon_t^b)$$

where $\overline{m}$ appears in front of the BR expectation operators. Their investment Euler equation [3] is changed to

$$i_t = \frac{1}{1+\beta\overline{\gamma}^{1-\sigma_c}} i_{t-1} + \frac{\beta\overline{\gamma}^{1-\sigma_c}\overline{m}}{1+\beta\overline{\gamma}^{1-\sigma_c}} \mathbb{E}_t i_{t+1} + \frac{1}{(1+\beta\overline{\gamma}^{1-\sigma})\overline{\gamma}^2\varphi} q_t + \varepsilon_t^i$$

where $\overline{m}$ appears in front of BR expected investment. Their arbitrage equation for the value of capital [4] is changed to

$$q_t = \beta\overline{\gamma}^{-\sigma_c}(1-\delta)\overline{m}\mathbb{E}_t q_{t+1} + [1-\beta\overline{\gamma}^{-\sigma_c}(1-\delta)]\overline{m}\mathbb{E}_t r_{t+1}^k - (r_t - \overline{m}\mathbb{E}_t \pi_{t+1} + \varepsilon_t^b)$$

where $\overline{m}$ appears in front of the BR expectation operators. The aggregate production function [5], capital laws of mation [6] and [8], capital utilization [7], and price mark-up [9] are identical since they do not involve BR expectations. The Smets and Wouters (2007) equivalent of our expression (16) (see pp. 18 in their Online Appendix) is given by

$$p_t^* = p_t + (1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}) \sum_{k=0}^{\infty} (\xi_p\beta\overline{\gamma}^{1-\sigma_c}\overline{m})^k \mathbb{E}_t \Big[(\pi_{t+1} + \cdots + \pi_{t+k}) - \omega(\pi_t + \cdots + \pi_{t+k-1}) + \frac{1}{(\phi_p - 1)\varepsilon_p + 1} \left(\widehat{\mathrm{mc}}_{t+k} + \widehat{\lambda}_{p,t+k} \right) \Big]$$

Price level dynamics in the Smets and Wouters (2007) follow the same dynamics as in (17). Following the steps in Appendix A.5, we can write the Smets and Wouters (2007) equivalent of our expression (18):

$$\pi_t = \pi_1\pi_{t-1} + \pi_2\mathbb{E}_t\pi_{t+1} - \pi_3\mu_t^p + \pi_3\widehat{\lambda}_{p,t}$$

which is the equivalent to condition [10] in their paper, where

$$\pi_1 = \frac{\iota_p}{1+\iota_p\beta\overline{\gamma}^{1-\sigma_c}\overline{m}\left[\xi_p + (1-\xi_p)\frac{1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}}{1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}\overline{m}}\right]}$$
$$\pi_2 = \frac{\beta\overline{\gamma}^{1-\sigma_c}\overline{m}\left[\xi_p + (1-\xi_p)\frac{1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}}{1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}\overline{m}}\right]}{1+\iota_p\beta\overline{\gamma}^{1-\sigma_c}\overline{m}\left[\xi_p + (1-\xi_p)\frac{1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}}{1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}\overline{m}}\right]}$$
$$\pi_3 = \frac{(1-\xi_p)(1-\xi_p\beta\overline{\gamma}^{1-\sigma_c})}{\xi_p[(\phi_p-1)\varepsilon_p+1]\left\{1+\beta\overline{\gamma}^{1-\sigma_c}\overline{m}\left[\xi_p + (1-\xi_p)\frac{1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}}{1-\xi_p\beta\overline{\gamma}^{1-\sigma_c}\overline{m}}\right]\right\}}$$

where in their paper $\varepsilon_t^p = \pi_3\widehat{\lambda}_{p,t}$. Similarly, the wage Phillips curve, equivalent to their condition [13], is given by

$$w_t = w_1 w_{t-1} + (1-w_1)(\mathbb{E}_t w_{t+1} + \mathbb{E}_t\pi_{t+1}) - w_2\pi_t + w_3\pi_{t-1} - w_4\mu_t^w + \varepsilon_t^w$$

