Online Appendix: Optimal Monetary Policy According to HANK

Sushant Acharya* Edouard Challe† Keshav Dogra‡

January 2023

A Proof of Proposition 1

The date s problem of an individual i born at date s can be written as:

$$\max_{\{c_t^s(i),\ell_t^s(i),a_{t+1}^s(i)\}} -\mathbb{E}_s \sum_{t=s}^{\infty} (\beta\vartheta)^{t-s} \left(\prod_{k=s}^{t-1} \zeta_k\right) \left\{ \frac{1}{\gamma} e^{-\gamma c_t^s(i)} + \rho e^{\frac{1}{\rho}[\ell_t^s(i)-\xi_t^s(i)]} \right\}$$

s.t.

$$c_t^s(i) + q_t a_{t+1}^s(i) \quad = \quad w_t \ell_t^s(i) + (1-\tau_t^a) a_t^s(i) + D_t - T_t \tag{A.1}$$

where $a_s^s(i) = 0$, $w_t = (1-\tau^w)\widetilde{w}_t$, $\tau_t^a = 0$ for $t > 0$ and ζ_t is the discount-factor shock introduced in Online Appendix J. The optimal labor supply decision of household i is given by:

$$\ell_t^s(i) = \rho \ln w_t - \gamma\rho c_t^s(i) + \xi_t^s(i) \tag{A.2}$$

and the Euler equation for all dates $t > 0$ is given by:

$$e^{-\gamma c_t^s(i)} = \beta \zeta_t R_t (1-\tau_{t+1}^a) \mathbb{E}_t e^{-\gamma c_{t+1}^s(i)} \tag{A.3}$$

where we have used the fact that $q_t = \frac{\vartheta}{R_t}$. Next, guess that the consumption decision rule takes the form:

$$c_t^s(i) = \mathcal{C}_t + \mu_t x_t^s(i) \tag{A.4}$$

where $x_t^s(i) = (1-\tau_t^a) a_t^s(i) + w_t \left(\xi_t^s(i) - \overline{\xi}\right)$ is de-meaned cash-on-hand and so, $x_{t+1}^s(i)$ is given by

$$x_{t+1}^s(i) = (1-\tau_{t+1}^a) a_{t+1}^s(i) + w_{t+1} \left(\xi_{t+1}^s(i) - \overline{\xi}\right)$$

*Bank of Canada and CEPR **Email:** sacharya@bankofcanada.ca
†European University Institute **Email:** edouard.challe@eui.eu
‡Federal Reserve Bank of New York **Email:** keshav.dogra@ny.frb.org

Substituting out tor $a^s_{t+1}(i)$ and $\ell^s_t(i)$ using (A.1) and (A.2), and using the definition of $x^s_t(i)$, the above expression can be written as

$$x^s_{t+1}(i) = \frac{(1-\tau^a_{t+1})R_t}{\vartheta}\Big\{x^s_t(i) + w_t\left(\rho \ln w_t + \bar{\xi}\right) + D_t - T_t - (1+\gamma\rho w_t)c^s_t(i)\Big\} + w_{t+1}\left(\xi^s_{t+1}(i) - \bar{\xi}\right)$$

Since $x^s_{t+1}(i)$ is normally distributed, given (A.4), $c^s_{t+1}(i)$ is also normally distributed with mean:

$$\mathbb{E}_t c^s_{t+1}(i) = \mathcal{C}_{t+1} + \mu_{t+1}\frac{(1-\tau^a_{t+1})R_t}{\vartheta}\left[x^s_t(i) + w_t\left(\rho \ln w_t + \bar{\xi}\right) + D_t - T_t - (1+\gamma\rho w_t)\, c^s_t(i)\right]$$

and variance:

$$\mathbb{V}_t\left(c^s_{t+1}(i)\right) = \mu^2_{t+1} w^2_{t+1} \sigma^2_{t+1}$$

Taking logs of (A.3) and using the two expressions above:

$$\begin{aligned}
c^s_t(i) &= -\frac{1}{\gamma}\ln[\beta\zeta_t R_t(1-\tau^a_{t+1})] - \frac{1}{\gamma}\ln \mathbb{E}_t e^{-\gamma c^s_{t+1}(i)} \\
&= -\frac{1}{\gamma}\ln[\beta\zeta_t R_t(1-\tau^a_{t+1})] + \mathbb{E}_t c^s_{t+1}(i) - \frac{\gamma}{2}\mathbb{V}_t\left(c^s_{t+1}(i)\right) \\
&= -\frac{1}{\gamma}\ln[\beta\zeta_t R_t(1-\tau^a_{t+1})] + \mathcal{C}_{t+1} \\
&\quad +\mu_{t+1}\frac{(1-\tau^a_{t+1})R_t}{\vartheta}\left[x^s_t(i) + w_t\left(\rho \ln w_t + \bar{\xi}\right) + D_t - T_t - (1+\gamma\rho w_t)\, c^s_t(i)\right] \\
&\quad -\frac{\gamma\mu^2_{t+1} w^2_{t+1}\sigma^2_{t+1}}{2}
\end{aligned}$$

Combining the $c^s_t(i)$ terms and using (A.4), the above can be rewritten as:

$$\begin{aligned}
\left[1 + \mu_{t+1}\frac{(1-\tau^a_{t+1})R_t}{\vartheta}(1+\gamma\rho w_t)\right] c^s_t(i) &= -\frac{1}{\gamma}\ln[\beta\zeta_t R_t(1-\tau^a_{t+1})] + \mathcal{C}_{t+1} - \frac{\gamma\mu^2_{t+1} w^2_{t+1}\sigma^2_{t+1}}{2} \\
&\quad +\mu_{t+1}\frac{(1-\tau^a_{t+1})R_t}{\vartheta}\left[x^s_t(i) + w_t\left(\rho \ln w_t + \bar{\xi}\right) + D_t - T_t\right]
\end{aligned}$$

Using $c^s_t(i) = \mathcal{C}_t + \mu_t x^s_t(i)$, we have:

$$\begin{aligned}
\left[1 + \mu_{t+1}\frac{(1-\tau^a_{t+1})R_t}{\vartheta}(1+\gamma\rho w_t)\right]\{\mathcal{C}_t + \mu_t x^s_t(i)\} &= -\frac{1}{\gamma}\ln[\beta\zeta_t R_t(1-\tau^a_{t+1})] + \mathcal{C}_{t+1} - \frac{\gamma\mu^2_{t+1} w^2_{t+1}\sigma^2_{t+1}}{2} \\
&\quad +\mu_{t+1}\frac{(1-\tau^a_{t+1})R_t}{\vartheta}\left[w_t\left(\rho \ln w_t + \bar{\xi}\right) + D_t - T_t\right] \\
&\quad +\mu_{t+1}\frac{(1-\tau^a_{t+1})R_t}{\vartheta} x^s_t(i)
\end{aligned}$$

Matching coefficients on $x^s_t(i)$, we have for all $t \geq 0$:

$$\mu^{-1}_t = 1 + \gamma\rho w_t + \frac{\vartheta}{(1-\tau^a_{t+1})R_t}\mu^{-1}_{t+1} \tag{A.5}$$

Notice that (A.5) is the same as (18) in the paper once we use the fact that $\tau^a_{t+1} = 0$ for all $t \geq 0$. Next, since the expression above must hold for all values of $x^s_t(i)$ including $x^s_t(i) = 0$, we have

$$\begin{aligned} \mathcal{C}_t &= -\frac{\vartheta\mu_t}{\mu_{t+1}(1-\tau^a_{t+1})R_t}\frac{1}{\gamma}\ln[\beta\zeta_t(1-\tau^a_{t+1})R_t] + \frac{\vartheta\mu_t}{\mu_{t+1}(1-\tau^a_{t+1})R_t}\mathcal{C}_{t+1} + \mu_t\left[w_t\left(\rho\ln w_t + \bar{\xi}\right) + D_t - T_t\right] \\ &\quad -\frac{\vartheta}{(1-\tau^a_{t+1})R_t}\frac{\mu_t}{\mu_{t+1}}\frac{\gamma\mu^2_{t+1}w^2_{t+1}\sigma^2_{t+1}}{2} \end{aligned} \tag{A.6}$$

Next, aggregate hours worked are given by $\ell_t = \rho\ln w_t - \gamma\rho\mathcal{C}_t + \bar{\xi}$ and hence aggregate income is $y_t = w_t\ell_t + D_t - T_t = w_t\rho\ln w_t - \gamma\rho w_t\mathcal{C}_t + w_t\bar{\xi} + D_t - T_t$. Using this in (A.6) together with $\mathcal{C}_t = y_t$ yields

$$\begin{aligned} \left[1-\mu_t(1+\gamma\rho w_t)\right]y_t &= -\frac{\vartheta\mu_t}{\mu_{t+1}(1-\tau^a_{t+1})R_t}\frac{1}{\gamma}\ln[\beta\zeta_t(1-\tau^a_{t+1})R_t] + \frac{\vartheta\mu_t}{\mu_{t+1}(1-\tau^a_{t+1})R_t}y_{t+1} \\ &\quad -\frac{\vartheta}{(1-\tau^a_{t+1})R_t}\frac{\mu_t}{\mu_{t+1}}\frac{\gamma\mu^2_{t+1}w^2_{t+1}\sigma^2_{t+1}}{2} \end{aligned}$$

Next, (A.5) implies that $1-\mu_t(1+\gamma\rho w_t) = \frac{\vartheta\mu_t}{\mu_{t+1}(1-\tau^a_{t+1})R_t}$ so dividing both sides of the equation above by $1-\mu_t(1+\gamma\rho w_t)$ yields

$$y_t = -\frac{1}{\gamma}\ln[\beta\zeta_t(1-\tau^a_{t+1})R_t] + y_{t+1} - \frac{\gamma\mu^2_{t+1}w^2_{t+1}\sigma^2_{t+1}}{2}$$

B Derivation of Σ recursion

B.1 Evolution of cash-on-hand within cohort

Given the consumption function and the definition of x, the evolution of cash on hand can be written as:

$$\begin{aligned} x^s_{t+1}(i) &= a^s_{t+1}(i) + w_{t+1}(\xi^s_{t+1}(i) - \bar{\xi}) \\ &= \frac{R_t}{\vartheta}\left[x^s_t(i) + w_t\left(\rho\ln w_t + \bar{\xi}\right) - T_t + D_t - \left(1+\rho\gamma w_t\right)y_t - \left(1+\rho\gamma w_t\right)\mu_t x^s_t(i)\right] + w_{t+1}(\xi^s_{t+1}(i) - \bar{\xi}) \\ &= \frac{R_t}{\vartheta}\left[1-\left(1+\rho\gamma w_t\right)\mu_t\right]x^s_t(i) + w_{t+1}(\xi^s_{t+1}(i) - \bar{\xi}) \end{aligned}$$

where we have used the fact that $\tau^a_t = 0$ for all dates $t > 0$. In the last line, we have used the definition of aggregate income $y_t = w_t(\rho\ln w_t - \gamma\rho y_t + \bar{\xi}) - T_t + D_t$. Multiplying both sides by μ_{t+1}:

$$\mu_{t+1}x^s_{t+1}(i) = \mu_{t+1}\frac{R_t}{\vartheta}\left[1-\left(1+\rho\gamma w_t\right)\mu_t\right]x^s_t(i) + \mu_{t+1}w_{t+1}(\xi^s_{t+1}(i) - \bar{\xi})$$

and using (18) in the paper, we have $\mu_{t+1}x^s_{t+1}(i) = \mu_t x^s_t(i) + \mu_{t+1}w_{t+1}(\xi^s_{t+1}(i) - \bar{\xi})$. That is, $\mu_t x^s_t(i)$ follows a random walk within cohort. This implies that in steady state with $\mu_t = \mu$, $x^s_t(i) \sim N\left(0, (t+1-s)w^2\sigma^2\right)$ and $a^s_t(i) \sim N\left(0, (t-s)w^2\sigma^2\right)$.

B.2 Objective function of planner

Substituting labor supply (16) in the paper into the objective function, we can write the date 0 expected utility of individual i from the cohort born at date s going forwards as:

$$W_0^s(i) = -\frac{1}{\gamma}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta_{t|0} \vartheta^t (1+\gamma\rho w_t) e^{-\gamma c_t^s(i)} = -\frac{1}{\gamma}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta_{t|0} \vartheta^t (1+\gamma\rho w_t) e^{-\gamma y_t - \gamma\mu_t x_t^s(i)}$$

where we have used the consumption function (15) in the paper and the fact that in equilibrium $\mathcal{C}_t = y_t$. We assume that the planner puts a weight of $\wp^s(i)$ on individual i born at date $s \leq 0$ and $\beta_{s|0} = \beta^s \prod_{k=0}^{s-1} \zeta_k$ on the lifetime welfare of individuals who will be born at date $s > 0$. Then the social welfare is:

$$\mathbb{W}_0 = \underbrace{(1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{-s} \int \wp^s(i) W_0^s(i) di}_{\text{welfare of those alive at date 0}} + \underbrace{(1-\vartheta) \sum_{s=1}^{\infty} \beta_{s|0} \int W_s^s(i) di}_{\text{welfare of the unborn at date 0}}$$

Using the definition of $W_0^s(i)$ and $W_s^s(i)$, notice that $\mathbb{W}_0$ can be written as:

$$\mathbb{W}_0 = -\frac{1}{\gamma} \sum_{t=0}^{\infty} \beta^t \underbrace{(1+\gamma\rho w_t) e^{-\gamma y_t}}_{\text{utility of rep. agent}} \Sigma_t$$

where Σ_t is defined as:

$$\begin{aligned} \Sigma_t &= (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{t-s} \int \wp^s(i) e^{-\gamma(c_t^s(i) - c_t)} di + (1-\vartheta) \sum_{s=1}^{t} \int \vartheta^{t-s} e^{-\gamma(c_t^s(i)-c_t)} di \\ &= (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{t-s} \int \wp^s(i) e^{-\gamma\mu_t x_t^s(i)} di + (1-\vartheta) \sum_{s=1}^{t} \int \vartheta^{t-s} e^{-\gamma\mu_t x_t^s(i)} di \end{aligned} \tag{B.1}$$

Thus, we can write $\mathbb{W}_0$ as:

$$\mathbb{W}_0 = \sum_{t=0}^{\infty} \beta_{t|0} \mathbb{U}_t \qquad \text{where} \qquad \mathbb{U}_t = -\frac{1}{\gamma}(1+\gamma\rho w_t) e^{-\gamma y_t} \Sigma_t$$

Next, we write (B.1) as:

$$\begin{aligned} \Sigma_t &= (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{t-s} \int \wp^s(i) e^{-\gamma\mu_t x_t^s(i)} di + (1-\vartheta) \sum_{s=1}^{t-1} \int \vartheta^{t-s} e^{-\gamma\mu_t x_t^s(i)} di + (1-\vartheta) \int e^{-\gamma\mu_t x_t^t(i)} di \\ &= \vartheta \left\{ (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{t-1-s} \int \wp^s(i) e^{-\gamma\mu_t x_t^s(i)} di + (1-\vartheta) \sum_{s=1}^{t-1} \int \vartheta^{t-1-s} e^{-\gamma\mu_t x_t^s(i)} di \right\} + (1-\vartheta) \int e^{-\gamma\mu_t x_t^t(i)} di \end{aligned}$$

Using $\mu_t x_t^s(i) = \mu_{t-1}x_{t-1}^s(i) + \mu_t w_t\left(\xi_t^s(i) - \overline{\xi}\right)$ from Appendix B.1:

$$\begin{aligned}
\Sigma_t &= \vartheta\Big\{ (1-\vartheta)\sum_{s=-\infty}^{0}\vartheta^{t-1-s}\int \wp^s(i)\, e^{-\gamma\left\{\mu_{t-1}x_{t-1}^s(i)+\mu_t w_t\left(\xi_t^s(i)-\overline{\xi}\right)\right\}}di \\
&\quad + (1-\vartheta)\sum_{s=1}^{t-1}\int \vartheta^{t-1-s}e^{-\gamma\left\{\mu_{t-1}x_{t-1}^s(i)+\mu_t w_t\left(\xi_t^s(i)-\overline{\xi}\right)\right\}}di\Big\} + (1-\vartheta)\int e^{-\gamma\mu_t x_t^t(i)}di \\
&= \vartheta e^{\frac{1}{2}\gamma^2\mu_t^2 w_t^2\sigma_t^2}\left\{(1-\vartheta)\sum_{s=-\infty}^{0}\vartheta^{t-1-s}\int \wp^s(i)\, e^{-\gamma\mu_{t-1}x_{t-1}^s(i)}di + (1-\vartheta)\sum_{s=1}^{t-1}\int \vartheta^{t-1-s}e^{-\gamma\mu_{t-1}x_{t-1}^s(i)}di\right\} \\
&\quad + (1-\vartheta)\int e^{-\gamma\mu_t x_t^t(i)}di \\
&= e^{\frac{1}{2}\gamma^2\mu_t^2 w_t^2\sigma_t^2}\left[1-\vartheta+\vartheta\Sigma_{t-1}\right]
\end{aligned}$$

Taking logs, this is the same for dates $t > 0$ as (27) in the paper. For date 0:

$$\begin{aligned}
\Sigma_0 &= (1-\vartheta)\sum_{s=-\infty}^{0}\vartheta^{-s}\int \wp^s(i)\, e^{-\gamma\mu_0 x_0^s(i)}di \\
&= (1-\vartheta)\sum_{s=-\infty}^{0}\vartheta^{-s}\int \wp^s(i)\, e^{-\gamma\mu_0\left(1-\tau_0^a\right)a_0^s(i)}e^{-\gamma\mu_0 w_0\left(\xi_0^s(i)-\overline{\xi}\right)}di \\
&= (1-\vartheta)\, e^{\frac{1}{2}\gamma^2\mu_0^2 w_0^2\sigma_0^2}\sum_{s=-\infty}^{0}\vartheta^{-s}\int \wp^s(i)\, e^{-\gamma\mu_0\left(1-\tau_0^a\right)a_0^s(i)}di
\end{aligned}$$

where we use the fact that $x_0^s(i) = (1-\tau_0^a)a_0^s(i) + w_0(\xi_0^s(i) - \overline{\xi})$. Next, we restrict $\wp^s(i) = e^{\gamma\alpha a_0^s(i)}$ where $\alpha \geq 0$ measures the planner's tolerance for pre-existing wealth inequality at date 0. Then we can write Σ_0 as:

$$\Sigma_0 = (1-\vartheta)\, e^{\frac{1}{2}\gamma^2\mu_0^2 w_0^2\sigma_0^2}\sum_{s=-\infty}^{0}\vartheta^{-s}\int e^{-\gamma\left[\alpha-\mu_0\left(1-\tau_0^a\right)\right]a_0^s(i)}di$$

Since $a_0^s(i) \sim N\left(0, -sw^2\sigma^2\right)$ for $s \leq 0$, this can be rewritten as:

$$\begin{aligned}
\Sigma_0 &= (1-\vartheta)\, e^{\frac{1}{2}\gamma^2\mu_0^2 w_0^2\sigma_0^2}\sum_{s=-\infty}^{0}\left(\vartheta e^{\frac{\gamma^2\mu^2 w^2\sigma^2}{2}\left[\frac{\alpha-\mu_0\left(1-\tau_0^a\right)}{\mu}\right]^2}\right)^{-s} \\
&= \frac{(1-\vartheta)\, e^{\frac{1}{2}\gamma^2\mu_0^2 w_0^2\sigma_0^2}}{1-\vartheta e^{\frac{\gamma^2\mu^2 w^2\sigma^2}{2}\left[\frac{\alpha-\mu_0\left(1-\tau_0^a\right)}{\mu}\right]^2}}
\end{aligned}$$

Taking logs, rewriting and using the definition $\Lambda = \gamma^2\mu^2 w^2\sigma^2$, this is the same as (45) in the paper and with $\alpha = 0$, this is the same as (28) in the paper.

B.2.1 The Utilitarian planner

The Utilitarian planner is one who assigns $\wp^s(i) = 1$ for all households alive at date 0. In this case the expression for Σ_0 can be simplified to:

$$\Sigma_t = (1-\vartheta)\sum_{s=-\infty}^{t}\vartheta^{t-s}e^{\frac{1}{2}\gamma^2\sigma_c^2(s,t)}$$

To see this, impose $\wp^s(i) = 1$ in (B.1), which can then be written as:

$$\Sigma_t = (1-\vartheta)\sum_{s=-\infty}^{t}\vartheta^{t-s}\int e^{-\gamma\mu_t x_t^s(i)}di$$

Given the consumption function (15) in the paper and the normality of shocks, the consumption of newly born individuals at any date s is normally distributed with mean y_s and variance $\sigma_c^2(s,s) = \mu_s^2 w_s^2 \sigma_s^2$ since they all have zero wealth. Given the linearity of the budget constraint, it follows that newly born agents' savings decisions $a_{s+1}^s(i)$ are also normally distributed with mean 0 and variance $\sigma_a^2(s+1,s) = \left(\frac{R_s}{\vartheta}\right)^2 [1-(1+\gamma\rho w_s)\mu_s]^2 w_s^2\sigma_s^2$. By induction, it follows that for any cohort born at date s, the cross-sectional distribution of consumption at any date $t > s$ is normal with mean y_t and variance

$$\sigma_c^2(t,s) = \mu_t^2\sigma_a^2(t,s) + \mu_t^2 w_t^2\sigma_t^2 \tag{B.2}$$

while the distribution of asset holdings is normal with mean 0 and variance

$$\sigma_a^2(t,s) = \frac{R_{t-1}^2}{\vartheta^2}[1-(1+\gamma\rho w_{t-1})\mu_{t-1}]^2\left[\sigma_a^2(t-1,s) + w_{t-1}^2\sigma_{t-1}^2\right] \tag{B.3}$$

C Some auxiliary results

In the proofs that follow, we shall make liberal use of the following assumptions and results.

Assumption 1. *Throughout the paper, we shall assume that:*

1. $\vartheta \geq \frac{1}{2}$
2. $\beta\vartheta > e^{-\frac{1}{2}} = 0.61$
3. $\sigma < \min\{\overline{\sigma}_1, \overline{\sigma}_2\}$ *where* $\overline{\sigma}_1 = \sqrt{\frac{2\rho^2\ln\vartheta^{-1}}{\left(\frac{\gamma\rho}{1+\gamma\rho}\right)^2\left(\frac{2\left(1-\frac{\wp}{\gamma}\right)\ln\vartheta^{-1}}{1+2\ln\vartheta}+(1-\beta)\right)^2}}$ *and* $\overline{\sigma}_2 = \frac{\rho}{\sqrt{(1-\beta\vartheta)(1+\gamma\rho+1-\beta\vartheta)}}$.

Lemma 1. *Given that* $\beta\vartheta > e^{-\frac{1}{2}}$*, we have* $\Lambda < 1$ *and* $\widetilde{\beta} < 1$.

Proof. Recall that in steady state, $\Lambda = \gamma^2\mu^2 w^2\sigma^2 > 0$, i.e.:

$$\Lambda = \frac{\sigma^2}{\rho^2}\left(\frac{\gamma\rho w}{1+\gamma\rho w}\right)^2\left(1-\beta\vartheta e^{\frac{\Lambda}{2}}\right)^2$$

Rearranging:

$$f(\Lambda) \equiv \frac{\Lambda}{\left(1 - \beta\vartheta e^{\frac{\Lambda}{2}}\right)^2} = \frac{\sigma^2}{\rho^2}\left(\frac{\gamma\rho w}{1+\gamma\rho w}\right)^2 \tag{C.1}$$

Now, $f(\Lambda)$ is increasing for $\Lambda < \Lambda^* \equiv -2\ln\beta\vartheta < 1$ given our assumption, and goes to ∞ as $\Lambda \to \Lambda^*$. For any values of σ and ρ, we can find some $0 < \overline{\Lambda} < \Lambda^*$ satisfying $f\left(\overline{\Lambda}\right) = \frac{\sigma^2}{\rho^2}$. Thus, any solution to (C.1) must satisfy $\Lambda \leq \overline{\Lambda} < \Lambda^* < 1$. By construction, for any $\Lambda < \Lambda^*$, $\widetilde{\beta} = \beta\theta e^{\frac{\Lambda}{2}} < 1$. $\square$

Lemma 2. *For $\sigma < [0, \overline{\sigma}_1)$, we have $\vartheta e^{\frac{\Lambda}{2}} < 1$.*

Proof. First we show that $\vartheta e^{\frac{\Lambda}{2}} = 1$ implies that $\sigma = \overline{\sigma}$. Starting from the expressions for wages in steady state, using $\vartheta e^{\frac{\Lambda}{2}} = 1$ we have:

$$\frac{w-1}{1+\gamma\rho w} = \frac{\Theta - 1 + \Lambda}{(1-\Lambda)(1-\widetilde{\beta})} = \frac{2\left(1 - \frac{\varphi}{\gamma}\right)\ln\vartheta^{-1}}{(1+2\ln\vartheta)(1-\beta)}$$

Add 1 to both sides and multiply by $\frac{\gamma\rho}{1+\gamma\rho}$ to get:

$$\frac{\gamma\rho w}{1+\gamma\rho w} = \left[\frac{2\ln\vartheta^{-1}\left(1 - \frac{\varphi}{\gamma}\right)}{(1+2\ln\vartheta)\left(1-\widetilde{\beta}\right)} + 1\right]\frac{\gamma\rho}{1+\gamma\rho}$$

Next, using the expression above in the definition of Λ, we have:

$$\sigma^2 = \frac{2\ln\vartheta^{-1}}{\left(\frac{\gamma\rho}{1+\gamma\rho}\right)^2\left(\frac{-2\ln\vartheta\left(1-\frac{\varphi}{\gamma}\right)}{(1+2\ln\vartheta)} + (1-\beta)\right)^2}$$

which is the same as $\overline{\sigma}_1$ defined in Assumption 1. Second, note that when $\sigma^2 = 0$, we have $\Lambda = 0$ and $\vartheta e^{\frac{\Lambda}{2}} = \vartheta < 1$. By continuity it follows that for $\sigma \in [0, \overline{\sigma}_1)$, we have $\vartheta e^{\frac{\Lambda}{2}} < 1$. $\square$

Corollary 1. *The following is true:*

$$1 - \beta^{-1}\widetilde{\beta}(1-\Lambda) > 0$$

Proof.

$$1 - \beta^{-1}\widetilde{\beta}(1-\Lambda) = 1 - \vartheta e^{\frac{\Lambda}{2}}(1-\Lambda) > 0$$

$\square$

D First-order condition of the planning problem

D.1 Optimally set fiscal instruments

The planner chooses τ_0^a and τ^w optimally absent aggregate shocks ($z_t = 1$ and $\varepsilon_t = \varepsilon \ \forall t$). This problem can be written as:

$$\max_{\{w_t, y_t, \mu_t, \Sigma_t, \Pi_t\}_{t=0}^{\infty}, \tau_0^a, \tau^w} \sum_{t=0}^{\infty} \beta^t \left\{ -\frac{1}{\gamma} \left(1 + \gamma \rho w_t\right) e^{-\gamma y_t} \Sigma_t \right\}$$

s.t.

$$\gamma y_t = \gamma y_{t+1} - \ln \beta \vartheta + \ln \mu_{t+1} + \ln \left[\mu_t^{-1} - (1 + \gamma \rho w_t)\right] - \frac{\gamma^2 \mu_{t+1}^2 w^2 \sigma^2}{2} e^{2\varphi(y_{t+1} - y)} \tag{D.1}$$

$$(\Pi_t - 1)\,\Pi_t = \frac{\varepsilon}{\Psi} \left[1 - \frac{1 - \tau^w}{w_t}\right] + \beta \left(\frac{y_{t+1} w_{t+1}}{y_t w_t}\right) (\Pi_{t+1} - 1)\,\Pi_{t+1} \tag{D.2}$$

$$\ln \Sigma_t = \frac{\gamma^2 \mu_t^2 w^2 \sigma^2}{2} e^{2\varphi(y_t - y)} + \ln \left[1 - \vartheta + \vartheta \Sigma_{t-1}\right] + \mathbb{I}(t = 0) \ln \left[\frac{1 - \vartheta e^{\frac{\Lambda}{2}}}{1 - \vartheta e^{\frac{\Lambda}{2}\left(\frac{\alpha - \left(1 - \tau_0^a\right)\mu_0}{\mu}\right)^2}} \right] \tag{D.3}$$

$$y_t = \frac{\rho \ln w_t + \overline{\xi}}{1 + \gamma \rho + \frac{\Psi}{2} (\Pi_t - 1)^2} \tag{D.4}$$

Let $M_{1,t}$ denote the multiplier on the date t aggregate Euler equation, $M_{2,t}$ that on the date t Phillips curve, $M_{3,t}$ that on the date t Σ recursion and $M_{4,t}$ that on the relationship between y_t, w_t and Π_t). The necessary conditions for optimality are as follows.
First-order condition with respect to w_t:

$$\mathbb{U}_t \frac{\gamma \rho w_t}{1 + \gamma \rho w_t} + M_{2,t-1} \left(\frac{y_t w_t}{y_{t-1} w_{t-1}}\right) (\Pi_t - 1)\,\Pi_t - M_{1,t} \frac{\gamma \rho w_t}{\mu_t^{-1} - (1 + \gamma \rho w_t)}$$
$$+ M_{2,t} \left\{ \frac{\varepsilon (1 - \tau^w)}{\Psi w_t} - \beta \left(\frac{y_{t+1} w_{t+1}}{y_t w_t}\right) (\Pi_{t+1} - 1)\,\Pi_{t+1} \right\} - \frac{M_{4,t}}{\gamma} \frac{\gamma \rho}{1 + \gamma \rho + \frac{\Psi}{2} (\Pi_t - 1)^2} = 0 \tag{D.5}$$