where

$$w_1 = \frac{1}{1+\beta\overline{\gamma}^{1-\sigma_c}\overline{m}\left[\xi_w + (1-\xi_w)\frac{1-\xi_w\beta\overline{\gamma}^{1-\sigma_c}}{1-\xi_w\beta\overline{\gamma}^{1-\sigma_c}\overline{m}}\right]}$$
$$w_2 = \frac{1+\beta\overline{\gamma}^{1-\sigma_c}\iota_w}{1+\beta\overline{\gamma}^{1-\sigma_c}\overline{m}\left[\xi_w + (1-\xi_w)\frac{1-\xi_w\beta\overline{\gamma}^{1-\sigma_c}}{1-\xi_w\beta\overline{\gamma}^{1-\sigma_c}\overline{m}}\right]}$$
$$w_3 = \iota_w w_1$$
$$w_4 = \frac{(1-\xi_w)(1-\xi_w\beta\overline{\gamma}^{1-\sigma_c})}{\xi_w[(\phi_w-1)\varepsilon_w+1]\left\{1+\beta\overline{\gamma}^{1-\sigma_c}\overline{m}\left[\xi_w + (1-\xi_w)\frac{1-\xi_w\beta\overline{\gamma}^{1-\sigma_c}}{1-\xi_w\beta\overline{\gamma}^{1-\sigma_c}\overline{m}}\right]\right\}}$$

The relation between the rental rate of capital and the capital-capital labor ratio and the real wage rate is identical to [11] in Smets and Wouters (2007), and so is the wage mark-up [12] and the nominal interest rate rule [14], since they do not involve BR expectations.

A.8.2. Empirics

We estimate the medium-scale DSGE model in Smets and Wouters (2007), extended with BR in the previous section. The model equations are outlined in Appendix A.8.1, and the empirical findings are reported in tables 3-4. We report the estimation of the model when $\overline{m} = 1$ because the original replication codes contain an error (see the discussion by Johannes Pfeifer here), using the same priors. We also remove the exogenous MA shocks, which simplifies the computation of the forecast underrevision coefficient.

In terms of the model parameters, we find that the BR model produces estimates that are closer to those found in the literature. We focus our discussion on the parameters that Smets and Wouters (2007) mention as not so close to the data or the literature. The authors estimate a trend growth rate of 0.43, whereas the average growth rate of output per capita over the sample is 0.502. In our estimation under $\overline{m} = 1$, the discrepancy persists ($\overline{\gamma} = 0.386$). However, once we estimate $\overline{m}$ to match the forecast underreaction coefficient in Coibion and Gorodnichenko (2015), $\overline{\gamma} = 0.508$. The implied duration of prices in Smets and Wouters (2007) is 8.82 months, in the upper range of the estimates in the micro-data (Nakamura and Steinsson 2008). Under BR, we estimate a price duration of 6.32 months, in the mid-range of the estimates reported in their table 1. Finally, Smets and Wouters (2007) estimate an excessive sensitivity of nominal interest rates to inflation (2.04, far from the estimates in Clarida et al. 2000). In our estimation under $\overline{m} = 1$, the inconsistency lingers: we find a Taylor rule coefficient of 1.833. Under BR, the estimated coefficient is lowered to 1.472, closer to the standard value assumed in the literature.

We find that adding more realistic elements improves the fit. In particular, the log data density of the rational model is –1624.463, whereas the log data density of the BR model is –1642.878. However, this model comparison exercise does not put any weight on expectation data: since we are only comparing the model performance with data on endogenous variables, we are not considering a penalty for not matching the evidence on the positive co-movement between ex-ante forecast errors and forecast revisions (Coibion and Gorodnichenko 2015). For instance, the estimation of the BR model yields $\overline{m} = 0.28$, which in turn implies an underrevision coefficient of 1.222. The baseline forecast-underrevision reported in Coibion and Gorodnichenko (2015) is 1.2306.

	Prior Distribution		Posterior Distribution	
		Mean	*1966:I to 2004:IV*	
		(S.d)	**SW**	**BSW**
φ	*Normal*	4	4.341	2.262
		(1.50)	(2.4451, 6.2669)	(2.0000, 2.6143)
σ_c	*Normal*	1.50	1.526	0.513
		(0.375)	(1.1334, 1.9435)	(0.3845, 0.6369)
h	*Beta*	0.70	0.537	0.743
		(0.10)	(0.3947, 0.6435)	(0.6602, 0.8279)
ξ_w	*Beta*	0.50	0.936	0.799
		(0.15)	(0.9200, 0.9500)	(0.7330, 0.8672)
σ_l	*Normal*	2	2.333	2.169
		(0.75)	(1.3518, 3.2770)	(1.1040, 3.2399)
ξ_p	*Beta*	0.50	0.538	0.525
		(0.10)	(0.5000, 0.5837)	(0.5000, 0.5559)
ι_w	*Beta*	0.50	0.661	0.211
		(0.15)	(0.5427, 0.7805)	(0.1068, 0.3151)
ι_p	*Beta*	0.50	0.093	0.107
		(0.15)	(0.0301, 0.1523)	(0.0324, 0.1737)
ψ	*Beta*	0.50	0.496	0.314
		(0.15)	(0.3246, 0.6663)	(0.1384, 0.4860)
Φ	*Normal*	1.25	1.684	1.664
		(0.125)	(1.5560, 1.8042)	(1.5247, 1.8035)
r_π	*Normal*	1.50	1.833	1.472
		(0.25)	(1.5350, 2.1353)	(1.1088, 1.8093)
ρ	*Beta*	0.75	0.913	0.937
		(0.10)	(0.8882, 0.9373)	(0.9118, 0.9634)
r_y	*Normal*	0.125	0.154	0.103
		(0.05)	(0.0955, 0.2143)	(0.0339, 0.1705)
$r_{\Delta y}$	*Normal*	0.125	0.172	0.096
		(0.05)	(0.1344, 0.2111)	(0.0675, 0.1236)
$\overline{\pi}$	*Gamma*	0.625	0.932	0.735
		(0.10)	(0.7664, 1.1032)	(0.6218, 0.8484)
$100(\beta^{-1}-1)$	*Gamma*	0.25	0.216	0.254
		(0.10)	(0.0756, 0.2741)	(0.0979, 0.4011)
$\overline{l}$	*Normal*	0	0.298	−1.061
		(2)	(-1.4166, 2.1641)	(-3.0317, 0.9002)
$\overline{\gamma}$	*Normal*	0.40	0.216	0.508
		(0.10)	(0.3349, 0.4468)	(0.4838, 0.5311)
α	*Normal*	0.30	0.172	0.139
		(0.05)	(0.1400, 0.2029)	(0.1083, 0.1701)
$\overline{m}$	*Implied*	—	1	0.28
		—	(—)	(—)
	Log data density		−1624.463	−1642.878