FOC wrt y_t:

$$-\gamma \mathbb{U}_t - \gamma M_{1,t} + \beta^{-1} M_{1,t-1} \left\{ \gamma - \varphi \gamma^2 \mu_t^2 w^2 \sigma^2 e^{2[\varphi(y_t - y)]} \right\} + M_{2,t-1} \left(\frac{w_t}{y_{t-1} w_{t-1}}\right) (\Pi_t - 1)\,\Pi_t$$
$$-\beta M_{2,t} \left(\frac{y_{t+1} w_{t+1}}{y_t^2 w_t}\right) (\Pi_{t+1} - 1)\Pi_{t+1} + M_{3,t} \varphi \gamma^2 \mu_t^2 w^2 \sigma^2 e^{2\varphi(y_t - y)} + M_{4,t} = 0 \tag{D.6}$$

FOC wrt μ_t:

$$-M_{1,t}\frac{\mu_t^{-1}}{\mu_t^{-1}-1-\gamma\rho w_t}+\beta^{-1}M_{1,t-1}\left[1-\gamma^2\sigma^2w^2\mu_t^2e^{2\varphi(y_t-y)}\right]+M_{3,t}\gamma^2\sigma^2w^2\mu_t^2e^{2\varphi(y_t-y)}$$
$$-\mathbb{I}(t=0)M_{3,0}\frac{\vartheta e^{\frac{\Lambda}{2}\left(\frac{\alpha-\left(1-\tau_0^a\right)\mu_0}{\mu}\right)^2}}{1-\vartheta e^{\frac{\Lambda}{2}\left(\frac{\alpha-\left(1-\tau_0^a\right)\mu_0}{\mu}\right)^2}}\Lambda\left(\frac{\alpha-\left(1-\tau_0^a\right)\mu_0}{\mu}\right)\left(1-\tau_0^a\right)\frac{\mu_0}{\mu}\quad=\quad 0 \tag{D.7}$$

FOC wrt Σ_t:

$$\mathbb{U}_t-M_{3,t}+\beta M_{3,t+1}\frac{\vartheta\Sigma_t}{1-\vartheta+\vartheta\Sigma_t}=0 \tag{D.8}$$

FOC wrt Π_t:

$$\left[\left(\frac{y_tw_t}{y_{t-1}w_{t-1}}\right)M_{2,t-1}-M_{2,t}\right](2\Pi_t-1)+\Psi M_{4,t}\frac{\rho\ln w_t+\overline{\xi}}{\left[1+\gamma\rho+\frac{\Psi}{2}\left(\Pi_t-1\right)^2\right]^2}(\Pi_t-1)=0 \tag{D.9}$$

FOC wrt τ_0^a:

$$M_{3,0}\frac{\vartheta e^{\frac{\Lambda}{2}\left(\frac{\alpha-\left(1-\tau_0^a\right)\mu_0}{\mu}\right)^2}}{1-\vartheta e^{\frac{\Lambda}{2}\left(\frac{\alpha-\left(1-\tau_0^a\right)\mu_0}{\mu}\right)^2}}\Lambda\left(\frac{\alpha-\left(1-\tau_0^a\right)\mu_0}{\mu}\right)\frac{\mu_0}{\mu}=0 \tag{D.10}$$

FOC wrt τ^w:

$$\sum_{t=0}^{\infty}\beta^t\frac{M_{2,t}}{w_t}=0 \tag{D.11}$$

We guess and verify that the optimal solution features $y_t=y, w_t=w, \mu_t=\mu$ and $\Pi_t=1$ such that $(1-\widetilde{\beta})\frac{w-1}{1+\gamma\rho w}=\Omega$. Plugging in the guesses into the FOCs, (D.9) implies $M_{2,t-1}=M_{2,t}$. Given this, (D.11) implies that $M_{2,t}=0$ for all $t\geq 0$. Using $\mu_t=\mu$ in (D.10), we have:

$$1-\tau_0^a=\frac{\alpha}{\mu}$$

as long as $M_{3,0}\neq 0$. In particular, if $\alpha=0$, i.e., the planner is utilitarian, we have $\tau_0^a=1$. Next, we show that $M_{3,0}\neq 0$. To see this, notice that (D.8) can be rewritten as:

$$1-\frac{M_{3,t}}{\mathbb{U}_t}+\beta\vartheta\frac{M_{3,t+1}}{\mathbb{U}_{t+1}}\frac{\Sigma_{t+1}}{1-\vartheta+\vartheta\Sigma_t}=0\qquad\Rightarrow\qquad\frac{M_{3,t}}{\mathbb{U}_t}=1+\widetilde{\beta}\frac{M_{3,t+1}}{\mathbb{U}_{t+1}} \tag{D.12}$$

where we have used the fact that $\mathbb{U}_{t+1}/\mathbb{U}_t=\Sigma_{t+1}/\Sigma_t$ and $\frac{\Sigma_{t+1}}{1-\vartheta+\vartheta\Sigma_t}=e^{\frac{\Lambda}{2}}$ since $y_t=y, w_t=w$ and $\mu_t=\mu$. Iterating forwards, we get $M_{3,t}/\mathbb{U}_t=(1-\widetilde{\beta})^{-1}\neq 0$.

Using this, (D.5), (D.6) and (D.7) become:

$$\frac{(1+\gamma\rho)w}{1+\gamma\rho w} + (1-\widetilde{\beta}^{-1})\frac{M_{1,t}}{\mathbb{U}_t}\frac{(1+\gamma\rho)w}{1+\gamma\rho w} - \frac{1}{\gamma}\frac{M_{4,t}}{\mathbb{U}_t} \quad = \quad 0 \tag{D.13}$$

$$-1 - \frac{M_{1,t}}{\mathbb{U}_t} + \beta^{-1}\frac{M_{1,t-1}}{\mathbb{U}_t}\left(1 - \frac{\varphi\Lambda}{\gamma}\right) + \frac{1}{1-\widetilde{\beta}}\frac{\varphi\Lambda}{\gamma} + \frac{1}{\gamma}\frac{M_{4,t}}{\mathbb{U}_t} \quad = \quad 0 \tag{D.14}$$

and

$$-\widetilde{\beta}^{-1}\frac{M_{1,t}}{\mathbb{U}_t} + \beta^{-1}(1-\Lambda)\frac{M_{1,t-1}}{\mathbb{U}_t} + \Lambda\frac{1}{1-\widetilde{\beta}} \quad = \quad 0 \tag{D.15}$$

where we have used $\mu^{-1} = \frac{1+\gamma\rho w}{1-\widetilde{\beta}}$. Next, combining (D.13) and (D.14), we get:

$$\frac{w-1}{1+\gamma\rho w} + \left[\left(1-\widetilde{\beta}^{-1}\right)\frac{w-1}{1+\gamma\rho w} - \widetilde{\beta}^{-1}\right]\frac{M_{1,t}}{\mathbb{U}_t} + \beta^{-1}\Theta\frac{M_{1,t-1}}{\mathbb{U}_t} + (1-\Theta)\frac{1}{1-\widetilde{\beta}} = 0 \tag{D.16}$$

Combining (D.15) with (D.16), we get:

$$\left[\frac{w-1}{1+\gamma\rho w} - \frac{\Theta - 1 + \Lambda}{\left(1-\widetilde{\beta}\right)(1-\Lambda)}\right]\left[1 + \beta^{-1}\left(1-\widetilde{\beta}\right)\frac{M_{1,t-1}}{\mathbb{U}_t}\right] = 0 \tag{D.17}$$

In particular, this must be true at date 0 when $M_{1,-1} = 0$. This requires:

$$\left(1-\widetilde{\beta}\right)\frac{w-1}{1+\gamma\rho w} = \frac{\Theta - 1 + \Lambda}{1-\Lambda}$$

which is the same as the definition of Ω in (29) in the main text. Given that w satisfies this restriction, (D.17) is also true at all subsequent dates. Since $\Pi = 1$, this implies from the Phillips curve that

$$1 - \tau^w = w^{-1} = \frac{1-\widetilde{\beta}+\Omega}{1-\widetilde{\beta}-\gamma\rho\Omega}$$

It follows that all FOCs and constraints are satisfied by our guesses and given the optimal values of τ_0^a and τ^w, the variables y_t, Π_t, μ_t, w_t remain at their steady state level absent aggregate shocks.

D.2 Steady state of the optimal plan

Imposing steady state on (D.3), one gets:

$$\Sigma = \frac{(1-\vartheta)\,e^{\frac{\Lambda}{2}}}{1-\vartheta e^{\frac{\Lambda}{2}}}$$

We already know from (D.12) in steady state that $m_3 = \frac{1}{1-\widetilde{\beta}}$ and that $m_2 = 0$ from (D.11) where $m_i = M_i/\mathbb{U}$ for $i = \{1, 2, 3, 4\}$. Next, imposing steady state in (D.15) yields:

$$m_1 = \frac{\widetilde{\beta}}{1-\widetilde{\beta}} \left[\frac{\Lambda}{1 - \beta^{-1}\widetilde{\beta}\,(1-\Lambda)} \right] \tag{D.18}$$

Notice that since $\Lambda = 0$ in RANK, we have $m_1 = 0$. Finally, using this in (D.13) and imposing steady state yields:

$$m_4 = \gamma \frac{\left(1 - \beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1 - \beta^{-1}\widetilde{\beta}\,(1-\Lambda)} \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \tag{D.19}$$

where $\Omega = \frac{\Theta - 1 + \Lambda}{1-\Lambda}$.

D.3 Optimal monetary policy given optimally set fiscal policy

The planning problem can be written as:

$$\max_{\{w_t, y_t, \mu_t, \Sigma_t, \Pi_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left(\prod_{k=0}^{t-1} \zeta_k \right) \left\{ -\frac{1}{\gamma} (1 + \gamma\rho w_t)\, e^{-\gamma y_t} \Sigma_t \right\} \tag{D.20}$$

s.t.

$$\gamma y_t = \gamma y_{t+1} - \ln \beta\vartheta - \ln \zeta_t + \ln \mu_{t+1} + \ln\left[\mu_t^{-1} - (1 + \gamma\rho w_t)\right] - \frac{\gamma^2 \mu_{t+1}^2 w^2 \sigma^2}{2} e^{2\varphi(y_{t+1}-y) + 2\varsigma_{t+1}} \tag{D.21}$$

$$(\Pi_t - 1)\,\Pi_t = \frac{\varepsilon_t}{\Psi}\left[1 - \frac{\varepsilon_t - 1}{\varepsilon_t} \frac{(1-\tau^w) z_t}{(1-\tau^\star) w_t}\right] + \beta \left(\frac{z_t w_{t+1} y_{t+1}}{z_{t+1} w_t y_t} \right) (\Pi_{t+1} - 1)\,\Pi_{t+1} \tag{D.22}$$

$$\ln \Sigma_t = \frac{\gamma^2 \mu_t^2 w^2 \sigma^2}{2} e^{2\{\varphi(y_t - y) + \varsigma_t\}} + \ln\left[1 - \vartheta + \vartheta \Sigma_{t-1}\right] + \mathbb{I}(t=0) \ln\left[\frac{1 - \vartheta e^{\frac{\Lambda}{2}}}{1 - \vartheta e^{\left(\frac{\alpha}{\mu}\right)^2 \frac{\Lambda}{2} \left(\frac{\mu - \mu_0}{\mu}\right)^2}} \right] \tag{D.23}$$

$$y_t = \frac{z_t \left(\rho \ln w_t + \overline{\xi}\right)}{1 + \gamma\rho z_t + \frac{\Psi}{2}(\Pi_t - 1)^2} \tag{D.24}$$

and $\Sigma_{-1} = 1$. The problem can be written as a Lagrangian:

$$\begin{aligned}
\mathcal{L} \;=\; & \sum_{t=0}^{\infty} \beta^t \left(\prod_{k=0}^{t-1} \zeta_k \right) \left\{ -\frac{1}{\gamma} \left(1 + \gamma\rho w_t\right) e^{-\gamma y_t} \Sigma_t \right\} \\
& + \sum_{t=0}^{\infty} \beta^t \left(\prod_{k=0}^{t-1} \zeta_k \right) M_{1,t} \Big\{ \gamma y_{t+1} - \ln \beta\vartheta - \ln \zeta_t + \ln \mu_{t+1} + \ln \left[\mu_t^{-1} - (1 + \gamma\rho w_t) \right] \\
& - \frac{\gamma^2 \mu_{t+1}^2 w^2 \sigma^2}{2} e^{2\varphi(y_{t+1} - y) + 2\varsigma_{t+1}} - \gamma y_t \Big\} \\
& + \sum_{t=0}^{\infty} \beta^t \left(\prod_{k=0}^{t-1} \zeta_k \right) M_{2,t} \left\{ \frac{\varepsilon_t}{\Psi} \left[1 - \frac{\varepsilon_t - 1}{\varepsilon_t} \frac{(1-\tau^w) z_t}{(1-\tau^\star) w_t} \right] + \beta \left(\frac{z_t w_{t+1} y_{t+1}}{z_{t+1} w_t y_t} \right) (\Pi_{t+1} - 1) \Pi_{t+1} - (\Pi_t - 1) \Pi_t \right\} \\
& + M_{3,0} \left\{ \frac{\gamma^2 \mu_0^2 w^2 \sigma^2}{2} e^{2\varphi(y_0 - y) + 2\varsigma_0} + \ln \left[1 - \vartheta + \vartheta \Sigma_{-1} \right] + \ln \left[\frac{1 - \vartheta e^{\frac{\Lambda}{2}}}{1 - \vartheta e^{\left(\frac{\alpha}{\mu}\right)^2 \frac{\Lambda}{2} \left(\frac{\mu - \mu_0}{\mu}\right)^2}} \right] - \ln \Sigma_0 \right\} \\
& + \sum_{t=1}^{\infty} \beta^t \left(\prod_{k=0}^{t-1} \zeta_k \right) M_{3,t} \left\{ \frac{\gamma^2 \mu_t^2 w^2 \sigma^2}{2} e^{2\varphi(y_t - y) + 2\varsigma_t} + \ln \left[1 - \vartheta + \vartheta \Sigma_{t-1} \right] - \ln \Sigma_t \right\} \\
& + \sum_{t=0}^{\infty} \beta^t \left(\prod_{k=0}^{t-1} \zeta_k \right) M_{4,t} \left\{ y_t - \frac{z_t \left(\rho \ln w_t + \bar{\xi} \right)}{1 + \gamma\rho z_t + \frac{\Psi}{2} (\Pi_t - 1)^2} \right\}
\end{aligned}$$

The optimal decisions satisfy:
FOC wrt w_t (multiplied through by w_t):

$$\begin{aligned}
& \mathbb{U}_t \frac{\gamma\rho w_t}{1 + \gamma\rho w_t} + \zeta_{t-1}^{-1} M_{2,t-1} \left(\frac{z_{t-1} y_t w_t}{z_t y_{t-1} w_{t-1}} \right) (\Pi_t - 1) \Pi_t - M_{1,t} \frac{\gamma\rho w_t}{\mu_t^{-1} - (1 + \gamma\rho w_t)} \\
+ M_{2,t} & \left\{ \frac{\varepsilon_t - 1}{\Psi} \frac{(1-\tau^w) z_t}{(1-\tau^\star) w_t} - \beta \left(\frac{z_t y_{t+1} w_{t+1}}{z_{t+1} y_t w_t} \right) (\Pi_{t+1} - 1) \Pi_{t+1} \right\} - \frac{M_{4,t}}{\gamma} z_t \frac{\gamma\rho}{1 + \gamma\rho z_t + \frac{\Psi}{2} (\Pi_t - 1)^2} \;=\; 0
\end{aligned} \tag{D.25}$$

FOC wrt y_t:

$$-\gamma \mathbb{U}_t - \gamma M_{1,t} + \beta^{-1} \zeta_{t-1}^{-1} M_{1,t-1} \left\{ \gamma - \varphi \gamma^2 \mu_t^2 w^2 \sigma^2 e^{2[\varphi(y_t - y) + \varsigma_t]} \right\} + \zeta_{t-1}^{-1} M_{2,t-1} \left(\frac{z_{t-1} w_t}{z_t y_{t-1} w_{t-1}} \right) (\Pi_t - 1) \Pi_t \tag{D.26}$$

$$-\beta M_{2,t} \left(\frac{z_t y_{t+1} w_{t+1}}{z_{t+1} y_t^2 w_t} \right) (\Pi_t - 1) \Pi_t + M_{3,t} \varphi \gamma^2 \mu_t^2 w^2 \sigma^2 e^{2\varphi(y_t - y) + 2\varsigma_t} + M_{4,t} \;=\; 0 \tag{D.27}$$

FOC wrt μ_t:

$$\begin{aligned}
& -M_{1,t} \frac{\mu_t^{-1}}{\mu_t^{-1} - 1 - \gamma\rho w_t} + \beta^{-1} \zeta_{t-1}^{-1} M_{1,t-1} \left[1 - \gamma^2 \sigma^2 w^2 \mu_t^2 e^{2\varphi(y_t - y) + 2\varsigma_t} \right] + M_{3,t} \gamma^2 \sigma^2 w^2 \mu_t^2 e^{2\varphi(y_t - y) + 2\varsigma_t} \\
& - \mathbb{I}(t=0) M_{3,0} \frac{\vartheta e^{\frac{\Lambda}{2} \left(\frac{\alpha}{\mu}\right)^2 \left(\frac{\mu - \mu_0}{\mu}\right)^2}}{1 - \vartheta e^{\frac{\Lambda}{2} \left(\frac{\alpha - \left(1 - \tau_0^a\right) \mu_0}{\mu} \right)^2}} \Lambda \left(\frac{\alpha}{\mu} \right)^2 \left(\frac{\mu - \mu_0}{\mu} \right) \frac{\mu_0}{\mu} \;=\; 0
\end{aligned} \tag{D.28}$$

FOC wrt Σ_t:

$$\mathbb{U}_t - M_{3,t} + \beta\zeta_t M_{3,t+1}\frac{\vartheta\Sigma_t}{1-\vartheta+\vartheta\Sigma_t} = 0 \tag{D.29}$$

FOC wrt Π_t:

$$\zeta_{t-1}^{-1}M_{2,t-1}\left(\frac{z_{t-1}w_t y_t}{z_t w_{t-1}y_{t-1}}\right)(2\Pi_t - 1) - M_{2,t}(2\Pi_t - 1) + \Psi M_{4,t}(\Pi_t - 1) = 0 \tag{D.30}$$

D.4 State contingent τ_0^a

Unlike in the main paper, if we allowed the planner to set τ_0^a in a state contingent fashion (varying with shocks), the optimality condition with respect to τ_0^a given by equation (D.10) holds for any μ_0, not just absent shocks. This implies that the tax is optimally set to

$$1 - \tau_0^{a\star} = \frac{\alpha}{\mu_0}$$

Consequently, (D.3) becomes

$$\ln\Sigma_t = \frac{\gamma^2\mu_t^2 w^2\sigma^2}{2}e^{2\varphi(y_t - y)} + \ln\left[1 - \vartheta + \vartheta\Sigma_{t-1}\right]$$

for any α at all dates $t \geq 0$. Since α does not appear explicitly in any of the other constraints or the objective function, it follows that the optimal path of all variables is the same as that chosen by the utilitarian planner.

E Local approximation

E.1 Log-linearized dynamic equations

All hatted variables denote log-deviations of from steady state, except for the hatted multipliers which denote deviations in levels. In the baseline model with all four shocks, the log-linearized equations describing aggregate dynamics are:

$$\begin{aligned}
\widehat{y}_t &= \Theta\widehat{y}_{t+1} - \frac{1}{\gamma y}\left(\widehat{i}_t - \pi_{t+1} + \widehat{\zeta}_t\right) - \frac{\Lambda}{\gamma y}\widehat{\mu}_{t+1} - \frac{\Lambda}{\gamma y}\widehat{\varsigma}_{t+1} && \text{(E.1)}\\
\widehat{\mu}_t &= -\left(1-\widetilde{\beta}\right)\frac{\gamma\rho w}{1+\gamma\rho w}\widehat{w}_t + \widetilde{\beta}\left(\widehat{\mu}_{t+1} + \widehat{i}_t - \pi_{t+1}\right) && \text{(E.2)}\\
\widehat{y}_t &= \frac{(\rho/y)}{1+\gamma\rho}\widehat{w}_t + \frac{1}{1+\gamma\rho}\widehat{z}_t && \text{(E.3)}\\
\pi_t &= \kappa\left(\widehat{y}_t - \widehat{y}_t^e\right) + \beta\pi_{t+1} + \frac{\varepsilon}{\Psi}\widehat{\varepsilon}_t && \text{(E.4)}
\end{aligned}$$

where $\kappa = \frac{\varepsilon(1+\gamma\rho)}{\Psi(\rho/y)}$. Using (E.2) and (E.3) to substitute out i_t and $\widehat{w}_t$ and using the fact that $\Omega = (1-\widetilde{\beta})\frac{w-1}{1+\gamma\rho w}$ and $1+\Omega = \frac{\Theta}{1-\Lambda}$, the IS equation (E.1) can be written as

$$\gamma y(1+\Omega)\widehat{y}_t + \widehat{\mu}_t = \widetilde{\beta}(1-\Lambda)\left\{\gamma y(1+\Omega)\widehat{y}_{t+1} + \widehat{\mu}_{t+1}\right\} - \widetilde{\beta}\widehat{\zeta}_t + \frac{\gamma y}{1+\gamma\rho}\left(1-\widetilde{\beta}+\Omega\right)\widehat{z}_t - \widetilde{\beta}\Lambda\widehat{\varsigma}_{t+1}$$

Solving this equation forwards yields:

$$\gamma y \left(1+\Omega\right) \widehat{y}_t + \widehat{\mu}_t = \Gamma_t \tag{E.5}$$

where

$$\begin{aligned}\Gamma_t &= \sum_{s=0}^{\infty} \widetilde{\beta}^s \left(1-\Lambda\right)^s \left\{ \frac{\gamma y}{1+\gamma\rho} \left(1-\widetilde{\beta}+\Omega\right) \widehat{z}_{t+s} - \widetilde{\beta}\widehat{\zeta}_{t+s} - \widetilde{\beta}\Lambda\widehat{\varsigma}_{t+1+s} \right\} \\ &= \frac{\gamma y}{1+\gamma\rho} \frac{1-\widetilde{\beta}+\Omega}{1-\widetilde{\beta}\varrho_z\left(1-\Lambda\right)} \widehat{z}_t - \frac{\widetilde{\beta}}{1-\widetilde{\beta}\varrho_\zeta\left(1-\Lambda\right)}\widehat{\zeta}_t - \frac{\widetilde{\beta}\varrho_\varsigma\Lambda}{1-\widetilde{\beta}\varrho_\varsigma\left(1-\Lambda\right)}\widehat{\varsigma}_t \end{aligned} \tag{E.6}$$

where we have used the fact that $\widehat{z}_{t+s} = \varrho_z^s \widehat{z}_t$, $\widehat{\zeta}_{t+s} = \varrho_\zeta^s \widehat{\zeta}_t$ and $\widehat{\varsigma}_{t+k} = \varrho_\varsigma^s \widehat{\varsigma}_t$ in the second equality. Next, the log-linearized Σ_t recursion is

$$\widehat{\Sigma}_t = -\gamma y \left(\Theta-1\right) \widehat{y}_t + \Lambda\left(\widehat{\mu}_t + \varsigma_t\right) + \beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1}$$

Using equation (E.5), we can substitute out $\widehat{\mu}_t$ from this expression

$$\widehat{\Sigma}_t = -\gamma y \left(\Theta-1\right) \widehat{y}_t + \Lambda\left[\Gamma_t - \gamma y\left(1+\Omega\right)\widehat{y}_t + \varsigma_t\right] + \beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1}$$

where Γ_t is defined in (E.6). Then we can write the log-linearized Σ_t recursion as

$$\widehat{\Sigma}_t = -\gamma y \Omega \widehat{y}_t + \Lambda\overline{\Gamma}_t + \beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1}$$

where $\overline{\Gamma}_t = \Gamma_z \widehat{z}_t + \Gamma_\zeta \widehat{\zeta}_t + \Gamma_\varsigma \widehat{\varsigma}_t$ where

$$\begin{aligned}\Gamma_z &= \frac{\gamma y}{1+\gamma\rho}\frac{1-\widetilde{\beta}+\Omega}{1-\widetilde{\beta}\left(1-\Lambda\right)\varrho_z} \\ \Gamma_\zeta &= -\frac{\widetilde{\beta}}{1-\widetilde{\beta}\left(1-\Lambda\right)\varrho_\zeta} \\ \Gamma_\varsigma &= \frac{1-\widetilde{\beta}\varrho_\varsigma}{1-\widetilde{\beta}\left(1-\Lambda\right)\varrho_\varsigma}\end{aligned}$$

Restricting attention to the case without demand shocks ($\widehat{\zeta}_t = \widehat{\varsigma}_t = 0$) as in our baseline model

$$\widehat{\Sigma}_t = -\gamma y \Omega \left[\widehat{y}_t - \underbrace{\left(\frac{1-\widetilde{\beta}+\Omega}{\Omega}\right)\frac{\Lambda}{1-\widetilde{\beta}\left(1-\Lambda\right)\varrho_z}\frac{1}{1+\rho/y}}_{=\varkappa(\Omega)} \widehat{y}_t^e \right] + \beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1}$$

where $\widehat{y}^e_t = \frac{1+\rho/y}{1+\gamma\rho}\widehat{z}_t$. When $\Omega \geq \Omega^c = \frac{\Lambda}{1-\Lambda}$, we clearly have $\varkappa(\Omega) > 0$; we also have

$$\begin{aligned}
\varkappa(\Omega) &= \left(\frac{1-\widetilde{\beta}+\Omega}{\Omega}\right)\frac{\Lambda}{1-\widetilde{\beta}(1-\Lambda)\varrho_z}\frac{1}{1+\rho/y} \\
&\leq \varkappa(\Omega^c) \\
&= \frac{1-\widetilde{\beta}(1-\Lambda)}{1-\widetilde{\beta}(1-\Lambda)\varrho_z}\frac{1}{1+\rho/y} \\
&< \frac{1}{1+\rho/y} < 1
\end{aligned}$$

Thus, for $\Omega \geq \Omega^c$ we have $\varkappa(\Omega) \in (0,1)$, as Lemma 1 claims.

E.2 Derivation of the Quadratic Loss function

As is well known, in the presence of a distorted steady state, maximizing a second-order approximation to the objective function (D.20) subject to first-order approximations of constraints (D.21)-(D.24), will not generally lead to a solution to the optimal policy problem which is accurate up to first-order. But following Benigno and Woodford (2005) and others, we obtain a valid linear-quadratic (LQ) approximation to the non-linear planning problem described in Appendix D.3 by using a second-order approximation of the constraints to eliminate the linear terms in the second-order approximations of the objective function.

Taking a second-order approximation to the planner's objective function $\mathbb{W}_0$, we have:[1]

$$\begin{aligned}
\mathbb{W}_0 \approx\ & \frac{\mathbb{U}}{1-\beta} \\
&+\mathbb{U}\sum_{t=0}^{\infty}\beta^t\left\{\frac{\gamma\rho w}{1+\gamma\rho w}\widehat{w}_t - \gamma y\widehat{y}_t + \widehat{\Sigma}_t + \frac{1}{2}(\gamma y)^2\widehat{y}_t^2 - \gamma y\widehat{y}_t\widehat{\Sigma}_t - \gamma y\frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\widehat{y}_t\widehat{w}_t + \frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\widehat{w}_t\widehat{\Sigma}_t\right\}
\end{aligned} \tag{E.7}$$

The second-order approximation to the IS curve at date t can be written as:

$$\begin{aligned}
\mathrm{g}_t^{\mathrm{IS}} =\ & \gamma y\Theta\widehat{y}_{t+1} + (1-\Lambda)\widehat{\mu}_{t+1} - \frac{1}{\widetilde{\beta}}\widehat{\mu}_t - \left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)\left(\frac{\gamma\rho w}{1+\gamma\rho w}\right)\widehat{w}_t - \gamma y\widehat{y}_t \\
&-(\gamma y)^2\frac{(1-\Theta)^2}{\Lambda}\widehat{y}_{t+1}^2 - 2\gamma y(1-\Theta)\widehat{\mu}_{t+1}\widehat{y}_{t+1} - \left(\frac{1+\Lambda}{2}\right)\widehat{\mu}_{t+1}^2 + \frac{1}{2}\left(2-\frac{1}{\widetilde{\beta}}\right)\frac{1}{\widetilde{\beta}}\widehat{\mu}_t^2 \\
&-\frac{1}{2}\left(\frac{1}{\widetilde{\beta}}\right)^2\left(\frac{\gamma\rho\left(1-\widetilde{\beta}+\Omega\right)}{1+\gamma\rho}\right)^2\widehat{w}_t^2 - \frac{1}{\widetilde{\beta}^2}\frac{\gamma\rho\left(1-\widetilde{\beta}+\Omega\right)}{1+\gamma\rho}\widehat{w}_t\widehat{\mu}_t
\end{aligned} \tag{E.8}$$

where we have used $\frac{\vartheta}{R_t} = \mu_{t+1}\left[\mu_t^{-1} - (1+\gamma\rho w_t)\right]$ to eliminate R_t. Next, since the steady state multiplier on the Phillips curve $M_2 = 0$, we can skip taking a second-order approximation of the Phillips curve. So,

[1] This approximation is valid for all specifications of Pareto weights considered in Sections 3, 4 and 5.2 in the main paper.

we proceed by taking a second-order approximation of the Σ_t recursion, we have:

$$\begin{aligned} \mathrm{g}_t^{\Sigma} \approx\ & \Lambda\widehat{\mu}_t + \gamma y(1-\Theta)\widehat{y}_t + \beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1} - \widehat{\Sigma}_t + \frac{1}{2}\widehat{\Sigma}_t^2 - \frac{1}{2}\left(\beta^{-1}\widetilde{\beta}\right)^2\widehat{\Sigma}_{t-1}^2 \\ & + (\gamma y)^2\frac{(1-\Theta)^2}{\Lambda}\widehat{y}_t^2 + 2\gamma y\,(1-\Theta)\,\widehat{\mu}_t\widehat{y}_t + \frac{\Lambda}{2}\widehat{\mu}_t^2 + \mathbb{I}(t=0)\frac{\vartheta}{1-\vartheta}\left(\frac{\alpha}{\mu}\right)^2\frac{\Lambda}{2}\widehat{\mu}_0^2 \end{aligned} \tag{E.9}$$