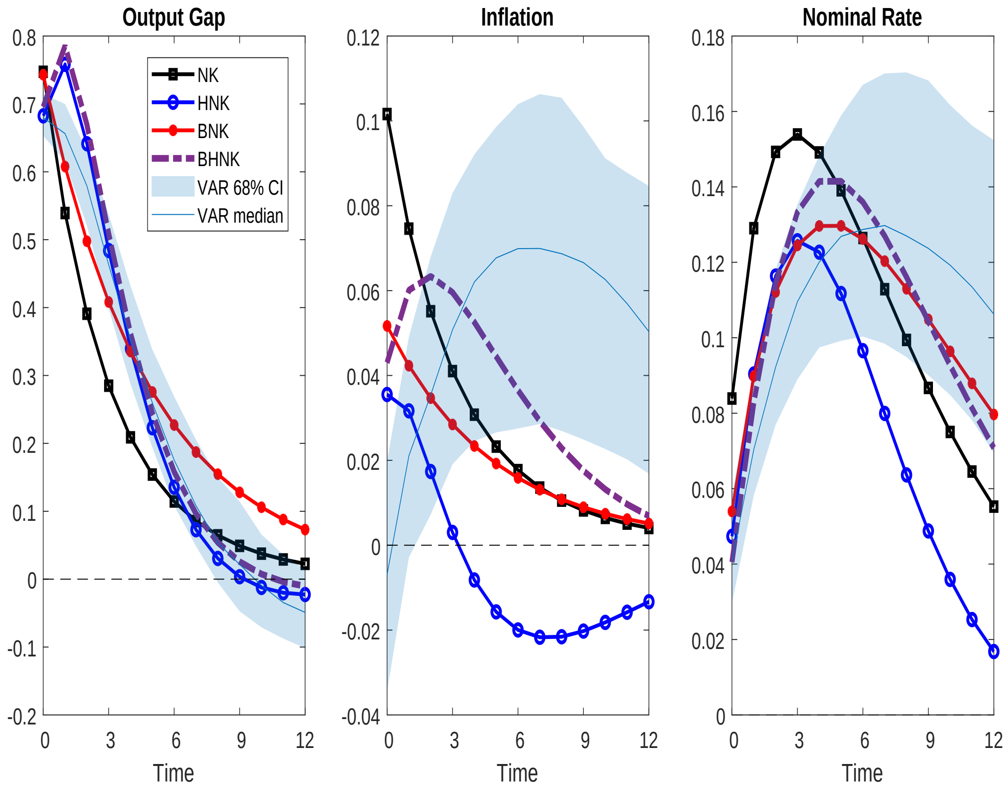
Note: Results are reported at the posterior mean. 90% confidence intervals in parenthesis. The model-implied forecast-underrevision coefficients are 0 (SW) and 1.2220 (BSW). The baseline forecast-underrevision reported in Coibion and Gorodnichenko (2015) is 1.2306.

TABLE 3. Estimated Structural Parameters: SW Models

	Prior Distribution		Posterior Distribution	
		Mean	*1966:I to 2004:IV*	
		(S.d)	**SW**	**BSW**
σ_a	*Inv. gamma*	0.10	0.519	0.521
		(2)	(0.4738, 0.5666)	(0.4754, 0.5655)
σ_b	*Inv. gamma*	0.10	0.146	0.816
		(2)	(0.0946, 0.1789)	(0.7251, 0.9015)
σ_g	*Inv. gamma*	0.10	0.742	0.743
		(2)	(0.6857, 0.7984)	(0.6833, 0.7990)
σ_I	*Inv. gamma*	0.10	0.453	1.360
		(2)	(0.3674, 0.5279)	(1.3604, 1.5980)
σ_r	*Inv. gamma*	0.10	0.228	0.208
		(2)	(0.2065 0.2479)	(0.1913, 0.2243)
σ_p	*Inv. gamma*	0.10	0.137	0.371
		(2)	(0.0909, 0.1829)	(0.3405, 0.4007)
σ_w	*Inv. gamma*	0.10	0.265	0.516
		(2)	(0.2321, 0.2996)	(0.4714, 0.5602)
ρ_a	*Beta*	0.50	0.994	0.972
		(0.20)	(0.9868, 0.9995)	(0.9544, 0.9904)
ρ_b	*Beta*	0.50	0.794	0.927
		(0.20)	(0.7314, 0.9164)	(0.8853, 0.9728)
ρ_g	*Beta*	0.50	0.993	0.975
		(0.20)	(0.9876, 0.9995)	(0.9571, 0.9930)
ρ_I	*Beta*	0.50	0.822	0.917
		(0.20)	(0.7258, 0.9255)	(0.8743, 0.9624)
ρ_r	*Beta*	0.50	0.143	0.161
		(0.20)	(0.0510, 0.2271)	(0.0656, 0.2549)
ρ_p	*Beta*	0.50	0.721	0.706
		(0.20)	(0.5911, 0.8425)	(0.6187, 0.8047)
ρ_w	*Beta*	0.50	0.176	0.465
		(0.20)	(0.0748, 0.2733)	(0.3522, 0.5780)