Finally, we can write the second-order approximation of (D.24) as:

$$\mathrm{g}_t^{y} \approx y\widehat{y}_t - \frac{y}{1+\gamma\rho}\widehat{z}_t - \frac{\rho}{1+\gamma\rho}\widehat{w}_t + \frac{1}{2}\frac{\rho}{1+\gamma\rho}\widehat{w}_t^2 - \frac{\rho}{(1+\gamma\rho)^2}\widehat{w}_t\widehat{z}_t + \frac{1}{2}\frac{\Psi y}{1+\gamma\rho}\pi_t^2 \tag{E.10}$$

Note that (E.8)-(E.10) equal 0 for any allocation satisfying the constraints up to second-order. Thus, we can use these equations together with the FOCs from the planner's problem absent shocks to eliminate first-order terms from the objective function (E.7). This yields the purely second-order approximation to (E.7):

$$\mathbb{W}_0 \approx \frac{\mathbb{U}}{1-\widetilde{\beta}} + \mathbb{U}\sum_{t=0}^{\infty}\beta^t\widetilde{\mathbb{U}}_t$$

where

$$\begin{aligned} \widetilde{\mathbb{U}}_t =\ & \frac{1}{2}(\gamma y)^2\widehat{y}_t^2 - \gamma y\widehat{y}_t\widehat{\Sigma}_t - \gamma y\frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\widehat{y}_t\widehat{w}_t + \frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\widehat{w}_t\widehat{\Sigma}_t \\ & + m_1\left\{-\beta^{-1}(\gamma y)^2\frac{(1-\Theta)^2}{\Lambda}\widehat{y}_t^2 - 2\beta^{-1}\gamma y\,(1-\Theta)\,\widehat{\mu}_t\widehat{y}_t - \beta^{-1}\left(\frac{1+\Lambda}{2}\right)\widehat{\mu}_t^2 + \frac{1}{2}\left(2-\frac{1}{\widetilde{\beta}}\right)\frac{1}{\widetilde{\beta}}\widehat{\mu}_t^2\right\} \\ & + m_1\left\{-\frac{1}{2}\left(\frac{1}{\widetilde{\beta}}\right)^2\left(\frac{\gamma\rho\left(1-\widetilde{\beta}+\Omega\right)}{1+\gamma\rho}\right)^2\widehat{w}_t^2 - \frac{1}{\widetilde{\beta}^2}\frac{\gamma\rho\left(1-\widetilde{\beta}+\Omega\right)}{1+\gamma\rho}\widehat{w}_t\widehat{\mu}_t\right\} \\ & + m_3\left\{\frac{1-\beta^{-1}\widetilde{\beta}^2}{2}\widehat{\Sigma}_t^2 + (\gamma y)^2\frac{(1-\Theta)^2}{\Lambda}\widehat{y}_t^2 + 2\gamma y\,(1-\Theta)\,\widehat{\mu}_t\widehat{y}_t + \frac{\Lambda}{2}\widehat{\mu}_t^2 + \mathbb{I}(t=0)\frac{\vartheta}{1-\vartheta}\left(\frac{\alpha}{\mu}\right)^2\frac{\Lambda}{2}\widehat{\mu}_0^2\right\} \\ & + m_4 y\left\{\frac{1}{2}\frac{\rho/y}{1+\gamma\rho}\widehat{w}_t^2 - \frac{\rho/y}{(1+\gamma\rho)^2}\widehat{w}_t\widehat{z}_t + \frac{1}{2}\frac{\Psi}{1+\gamma\rho}\pi_t^2\right\} \end{aligned} \tag{E.11}$$

where $m_i = M_i/\mathbb{U}$ denote the normalized steady state multipliers as above. Clearly maximizing $\mathbb{W}_0$ is equivalent to minimizing $\sum_{t=0}^{\infty}\beta^t\widetilde{\mathbb{U}}_t$ since $\mathbb{U} < 0$.

Using the expressions derived above for steady state multipliers and substituting out for $\widehat{w}_t$ using $\widehat{w}_t = \frac{1+\gamma\rho}{\rho/y}\widehat{y}_t - \frac{1}{\rho/y}\widehat{z}_t$ and $\widehat{\mu}_t$ using $\widehat{\mu}_t = \Gamma_z\widehat{z}_t - \gamma y\,(1+\Omega)\,\widehat{y}_t$, we can obtain a loss function in $\widehat{y}_t, \pi_t, z_t$ and

Σ_t (ignoring terms independent of policy) for $t > 0$:

$$
\begin{aligned}
\widetilde{\mathbb{U}}_t &= \frac{1}{2}\gamma y \left[\left(\frac{y}{\rho}\right) \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \left(1+\frac{\Omega}{1-\widetilde{\beta}}\right) \Upsilon(\Omega) + (\gamma y)\frac{\Omega^2}{1-\widetilde{\beta}} \right] \widehat{y}_t^2 \\
&\quad - \frac{\gamma y}{1+\gamma\rho} \left[\frac{\gamma y \Omega}{1-\widetilde{\beta}\varrho_z(1-\Lambda)} + \delta(\Omega)\Upsilon(\Omega)\left(\frac{y}{\rho}+1\right) \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \right] \left(1+\frac{\Omega}{1-\widetilde{\beta}}\right) \widehat{y}_t \widehat{z}_t \\
&\quad + \frac{1}{2}\left(\frac{1-\beta^{-1}\widetilde{\beta}^2}{1-\widetilde{\beta}}\right) \widehat{\Sigma}_t^2 + \gamma y \frac{\Omega}{1-\widetilde{\beta}} \widehat{y}_t \widehat{\Sigma}_t - \frac{\gamma y}{1+\gamma\rho}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right) \widehat{z}_t \widehat{\Sigma}_t \\
&\quad + \frac{1}{2} \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \left(1+\frac{\Omega}{1-\widetilde{\beta}}\right) \frac{\Psi\gamma y}{1+\gamma\rho} \pi_t^2
\end{aligned}
\tag{E.12}
$$

where $\Upsilon(\Omega)$ and $\delta(\Omega)$ are given by

$$
\Upsilon(\Omega) = 1 + \gamma\rho \frac{\Omega}{1-\widetilde{\beta}+\Omega} \left\{ \Omega\left(\frac{2}{\Lambda(1-\Lambda)} - 1\right) - 1 \right\} \tag{E.13}
$$

$$
\delta(\Omega) = \frac{1}{\Upsilon(\Omega)} \left[1 + \left(\frac{1+\Lambda}{1-\Lambda}\right) \frac{\gamma\rho\Omega}{1-\widetilde{\beta}\varrho_z(1-\Lambda)} \frac{1}{1+\rho/y} \right] \tag{E.14}
$$

For $t = 0$ we have:

$$
\begin{aligned}
\widetilde{\mathbb{U}}_0 &= \frac{1}{2}\gamma y \left[\left(\frac{y}{\rho}\right) \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \left(1+\frac{\Omega}{1-\widetilde{\beta}}\right) \Upsilon_0(\Omega) + (\gamma y)\frac{\Omega^2}{1-\widetilde{\beta}} \right] \widehat{y}_0^2 \\
&\quad - \frac{\gamma y}{1+\gamma\rho} \left[\frac{\gamma y \Omega}{1-\widetilde{\beta}\varrho_z(1-\Lambda)} + \delta_0(\Omega)\Upsilon_0(\Omega)\left(\frac{y}{\rho}+1\right) \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \right] \left(1+\frac{\Omega}{1-\widetilde{\beta}}\right) \widehat{y}_0 \widehat{z}_0 \\
&\quad + \frac{1}{2}\left(\frac{1-\beta^{-1}\widetilde{\beta}^2}{1-\widetilde{\beta}}\right) \widehat{\Sigma}_0^2 + \gamma y \frac{\Omega}{1-\widetilde{\beta}} \widehat{y}_0 \widehat{\Sigma}_0 - \frac{\gamma y}{1+\gamma\rho}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right) \widehat{z}_0 \widehat{\Sigma}_0 \\
&\quad + \frac{1}{2} \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \left(1+\frac{\Omega}{1-\widetilde{\beta}}\right) \frac{\Psi\gamma y}{1+\gamma\rho} \pi_0^2
\end{aligned}
\tag{E.15}
$$

where $\Upsilon_0(\Omega)$ and $\delta_0(\Omega)$ are given by

$$
\Upsilon_0(\Omega) = \Upsilon(\Omega) + (1+\Omega)\,\mathbb{G} \tag{E.16}
$$

$$
\delta_0(\Omega) = \frac{\Upsilon(\Omega)}{\Upsilon_0(\Omega)}\delta(\Omega) + \frac{1}{\Upsilon_0(\Omega)} \frac{1-\widetilde{\beta}+\Omega}{1-\widetilde{\beta}\varrho_z(1-\Lambda)} \frac{1}{1+\rho/y}\mathbb{G} \tag{E.17}
$$

where

$$
\mathbb{G} = \gamma\rho \left[\frac{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)} \right] \left[\frac{1+\Omega}{1-\widetilde{\beta}+\Omega} \right] \left(\frac{\alpha}{\mu}\right)^2 \left(\frac{\vartheta}{1-\vartheta}\right)\Lambda
$$

Notice that when $\alpha = 0$, $\mathbb{G} = 0$, $\Upsilon(\Omega) = \Upsilon_0(\Omega)$ and $\delta(\Omega) = \delta_0(\Omega)$, and there is no difference between the two expressions above. In principle, one could derive optimal policy by minimizing $\sum_{t=0}^{\infty} \beta^t \widetilde{\mathbb{U}}_t$ subject to the linearized Phillips curve (25) and the linearized Σ recursion (32). However, it is useful to use (32) to substitute out for $\widehat{\Sigma}_t$ and obtain a loss function purely in terms of $\widehat{y}_t, \pi_t$ and $\widehat{z}_t$. The terms involving $\widehat{\Sigma}_t$ in the objective function can be written as

$$L_\Sigma = \sum_{t=0}^{\infty} \beta^t \left\{ \frac{1}{2} \left(\frac{1-\beta^{-1}\widetilde{\beta}^2}{1-\widetilde{\beta}} \right) \widehat{\Sigma}_t^2 + \gamma y \frac{\Omega}{1-\widetilde{\beta}} \widehat{y}_t \widehat{\Sigma}_t - \frac{\gamma y}{1+\gamma\rho} \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \widehat{z}_t \widehat{\Sigma}_t \right\} \tag{E.18}$$

Next, solving (32) back to date -1 and using the definition of $\widehat{y}_t^e = \frac{1+\rho/y}{1+\gamma\rho} \widehat{z}_t$, we have

$$\widehat{\Sigma}_t = -\gamma y \Omega \sum_{k=0}^{t} \left(\frac{\widetilde{\beta}}{\beta}\right)^{t-k} \widehat{y}_k + \Lambda \frac{1-\widetilde{\beta}+\Omega}{1-\widetilde{\beta}\varrho_z(1-\Lambda)} \frac{\gamma y}{1+\gamma\rho} \sum_{k=0}^{t} \left(\frac{\widetilde{\beta}}{\beta}\right)^{t-k} \widehat{z}_k + \left(\frac{\widetilde{\beta}}{\beta}\right)^{t+1} \widehat{\Sigma}_{-1} \tag{E.19}$$

Substituting (E.19) into (E.18) yields, after some algebra

$$L_\Sigma = -\frac{1}{2}(\gamma y)^2 \frac{\Omega^2}{1-\widetilde{\beta}} \sum_{t=0}^{\infty} \beta^t \widehat{y}_t^2 + \frac{(\gamma y)^2}{1+\gamma\rho} \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \frac{\Omega}{1-\widetilde{\beta}\varrho_z(1-\Lambda)} \sum_{t=0}^{\infty} \beta^t \widehat{y}_t \widehat{z}_t$$

Substituting this expression into (E.12) and (E.15) yields the expression

$$\begin{aligned}
\widetilde{\mathbb{U}}_t \;=\; & \frac{\gamma y}{2} \left[\left(\frac{y}{\rho}\right) \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \Upsilon_0(\Omega) \right] \widehat{y}_0^2 \\
& + \frac{\gamma y}{2} \left[\left(\frac{y}{\rho}\right) \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \Upsilon(\Omega) \right] \sum_{t=1}^{\infty} \beta^t \widehat{y}_t^2 \\
& - \frac{\gamma y}{1+\gamma\rho} \left[\delta_0(\Omega) \Upsilon_0(\Omega) \left(\frac{y}{\rho}+1\right) \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \right] \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \widehat{y}_0 \widehat{z}_0 \\
& - \frac{\gamma y}{1+\gamma\rho} \left[\delta(\Omega) \Upsilon(\Omega) \left(\frac{y}{\rho}+1\right) \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \right] \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \sum_{t=1}^{\infty} \beta^t \widehat{y}_t \widehat{z}_t \\
& + \frac{\gamma y}{2} \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \frac{\Psi}{1+\gamma\rho} \sum_{t=0}^{\infty} \beta^t \pi_t^2
\end{aligned} \tag{E.20}$$

Dividing by $\frac{(1-\beta^{-1}\widetilde{\beta})(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)} \left(1 + \frac{\Omega}{1-\widetilde{\beta}}\right) \gamma y \left(\frac{y}{\rho}\right)$, using the fact that $\varepsilon/\kappa = \frac{\Psi(\rho/y)}{1+\gamma\rho}$ and using the definition of $\widehat{y}_t^e = \frac{1+\rho/y}{1+\gamma\rho} \widehat{z}_t$, yields the objective function in the main text in Proposition 9 in the paper

$$\frac{1}{2} \left\{ \Upsilon_0(\Omega) \left(\widehat{y}_0 - \delta_0(\Omega) \widehat{y}_0^e \right)^2 + \frac{\varepsilon}{\kappa} \pi_0^2 \right\} + \frac{1}{2} \sum_{t=1}^{\infty} \beta^t \left\{ \Upsilon(\Omega) \left(\widehat{y}_t - \delta(\Omega) \widehat{y}_t^e \right)^2 + \frac{\varepsilon}{\kappa} \pi_t^2 \right\} \tag{E.21}$$

For the utilitarian planner, $\Upsilon_0(\Omega) = \Upsilon(\Omega)$ and $\delta_0(\Omega) = \delta(\Omega)$ and the expression simplifies to the expression in Proposition 3 in the paper:

$$\frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left\{\Upsilon\left(\Omega\right)\left(\widehat{y}_t-\delta(\Omega)\widehat{y}_t^e\right)^2+\frac{\varepsilon}{\kappa}\pi_t^2\right\} \tag{E.22}$$

The optimal policy problem can now simply be specified as minimizing (E.21) subject to the linearized Phillips curve (30) in the paper. In Lagrangian form:

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\left\{\Upsilon_0\left(\Omega\right)\left(\widehat{y}_0-\delta_0(\Omega)\widehat{y}_0^e\right)^2+\frac{\varepsilon}{\kappa}\pi_0^2\right\}+\frac{1}{2}\sum_{t=1}^{\infty}\beta^t\left\{\Upsilon\left(\Omega\right)\left(\widehat{y}_t-\delta(\Omega)\widehat{y}_t^e\right)^2+\frac{\varepsilon}{\kappa}\pi_t^2\right\}\\ &+\sum_{t=0}^{\infty}\beta^t F_t\left\{\beta\pi_{t+1}+\kappa\left(\widehat{y}_t-\widehat{y}_t^e\right)+\frac{\varepsilon}{\Psi}\widehat{\varepsilon}_t-\pi_t\right\}\end{aligned}$$

The FOC w.r.t. $\widehat{y}_t$ can be written as:

$$\begin{aligned}\Upsilon_0(\Omega)\left(\widehat{y}_0-\delta_0(\Omega)\widehat{y}_0^e\right)+\kappa F_0=0 &\qquad \text{for } t=0\\ \Upsilon(\Omega)\left(\widehat{y}_t-\delta(\Omega)\widehat{y}_t^e\right)+\kappa F_t=0 &\qquad \text{for } t>0\end{aligned}$$

The FOC w.r.t. π_t can be written as

$$\frac{\varepsilon}{\kappa}\pi_t-F_t+F_{t-1}=0 \qquad \Leftrightarrow \qquad F_t=\frac{\varepsilon}{\kappa}\widehat{p}_t$$

where $\kappa=\frac{\varepsilon}{\Psi}\frac{1+\gamma\rho}{\rho/y}$. Combining the two FOCs we can derive the target criterion:

$$\begin{aligned}\widehat{y}_0-\delta_0(\Omega)\widehat{y}_0^e+\frac{\varepsilon}{\Upsilon_0(\Omega)}\widehat{p}_0=0 &\qquad \text{for } t=0\\ \widehat{y}_t-\delta(\Omega)\widehat{y}_t^e+\frac{\varepsilon}{\Upsilon(\Omega)}\widehat{p}_t=0 &\qquad \text{for } t>0\end{aligned}$$

E.3 Properties of loss function weights

Claim 1. $\Upsilon(\Omega)>1$ *with countercyclical risk*

Proof.

$$\begin{aligned}\Upsilon(\Omega) &= 1+\frac{\rho\gamma\Omega}{1-\widetilde{\beta}+\Omega}\left[\left(\frac{2}{\Lambda\left(1-\Lambda\right)}-1\right)\Omega-1\right]\\ &> 1+\frac{\rho\gamma\Omega}{1-\widetilde{\beta}+\Omega}\left[\left(\frac{2}{\Lambda\left(1-\Lambda\right)}-1\right)\frac{\Lambda}{1-\Lambda}-1\right]\end{aligned}$$

where we have used the fact that $\Omega=\frac{\Theta-1+\Lambda}{1-\Lambda}$ and for countercyclical risk ($\Theta>1$), we have $\Omega>\frac{\Lambda}{1-\Lambda}$. Then, the above can be simplified to:

$$\Upsilon(\Omega)>1+\frac{\rho\gamma\Omega}{1-\widetilde{\beta}+\Omega}\frac{1+\Lambda}{\left(1-\Lambda\right)^2}>1$$

□

Claim 2. $0 < \delta(\Omega) < 1$ *with countercyclical risk*

Proof. Using the expression for $\Upsilon(\Omega)$ in $\delta(\Omega)$, we have:

$$\delta(\Omega) = \frac{\left(1 - \widetilde{\beta} + \Omega\right)\left(1 + \frac{\gamma\rho\Omega}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{1}{1+(\rho/y)}\left(\frac{1+\Lambda}{1-\Lambda}\right)\right)}{1 - \widetilde{\beta} + (1-\gamma\rho)\,\Omega + \gamma\rho\Omega^2\left(\frac{2}{\Lambda(1-\Lambda)} - 1\right)} \tag{E.23}$$

We need to show that $\delta(\Omega) < 1$, i.e.

$$\left(1 - \widetilde{\beta} + \Omega\right)\left(1 + \frac{\gamma\rho\Omega}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{1}{1+(\rho/y)}\left(\frac{1+\Lambda}{1-\Lambda}\right)\right) < 1 - \widetilde{\beta} + (1-\gamma\rho)\,\Omega + \gamma\rho\Omega^2\left(\frac{2}{\Lambda(1-\Lambda)} - 1\right)$$

This expression can be simplified to yield:

$$1 + \frac{\left(1-\widetilde{\beta}\right)}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{y}{\rho+y}\left(\frac{1+\Lambda}{1-\Lambda}\right) < \Omega\left[\left(\frac{2}{\Lambda(1-\Lambda)} - 1\right) - \frac{1}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{y}{\rho+y}\left(\frac{1+\Lambda}{1-\Lambda}\right)\right] \tag{E.24}$$

First, we show that the term in the square brackets on the RHS of (E.24) is positive, i.e.

$$2 > \Lambda\left[1 - \Lambda + \frac{1+\Lambda}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{y}{\rho+y}\right]$$

The worst case for this to be true is if y is very large and $\varrho_z = 1$. In that case, for the expression above to be true, it must be that:

$$\widetilde{\beta} < \frac{2}{2-(1-\Lambda)\Lambda}$$

which is true since $\widetilde{\beta} < 1$ and $\frac{2}{2-(1-\Lambda)\Lambda} > 1$ since we know that $0 < \Lambda < 1$ from Appendix C. Thus, the term in the square brackets on the RHS of (E.24) is positive. Next, to show that (E.24) holds with countercyclical risk, it suffices to show that it holds for the lowest Ω consistent with non-procyclical risk, i.e. $\Omega = \frac{\Lambda}{1-\Lambda}$. Plug in $\Omega = \frac{\Lambda}{1-\Lambda}$ into (E.24), i.e:

$$1 + \frac{\left(1-\widetilde{\beta}\right)(1+\Lambda)}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{y}{\rho+y} < \left[\Lambda\left(\frac{2}{\Lambda(1-\Lambda)} - 1\right) - \frac{1+\Lambda}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{y}{\rho+y}\left(\frac{\Lambda}{1-\Lambda}\right)\right]$$

Again the worst case for this condition to be satisfied is if $\varrho_z = 1$. Suppose that is the case. Then, the expression can be further simplified to:

$$\frac{y}{\rho+y} < 1$$

which is true since steady state output is positive. □

Claim 3. $\Upsilon(0) = \delta(0) = 1$ *when* $\alpha = 0$.

Proof. True by inspection of equations (E.13), (E.14). □

Claim 4. $\Upsilon_0(\Omega) > \Upsilon(\Omega)$ *when* $\alpha \neq 0$

Proof. The claim that $\Upsilon_0 > \Upsilon$ is true by inspection of equations (E.16) since $\mathbb{G} > 0$ for $\Omega \geq \Omega^c > 0$. □

Claim 5. *$\Upsilon_0(\Omega)$ is increasing in α for $\alpha > 0$.*

Proof. Substituting the definition of $\mathbb{G}$ into (E.16):

$$\Upsilon_0(\Omega) = \Upsilon(\Omega) + \gamma(1+\Omega)^2 \frac{\gamma\rho}{m_4}\Lambda\left(\frac{\alpha}{\mu}\right)^2\left(\frac{\vartheta}{1-\vartheta}\right)\frac{m_3}{\mu}$$

This is clearly increasing in α for $\alpha > 0$. □

Claim 6. *$\varpi(\Omega) \in (0,1)$ for $\Omega \geq \Omega^c$*

Proof. Define:

$$\varpi(\Omega) = \frac{\delta(\Omega) - \varkappa(\Omega)}{1 - \varkappa(\Omega)}$$

where $\varkappa(\Omega) = \left(\frac{1-\widetilde{\beta}+\Omega}{\Omega}\right)\frac{\Lambda}{1-\widetilde{\beta}(1-\Lambda)\varrho_z}\frac{1}{1+\rho/y} \in (0,1)$ for $\Omega \geq \Omega^c$. Recall that Claim 2 above showed that $\delta(\Omega) < 1$ for $\Omega \geq \Omega^c$. Thus, we have $\varpi(\Omega) < 1$ for $\Omega \geq \Omega^c$. It remains to show that $\varpi(\Omega) > 0$, i.e., $\delta(\Omega) > \varkappa(\Omega)$. Recall from (E.14) that $\delta(\Omega)$ is given by:

$$\begin{aligned}\delta(\Omega) &= \frac{1}{\Upsilon(\Omega)}\left[1 + \left(\frac{1+\Lambda}{1-\Lambda}\right)\frac{\gamma\rho\Omega}{1-\widetilde{\beta}\varrho_z(1-\Lambda)}\frac{1}{1+\rho/y}\right] \\ &= \frac{1}{\Upsilon(\Omega)}\left[1 + \frac{\gamma\rho\Omega}{1-\widetilde{\beta}+\Omega}\left(\frac{1+\Lambda}{1-\Lambda}\right)\frac{\Omega}{\Lambda}\varkappa(\Omega)\right]\end{aligned}$$

So we have:

$$\begin{aligned}\delta(\Omega) - \varkappa(\Omega) &= \frac{1}{\Upsilon(\Omega)}\left[1 + \frac{\gamma\rho\Omega}{1-\widetilde{\beta}+\Omega}\left(\frac{1+\Lambda}{1-\Lambda}\right)\frac{\Omega}{\Lambda}\varkappa(\Omega) - \varkappa(\Omega)\Upsilon(\Omega)\right] \\ &= \frac{1}{\Upsilon(\Omega)}\left(1 - \frac{\gamma\rho(1-\Lambda)\Omega + \left(\frac{1-\widetilde{\beta}+\Omega}{\Omega}\right)\Lambda - \gamma\rho\Lambda}{1-\widetilde{\beta}(1-\Lambda)\varrho_z}\frac{1}{1+\rho/y}\right),\end{aligned}$$

where we have used the definitions of $\varkappa(\Omega)$ and $\Upsilon(\Omega)$. Since $\Upsilon(\Omega) > 0$, the expression above is positive if

$$1 \geq \frac{\gamma\rho(1-\Lambda)\Omega + \left(\frac{1-\widetilde{\beta}+\Omega}{\Omega}\right)\Lambda - \gamma\rho\Lambda}{1-\widetilde{\beta}(1-\Lambda)\varrho_z}\frac{1}{1+\rho/y}$$

or

$$(1+\rho/y)\left[1-\widetilde{\beta}\varrho_z(1-\Lambda)\right] > \Lambda\left(\frac{1-\widetilde{\beta}+\Omega}{\Omega}\right) - \gamma\rho\Lambda + \gamma\rho(1-\Lambda)\Omega \equiv \sqsupset(\Omega)$$

Clearly, $\sqsupset(\Omega)$ is a convex function of Ω. Since $\Omega = (1-\widetilde{\beta})\frac{w-1}{1+\gamma\rho w} < \lim_{w\to\infty}(1-\widetilde{\beta})\frac{w-1}{1+\gamma\rho w} = \frac{1-\widetilde{\beta}}{\gamma\rho}$, Ω is contained on the interval $\left[\Omega^c, \frac{1-\widetilde{\beta}}{\gamma\rho}\right]$. Thus,

$$\sqsupset(\Omega) \leq \max\left\{\sqsupset(\Omega^c), \sqsupset\left(\frac{1-\widetilde{\beta}}{\gamma\rho}\right)\right\}$$

where $\beth(\Omega^c) = \beth\left(\frac{1-\widetilde{\beta}}{\gamma\rho}\right) = 1 - \widetilde{\beta}\,(1-\Lambda)$. Thus, $\beth(\Omega) \leq 1 - \widetilde{\beta}(1-\Lambda)$. Clearly, we have

$$(1+\rho/y)\left[1 - \widetilde{\beta}\varrho_z\,(1-\Lambda)\right] > 1 - \widetilde{\beta}\,(1-\Lambda) \geq \beth(\Omega),$$

since $\rho/y > 0$ and $\varrho_z \in [0,1)$. Thus, $\delta(\Omega) - \varkappa(\Omega) > 0$ and $\varpi(\Omega) > 0$ when $\Omega \geq \Omega^c$. □

E.4 Deriving the target-criterion allowing for demand shocks

To derive a more general target criterion which allows for demand shocks in addition to aggregate productivity and markup shocks, we proceed by linearizing the first-order conditions of the non-linear planner's problem rather than adopting an LQ approach. Linearizing the first-order conditions (D.25)-(D.30) and constraints (D.1)-(D.4) around the steady state described in Appendix D.2 yields the following

FOC wrt w:

$$-(\gamma y)\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\widehat{y}_t + \left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\widehat{\Sigma}_t - \left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}}\right)\widehat{m}_{1,t} - m_1\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}}\right)^2\frac{\gamma\rho}{1+\gamma\rho}\widehat{w}_t$$
$$-\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}^2}\right)m_1\widehat{\mu}_t + \left(\frac{1+\gamma\rho}{\gamma\rho}\right)\frac{\varepsilon}{\Psi}\widehat{m}_{2,t} - \frac{\widehat{m}_{4,t}}{\gamma} + \frac{m_4}{\gamma}\widehat{w}_t - \frac{m_4}{\gamma}\frac{1}{1+\gamma\rho}\widehat{z}_t = 0 \quad \text{(E.25)}$$

FOC wrt y:

$$-\frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\widehat{w}_t + (\gamma y)\left[1 + 2\frac{(1-\Theta)^2}{\Lambda}\left(m_3 - \frac{m_1}{\beta}\right)\right]\widehat{y}_t - \widehat{\Sigma}_t - \widehat{m}_{1,t} + \frac{\Theta}{\beta}\left(\widehat{m}_{1,t-1} - m_1\widehat{\zeta}_{t-1}\right)$$
$$+2(1-\Theta)\left(m_3 - \frac{m_1}{\beta}\right)\widehat{\mu}_t + (1-\Theta)\,\widehat{m}_{3,t} + \frac{\widehat{m}_{4,t}}{\gamma} + 2(1-\Theta)\left(m_3 - \frac{m_1}{\beta}\right)\widehat{\varsigma}_t = 0 \quad \text{(E.26)}$$

FOC wrt π:

$$\Delta\widehat{m}_{2,t} = \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}\,(1-\Lambda)}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\frac{(\gamma y)\,\Psi}{1+\gamma\rho}\pi_t \;\Rightarrow\; \widehat{m}_{2,t} = \frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}\,(1-\Lambda)}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\frac{(\gamma y)\,\Psi}{1+\gamma\rho}\widehat{p}_t \quad \text{(E.27)}$$

FOC wrt μ:

$$-\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}^2}\right)\frac{\gamma\rho}{1+\gamma\rho}m_1\widehat{w}_t + \left[2\Lambda\left(m_3 - \frac{m_1}{\beta}\right) - \frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}m_1\right]\widehat{\mu}_t + \Lambda\widehat{m}_{3,t}$$
$$+2(\gamma y)(1-\Theta)\left(m_3 - \frac{m_1}{\beta}\right)\widehat{y}_t - \frac{1}{\widetilde{\beta}}\left(\widehat{m}_{1,t} - \frac{\widetilde{\beta}}{\beta}(1-\Lambda)\left(\widehat{m}_{1,t-1} - m_1\widehat{\zeta}_{t-1}\right)\right) + 2\Lambda\left(m_3 - \frac{m_1}{\beta}\right)\widehat{\varsigma}_t$$
$$+\mathbb{I}(t=0)\Lambda\left(\frac{\alpha}{\mu}\right)^2 m_3\left(\frac{\vartheta}{1-\vartheta}\right)\widehat{\mu}_0 = 0 \quad \text{(E.28)}$$

FOC wrt Σ:

$$\frac{\gamma\rho}{1+\gamma\rho}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\widehat{w}_t - (\gamma y)\,\widehat{y}_t - \widehat{m}_{3,t} + \widetilde{\beta}\widehat{m}_{3,t+1} + \frac{1-\beta^{-1}\widetilde{\beta}^2}{1-\widetilde{\beta}}\widehat{\Sigma}_t + \widetilde{\beta}m_3\widehat{\zeta}_t = 0 \quad \text{(E.29)}$$

where $\widehat{m}_i = \frac{\widehat{M}_i}{\mathbb{U}}$ for $i \in \{1,2,3,4\}$.