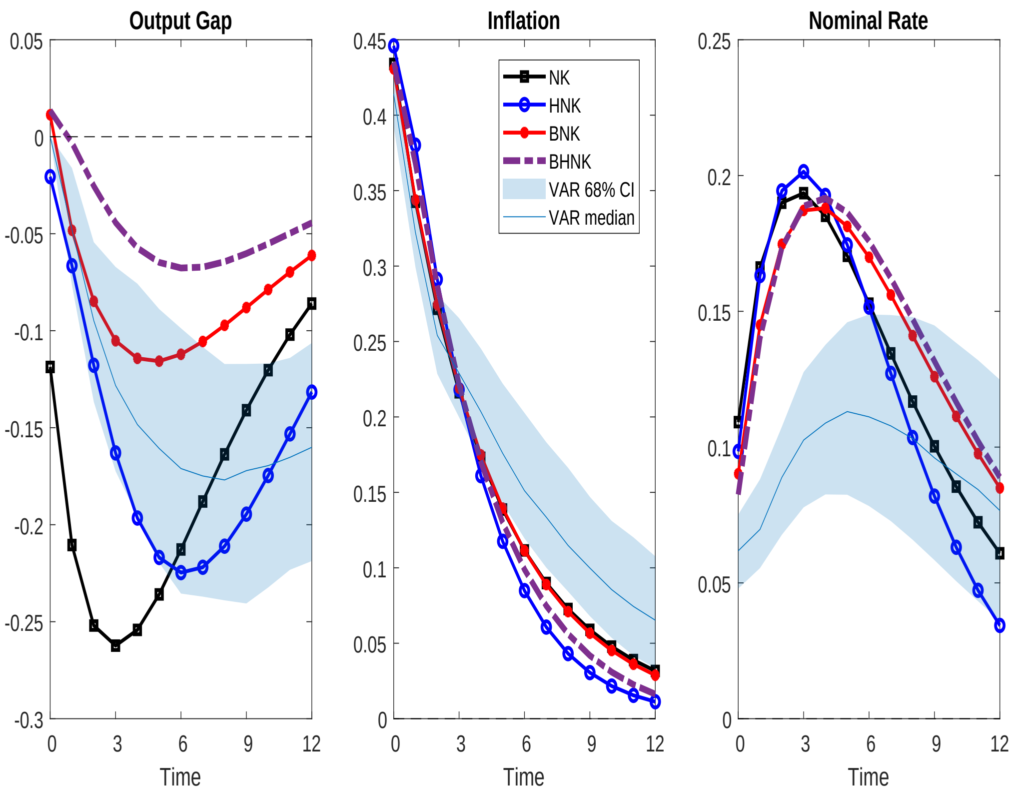
TABLE 4. Estimated Structural Parameters: SW Models

A.9. Impulse Response Functions to Demand and Supply Shocks



Note: The dynamic paths for the variables are reported under different model specifications after an aggregate demand shock: (i) a standard NK model in black lines (squares), (ii) a hybrid NK model in blue lines (circles), (iii) a behavioral NK model in red lines (asterisks), and (iv) a behavioral hybrid NK model in purple lines (dashed). The VAR-based demand shock is identified by means of a recursive identification, where the order of the variables is as follows: output gap, inflation and nominal interest rate. The horizontal axis displays the time which is measured in quarters. Vertical axis values refer to deviations from steady state in percentage.

FIGURE 2. Dynamic Responses to an Aggregate Demand Shock



Note: The dynamic paths for the variables are reported under different model specifications after an aggregate supply shock: (i) a standard NK model in black lines (squares), (ii) a hybrid NK model in blue lines (circles), (iii) a behavioral NK model in red lines (asterisks), and (iv) a behavioral hybrid NK model in purple lines (dashed). The VAR-based supply shock is identified by means of a recursive identification, where the order of the variables is as follows: output gap, inflation and nominal interest rate. The horizontal axis displays the time which is measured in quarters. Vertical axis values refer to deviations from steady state in percentage.

FIGURE 3. Dynamic Responses to an Aggregate Supply Shock

References

Clarida, Richard, Jordi Gali, and Mark Gertler, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *The Quarterly Journal of Economics*, 2000, *115* (1), 147–180.

Coibion, Olivier and Yuriy Gorodnichenko, "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 2015.

Nakamura, Emi and Jón Steinsson, "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics*, 2008, *123* (4), 1415–1464.

Smets, Frank and Rafael Wouters, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, June 2007, *97* (3), 586–606.