E.4.1 Deriving the target criterion

Add the FOC wrt w (E.25) to the FOC wrt y (E.26) to obtain:

$$
\begin{aligned}
&-(\gamma y)\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\widehat{y}_t + \frac{\Omega}{1-\widetilde{\beta}}\widehat{\Sigma}_t - \frac{\gamma\rho}{1+\gamma\rho}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\left[m_1\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)^2\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)+1\right]\widehat{w}_t + \frac{m_4}{\gamma}\widehat{w}_t \\
&+\left(\frac{1+\gamma\rho}{\gamma\rho}\right)\frac{\varepsilon}{\Psi}\widehat{m}_{2,t} - \frac{m_4}{\gamma}\frac{1}{1+\gamma\rho}\widehat{z}_t + \frac{\Theta}{\beta}\left(\widehat{m}_{1,t-1} - m_1\widehat{\zeta}_{t-1}\right) - \left[\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}}\right)+1\right]\widehat{m}_{1,t} \\
&+(\gamma y)\left[1+2\frac{(1-\Theta)^2}{\Lambda}\left(m_3-\frac{m_1}{\beta}\right)\right]\widehat{y}_t - \left[\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}^2}\right)m_1 - 2(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)\right]\widehat{\mu}_t \\
&+2(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)\widehat{\varsigma}_t + (1-\Theta)\widehat{m}_{3,t} \qquad \text{(E.30)}
\end{aligned}
$$

Combine with (E.28):

$$
\begin{aligned}
&(\gamma y)\left\{-\frac{\Omega}{1-\widetilde{\beta}} + 2\frac{(1-\Theta)^2}{\Lambda}\left(m_3-\frac{m_1}{\beta}\right) - 2(1-\Theta)(1+\Omega)\left(m_3-\frac{m_1}{\beta}\right)\right\}\widehat{y}_t + \frac{\Omega}{1-\widetilde{\beta}}\widehat{\Sigma}_t \\
&+\left\{m_1\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}^2}\right)\frac{\gamma\rho}{1+\gamma\rho}(1+\Omega) - \frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\left[m_1\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)^2\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)+1\right] + \frac{m_4}{\gamma}\right\}\widehat{w}_t \\
&-\left[2\left(m_3-\frac{m_1}{\beta}\right)+\frac{1}{\widetilde{\beta}}m_1\right]\Omega\widehat{\mu}_t \\
&-\Omega\widehat{m}_{3,t} + \left(\frac{1+\gamma\rho}{\gamma\rho}\right)\frac{\varepsilon}{\Psi}\widehat{m}_{2,t} - \frac{m_4}{\gamma}\frac{1}{1+\gamma\rho}\widehat{z}_t - 2\Omega\left(m_3-\frac{m_1}{\beta}\right)\widehat{\varsigma}_t \\
&-\mathbb{I}(t=0)\Lambda\left(\frac{\alpha}{\mu}\right)^2 m_3(1+\Omega)\left(\frac{\vartheta}{1-\vartheta}\right)\widehat{\mu}_0 = 0
\end{aligned}
$$

Next, use the GDP definition (E.3) to substitute out for $\widehat{w}_t$:

$$
\begin{aligned}
&(\gamma y)\left\{-\frac{\Omega}{1-\widetilde{\beta}} + 2\frac{(1-\Theta)^2}{\Lambda}\left(m_3-\frac{m_1}{\beta}\right) - 2(1-\Theta)(1+\Omega)\left(m_3-\frac{m_1}{\beta}\right)\right\}\widehat{y}_t + \frac{\Omega}{1-\widetilde{\beta}}\widehat{\Sigma}_t \\
&+\left\{m_1\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}^2}\right)\frac{\gamma\rho}{1+\gamma\rho}(1+\Omega) - \frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\left[m_1\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)^2\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)+1\right] + \frac{m_4}{\gamma}\right\}\frac{1+\gamma\rho}{\rho/y}\widehat{y}_t \\
&-\left\{m_1\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}^2}\right)\frac{\gamma\rho}{1+\gamma\rho}(1+\Omega) - \frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\left[m_1\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)^2\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)+1\right] + \frac{m_4}{\gamma}\right\}\frac{1}{\rho/y}\widehat{z}_t \\
&-\left[2\left(m_3-\frac{m_1}{\beta}\right)+\frac{1}{\widetilde{\beta}}m_1\right]\Omega\widehat{\mu}_t - \Omega\widehat{m}_{3,t} + \left(\frac{1+\gamma\rho}{\gamma\rho}\right)\frac{\varepsilon}{\Psi}\widehat{m}_{2,t} - \frac{m_4}{\gamma}\frac{1}{1+\gamma\rho}\widehat{z}_t \\
&-2\Omega\left(m_3-\frac{m_1}{\beta}\right)\widehat{\varsigma}_t - \mathbb{I}(t=0)\Lambda\left(\frac{\alpha}{\mu}\right)^2(1+\Omega)m_3\left(\frac{\vartheta}{1-\vartheta}\right)\widehat{\mu}_0 = 0 \qquad \text{(E.31)}
\end{aligned}
$$

Next, using (E.5) to substitute out for $\widehat{\mu}_t$, using (E.27) to eliminate $\widehat{m}_{2,t}$ and using the definitions of m_1, m_3 and m_4, (E.31) becomes

$$\begin{aligned}(\gamma y)\frac{\Omega}{1-\widetilde{\beta}}\left\{-1-2\left(\frac{1-\Theta}{\Lambda}\right)\left[\frac{1-\beta^{-1}\widetilde{\beta}}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}\right]+\left[\frac{2\left(1-\beta^{-1}\widetilde{\beta}\right)+\Lambda}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}\right]\frac{\Theta}{1-\Lambda}\right\}\widehat{y}_t+\frac{\Omega}{1-\widetilde{\beta}}\widehat{\Sigma}_t \\ +\frac{1}{\rho/y}\frac{\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}\left(\widehat{y}_t-\frac{1+\rho/y}{1+\gamma\rho}\widehat{z}_t\right) \\ -\frac{\Omega}{1-\widetilde{\beta}}\left[\frac{2\left(1-\beta^{-1}\widetilde{\beta}\right)+\Lambda}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}\right]\Gamma_t-\Omega\widehat{m}_{3,t}-2\frac{\Omega}{1-\widetilde{\beta}}\left[\frac{1-\beta^{-1}\widetilde{\beta}}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}\right]\widehat{\varsigma}_t \\ +\mathbb{I}(t=0)(\gamma y)\Lambda\left(\frac{\alpha}{\mu}\right)^2\frac{(1+\Omega)^2}{1-\widetilde{\beta}}\left(\frac{\vartheta}{1-\vartheta}\right)\widehat{y}_0 \\ -\mathbb{I}(t=0)\Lambda\left(\frac{\alpha}{\mu}\right)^2\frac{(1+\Omega)}{1-\widetilde{\beta}}\left(\frac{\vartheta}{1-\vartheta}\right)\Gamma_0+\left(\frac{\varepsilon y}{\rho}\right)\frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\widehat{p}_t &= 0\end{aligned} \tag{E.32}$$

Guess that

$$\widehat{m}_{3,t}=\frac{1}{1-\widetilde{\beta}}\widehat{\Sigma}_t+\gamma y\frac{\Omega}{1-\widetilde{\beta}}\widehat{y}_t+a_z\widehat{z}_t+a_\zeta\widehat{\zeta}_t+a_\varsigma\widehat{\varsigma}_t \tag{E.33}$$

and use this in (E.29) with $\widehat{w}_t$ substituted out using the definition of GDP:

$$\begin{aligned}\gamma y\Omega\widehat{y}_{t+1}-\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\left[\frac{\gamma y}{1+\gamma\rho}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)+a_z\left(1-\widetilde{\beta}\varrho_z\right)\right]\widehat{z}_t-\beta^{-1}\widetilde{\beta}\widehat{\Sigma}_t+\frac{\left(1-\widetilde{\beta}\right)}{\widetilde{\beta}}\left(\widetilde{\beta}\varrho_\zeta-1\right)a_\zeta\widehat{\zeta}_t \\ +\frac{\left(1-\widetilde{\beta}\right)}{\widetilde{\beta}}\left(\widetilde{\beta}\varrho_\varsigma-1\right)a_\varsigma\widehat{\varsigma}_t+\widehat{\Sigma}_{t+1}+\widehat{\zeta}_t &= 0\end{aligned}$$

using the fact that $\widehat{z}_{t+1}=\varrho_z\widehat{z}_t, \widehat{\zeta}_{t+1}=\varrho_\zeta\widehat{\zeta}_t$ and $\widehat{\varsigma}_{t+1}=\varrho_\varsigma\widehat{\varsigma}_t$. Using the expression for $\widehat{\Sigma}_{t+1}$ in (32) and the definition for $\widehat{y}^e_t$ in the equation above, we have

$$\begin{aligned}\left[\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)\left(\widetilde{\beta}\varrho_z-1\right)a_z-\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}}\right)\frac{\gamma y}{1+\gamma\rho}+\Lambda\frac{\gamma y}{1+\gamma\rho}\frac{1-\widetilde{\beta}+\Omega}{1-\widetilde{\beta}\varrho_z(1-\Lambda)}\varrho_z\right]\widehat{z}_t \\ +\left[\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)\left(\widetilde{\beta}\varrho_\zeta-1\right)a_\zeta+1-\frac{\widetilde{\beta}\varrho_\zeta\Lambda}{1-\widetilde{\beta}\varrho_\zeta(1-\Lambda)}\right]\widehat{\zeta}_t \\ +\left[\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)\left(\widetilde{\beta}\varrho_\varsigma-1\right)a_\varsigma-\widetilde{\beta}\frac{\varrho_\varsigma^2\Lambda^2}{1-\widetilde{\beta}\varrho_z(1-\Lambda)}+\varrho_\varsigma\Lambda\right]\widehat{\varsigma}_t &= 0\end{aligned}$$

which implies that a_z, a_ζ and a_ς must satisfy:

$$a_z = -\frac{\gamma y}{1+\gamma\rho}\left[\frac{1+\frac{\Omega}{1-\widetilde{\beta}}}{1-\widetilde{\beta}\varrho_z(1-\Lambda)}\right] \tag{E.34}$$

$$a_\zeta = \frac{\widetilde{\beta}}{1-\widetilde{\beta}}\left[\frac{1}{1-\widetilde{\beta}(1-\Lambda)\varrho_\zeta}\right] \tag{E.35}$$

$$a_\varsigma = \frac{1}{1-\widetilde{\beta}}\left[\frac{\widetilde{\beta}\varrho_\varsigma\Lambda}{1-\widetilde{\beta}(1-\Lambda)\varrho_\varsigma}\right] \tag{E.36}$$

Using the expression (E.33) for $\widehat{m}_{3,t}$ in (E.32):

$$\widehat{y}_0 - \delta_0(\Omega)\left(\frac{1+\rho/y}{1+\gamma\rho}\right)\widehat{z}_t + \chi_0(\Omega)\widehat{\zeta}_t - \Xi_0(\Omega)\widehat{\varsigma}_t + \frac{\varepsilon}{\Upsilon_0(\Omega)}\widehat{p}_t = 0 \tag{E.37}$$

$$\widehat{y}_t - \delta(\Omega)\left(\frac{1+\rho/y}{1+\gamma\rho}\right)\widehat{z}_t + \chi(\Omega)\widehat{\zeta}_t - \Xi(\Omega)\widehat{\varsigma}_t + \frac{\varepsilon}{\Upsilon(\Omega)}\widehat{p}_t = 0 \qquad \text{for } t>0 \tag{E.38}$$

where $\Upsilon(\Omega), \delta(\Omega), \Upsilon_0(\Omega)$ and $\delta_0(\Omega)$ are the same as in (E.13), (E.14), (E.16) and (E.17) and

$$\chi(\Omega) = \frac{1}{\Upsilon(\Omega)}\frac{\Omega}{1-\widetilde{\beta}+\Omega}\left(\frac{1+\Lambda}{1-\Lambda}\right)\left[\frac{\widetilde{\beta}(\rho/y)}{1-\widetilde{\beta}\varrho_\zeta(1-\Lambda)}\right] \tag{E.39}$$

$$\Xi(\Omega) = \frac{1}{\Upsilon(\Omega)}(\rho/y)\frac{\Omega}{1-\widetilde{\beta}+\Omega}\frac{2\left(1-\widetilde{\beta}\varrho_\varsigma\right)+\widetilde{\beta}\varrho_\varsigma\Lambda(1-\Lambda)}{(1-\Lambda)\left[1-\widetilde{\beta}\varrho_\varsigma(1-\Lambda)\right]} \tag{E.40}$$

$$\chi_0(\Omega) = \frac{\Upsilon(\Omega)}{\Upsilon_0(\Omega)}\chi(\Omega) + \frac{1}{\Upsilon_0(\Omega)}\frac{1}{\gamma y}\frac{\widetilde{\beta}}{1-\widetilde{\beta}\varrho_\zeta(1-\Lambda)}\mathbb{G} \tag{E.41}$$

$$\Xi_0(\Omega) = \frac{\Upsilon(\Omega)}{\Upsilon_0(\Omega)}\Xi(\Omega) - \frac{1}{\Upsilon_0(\Omega)}\frac{1}{\gamma y}\frac{\widetilde{\beta}\varrho_\varsigma\Lambda}{1-\widetilde{\beta}\varrho_\varsigma(1-\Lambda)}\mathbb{G} \tag{E.42}$$

where

$$\mathbb{G} = \gamma\rho\left[\frac{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}\right]\left[\frac{1+\Omega}{\left(1-\widetilde{\beta}+\Omega\right)}\right]\left(\frac{\alpha}{\mu}\right)^2\left(\frac{\vartheta}{1-\vartheta}\right)\Lambda$$

Note that in the baseline with the utilitarian planner ($\alpha = 0$), we have $\mathbb{G} = 0$ and $\Xi_0(\Omega) = \Xi(\Omega)$ and $\chi(\Omega) = \chi_0(\Omega)$. This general target criterion can be specialized to yield the target criterion in Proposition 3 in the paper for the utilitarian planner (setting $\alpha = 0$ and $\widehat{\zeta}_t = \widehat{\varsigma}_t = 0$), Proposition 9 in the paper for the non-utilitarian planner (again, setting $\widehat{\zeta}_t = \widehat{\varsigma}_t = 0$), i.e., it yields the same target criterion as the LQ approach. It can also be specialized to yield Proposition J.1 for demand shocks (setting $\alpha = 0$ and $\widehat{z}_t = \widehat{\varepsilon}_t = 0$).

Claim 7. $\chi(\Omega) > 0$ *with countercyclical risk*

Proof. It is clear from the expression for $\chi(\Omega)$ that for countercyclical risk $\Omega \geq \Omega^c > 0$, $\chi(\Omega) > 0$. □

Claim 8. $\Xi(\Omega) > 0$ *with countercyclical risk.*

Proof. For $\Omega \geq \Omega^c > 0$, it is clear from the expression for $\Xi(\Omega)$ that $\Xi(\Omega) > 0$. □

F Optimal Dynamics

As shown in Appendix E.4.1, the dynamics of x_t and π_t are given by the target criterion (36):

$$x_t - x_{t-1} + \bar{\varepsilon}\pi_t \quad = \quad 0 \tag{F.1}$$

and the Phillips curve

$$\pi_t = \beta\pi_{t+1} + \kappa\left(x_t - [1-\delta(\Omega)]\frac{1+(\rho/y)}{1+\gamma\rho}\widehat{z}_t - \chi(\Omega)\widehat{\zeta}_t + \Xi(\Omega)\widehat{\varsigma}_t + \frac{\rho/y}{1+\gamma\rho}\widehat{\varepsilon}_t\right) \tag{F.2}$$

where

$$x_t = \widehat{y}_t - \delta\left(\Omega\right)\left(\frac{1+\rho/y}{1+\gamma\rho}\right)\widehat{z}_t + \chi\left(\Omega\right)\widehat{\zeta}_t - \Xi\left(\Omega\right)\widehat{\varsigma}_t,$$

$\bar{\varepsilon} = \varepsilon/\Upsilon(\Omega)$ and $\kappa = \frac{\varepsilon}{\Psi}\frac{1+\gamma\rho}{\rho/y}$. Substituting the target criterion into the Phillips curve, we get a second-order difference equation:

$$x_{t+1} - \left[1 + \frac{\kappa\varepsilon+1}{\beta}\right]x_t + \frac{1}{\beta}x_{t-1} = \frac{\bar{\varepsilon}\kappa}{\beta}\left[-[1-\delta(\Omega)]\frac{1+(\rho/y)}{1+\gamma\rho}\widehat{z}_t - \chi(\Omega)\widehat{\zeta}_t + \Xi(\Omega)\widehat{\varsigma}_t + \frac{\rho/y}{1+\gamma\rho}\widehat{\varepsilon}_t\right]$$

The solution to this system has the form:

$$x_t \quad = \quad \mathcal{A}_x x_{t-1} + \mathcal{A}_z\widehat{z}_t + \mathcal{A}_\zeta\widehat{\zeta}_t + \mathcal{A}_\varepsilon\widehat{\varepsilon}_t + \mathcal{A}_\varsigma\widehat{\varsigma}_t \tag{F.3}$$

$$\pi_t \quad = \quad \mathcal{B}_x x_{t-1} + \mathcal{B}_z\widehat{z}_t + \mathcal{B}_\zeta\widehat{\zeta}_t + \mathcal{B}_\varepsilon\widehat{\varepsilon}_t + \mathcal{B}_\varsigma\widehat{\varsigma}_t \tag{F.4}$$

Using the method of undetermined coefficients, it is straightforward to see that $\mathcal{A}_x$ satisfies the characteristic polynomial:

$$\mathcal{P}(\mathcal{A}_x) = \mathcal{A}_x^2 - \left[1 + \frac{\kappa\bar{\varepsilon}+1}{\beta}\right]\mathcal{A}_x + \frac{1}{\beta} = 0 \tag{F.5}$$

We know that $\mathcal{P}(0) = \beta^{-1} > 0$ and $\mathcal{P}(1) = -\beta^{-1}\kappa\bar{\varepsilon} < 0$. Thus, we have $\mathcal{A}_x \in (0,1)$ and the coefficients can be written as:

$$\mathcal{A}_x \quad = \quad \frac{1}{2}\left(1 + \frac{\kappa\bar{\varepsilon}+1}{\beta} - \sqrt{\left[1 + \frac{\kappa\bar{\varepsilon}+1}{\beta}\right]^2 - \frac{4}{\beta}}\right) \in (0,1) \tag{F.6}$$

$$\mathcal{A}_z \quad = \quad \frac{\kappa\beta^{-1}\bar{\varepsilon}}{\kappa\beta^{-1}\bar{\varepsilon} + (1-\mathcal{A}_x) + \left(\frac{1}{\beta} - \varrho_z\right)}[1-\delta(\Omega)]\frac{1+(\rho/y)}{1+\gamma\rho} \tag{F.7}$$

$$\mathcal{A}_\zeta = \frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon} + (1-\mathcal{A}_x) + \left(\frac{1}{\beta} - \varrho_\zeta\right)}\chi(\Omega) \tag{F.8}$$

$$\mathcal{A}_\varepsilon = -\frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon} + (1-\mathcal{A}_x) + \left(\frac{1}{\beta} - \varrho_\varepsilon\right)}\frac{\rho/y}{1+\gamma\rho} \tag{F.9}$$

$$\mathcal{A}_\varsigma = -\frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon} + (1-\mathcal{A}_x) + \left(\frac{1}{\beta} - \varrho_\varsigma\right)}\Xi(\Omega) \tag{F.10}$$

$$\mathcal{B}_x = \frac{1-\mathcal{A}_x}{\overline{\varepsilon}} \tag{F.11}$$

$$\mathcal{B}_i = -\frac{1}{\overline{\varepsilon}}\mathcal{A}_i \qquad \text{for } i \in \{z, \zeta, \varepsilon, \varsigma\} \tag{F.12}$$

Claim 9. *The following statements are true:*

1. $\frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon}+(1-\mathcal{A}_x)+\left(\frac{1}{\beta}-\varrho_i\right)} \in (0,1)$ *for* $i \in \{z, \zeta, \varepsilon, \varsigma\}$
2. $\mathcal{B}_x > 0$

Proof. To see that $\frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon}+(1-\mathcal{A}_x)+\left(\frac{1}{\beta}-\varrho_i\right)} \in (0,1)$ for $i \in \{z, \zeta, \varepsilon, \varsigma\}$, notice that since $\mathcal{A}_x \in (0,1)$, we have $1-\mathcal{A}_x \geq 0$. Furthermore, since $\beta < 1$, so $\beta^{-1} - \varrho_i > 0$ for $i \in \{z, \zeta, \varepsilon, \varsigma\}$ as long as $\varrho_i \in (0,1)$, which is a maintained assumption. Again, since $\mathcal{A}_x \in (0,1)$, it is immediate that $\mathcal{B}_x > 0$. It follows that $\mathcal{A}_z > 0$, $\mathcal{A}_\zeta > 0$, and $\mathcal{A}_\varepsilon < 0$.

□

F.1 Proof of Propositions 5, 6, J.2 and J.3

F.1.1 Impact effects following a productivity shock

Since $x_{-1} = 0$, it follows from equations (F.3) and (F.4) that the impact effect of a productivity shock is:

$$\frac{\partial x_0}{\partial \widehat{z}_0} = \mathcal{A}_z > 0 \qquad \text{and} \qquad \frac{\partial \pi_0}{\partial \widehat{z}_0} = \mathcal{B}_z < 0$$

Using $\widehat{y}_t = x_t + \delta(\Omega)\frac{1+(\rho/y)}{1+\gamma\rho}\widehat{z}_t$, we have:

$$\begin{aligned}
\frac{\partial \widehat{y}_0}{\partial \widehat{z}_0} &= \mathcal{A}_z + \delta(\Omega)\frac{1+(\rho/y)}{1+\gamma\rho} \\
&= \frac{1+(\rho/y)}{1+\gamma\rho}\left(\delta(\Omega) + [1-\delta(\Omega)]\frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon} + (1-\mathcal{A}_x) + \left(\frac{1}{\beta} - \varrho_z\right)}\right) \in \left(0, \frac{1+(\rho/y)}{1+\gamma\rho}\right)
\end{aligned}$$

where we have used the fact that $\delta(\Omega) \in (0,1)$ for $\Omega \geq \Omega^c$. In other words, $\widehat{y}_0$ falls less than $\widehat{y}_0^n = \frac{1+(\rho/y)}{1+\gamma\rho}\widehat{z}_0$ for $\widehat{z}_0 < 0$.

F.1.2 Impact effects following a markup shock

Since $x_{-1} = 0$ and $\widehat{y}_t = x_t$ (since all shocks other than markup shocks are 0 in this case), it follows immediately from equations (F.3) and (F.4) that:

$$\begin{aligned}\frac{\partial \widehat{y}_0}{\partial \widehat{\varepsilon}_0} &= \mathcal{A}_\varepsilon < 0 \\ \frac{\partial \pi_0}{\partial \widehat{\varepsilon}_0} &= \mathcal{B}_\varepsilon > 0\end{aligned}$$

Following a markup shock, the dynamics of π_t and x_t are described by the same equations (F.1) and (F.2) except that $\overline{\varepsilon}(\Omega) = \frac{\varepsilon}{\Upsilon(\Omega)}$ is smaller in HANK since $\Upsilon(\Omega) > 1$ while $\Upsilon = 1$ in RANK. Thus, to show that in HANK, output decreases less and inflation increases more following a positive markup shock, it suffices to show that $\frac{\partial \mathcal{A}_\varepsilon}{\partial \overline{\varepsilon}} < 0$ (output falls less on impact when $\overline{\varepsilon}$ is lower) and $\frac{\partial \mathcal{B}_\varepsilon}{\partial \overline{\varepsilon}} < 0$ (inflation increases more on impact when $\overline{\varepsilon}$ is lower). We have:

$$\begin{aligned}\mathcal{A}_\varepsilon &= -\frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon} + (1-\mathcal{A}_x) + \left(\frac{1}{\beta} - \varrho_\varepsilon\right)}\frac{\rho/y}{1+\gamma\rho} \\ &= -\frac{2\beta^{-1}\kappa}{\sqrt{\frac{[1+\beta^{-1}(\kappa\overline{\varepsilon}+1)]^2-4\beta^{-1}}{\overline{\varepsilon}^2}} + \left(\frac{1-\varrho_\varepsilon}{\overline{\varepsilon}}\right) + \frac{\beta^{-1}-\varrho_\varepsilon}{\overline{\varepsilon}} + \beta^{-1}\kappa}\frac{\rho/y}{1+\gamma\rho}\end{aligned}$$

where we have plugged in the expression for $\mathcal{A}_x$ from (F.6) in the second line. Since

$$\frac{\partial}{\partial \overline{\varepsilon}}\left[\frac{\left[1+\beta^{-1}\left(\kappa\overline{\varepsilon}+1\right)\right]^2 - 4\beta^{-1}}{\overline{\varepsilon}^2}\right] = -2\left[\frac{\left[1+\beta^{-1}\left(\kappa\overline{\varepsilon}+1\right)\right]\left(1+\beta^{-1}\right)+4\beta^{-1}}{\overline{\varepsilon}^3}\right] < 0$$

it is clear that the denominator of $\mathcal{A}_\varepsilon$ is decreasing in $\overline{\varepsilon}$. Since the numerator is negative, it follows that $\mathcal{A}_\varepsilon$ is decreasing in $\overline{\varepsilon}$. We also know that $\mathcal{B}_\varepsilon = -\frac{1}{\overline{\varepsilon}}\mathcal{A}_\varepsilon$ which implies:

$$\mathcal{B}_\varepsilon = \frac{2\beta^{-1}\kappa}{\sqrt{\left[1+\beta^{-1}\left(\kappa\overline{\varepsilon}+1\right)\right]^2 - 4\beta^{-1}} + (1-\varrho_\varepsilon) + (\beta^{-1}-\varrho_\varepsilon) + \beta^{-1}\kappa\overline{\varepsilon}}\frac{\rho/y}{1+\gamma\rho}$$

Clearly, the denominator is increasing in ε so $\mathcal{B}_\varepsilon$ is decreasing in $\overline{\varepsilon}$.

F.1.3 Impact effects following a discount factor shock

Since $x_{-1} = 0$ and $y_t = x_t - \chi(\Omega)\widehat{\zeta}_t$, the response of $\widehat{y}_0$ to $\widehat{\zeta}_0$ is:

$$\frac{d\widehat{y}_0}{d\widehat{\zeta}_0} = \mathcal{A}_\zeta - \chi\left(\Omega\right) = -\left[1 - \frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon} + (1-\mathcal{A}_x) + \left(\frac{1}{\beta} - \varrho_\zeta\right)}\right]\chi\left(\Omega\right) < 0$$

while the impact response of π_0 is given by $\mathcal{B}_\zeta = -\frac{1}{\overline{\varepsilon}}\mathcal{A}_\zeta < 0$.

F.2 Impact effects following a risk shock

Since $x_{-1} = 0$ and $\widehat{y}_t = x_t + \Xi(\Omega)\widehat{\varsigma}_t$, the response of $\widehat{y}_0$ to $\widehat{\varsigma}_0$ is:

$$\frac{d\widehat{y}_0}{d\widehat{\varsigma}_0} = \mathcal{A}_\varsigma + \Xi(\Omega) = \left[1 - \frac{\kappa\beta^{-1}\overline{\varepsilon}}{\kappa\beta^{-1}\overline{\varepsilon} + (1 - \mathcal{A}_x) + \left(\frac{1}{\beta} - \varrho_\varsigma\right)}\right] \Xi(\Omega) > 0$$

for $\Omega \geq \Omega^c$. Similarly, the impact response of π_0 is given by $\mathcal{B}_\varsigma = -\frac{1}{\overline{\varepsilon}}\mathcal{A}_\varsigma > 0$.

F.3 Response of $\widehat{y}_t - \widehat{y}_t^n$ and π_t for large t

The following Lemma characterizes the behavior of $\widehat{y}_t - \widehat{y}_t^n$ and π_t following a generic shock $\mathcal{S}_0$ where $\mathcal{S}_0 \in \{\widehat{z}_0, \widehat{\varepsilon}_0, \widehat{\zeta}_0, \widehat{\varsigma}_0\}$ for large t. In doing so, the Lemma provides a proof of the claims made in Propositions 5, 6, J.2 and J.3 about long-run behavior of $\widehat{y}_t - \widehat{y}_t^n$ and π_t.

Lemma 3. *After any date 0 shock $\mathcal{S}_0$ where $\mathcal{S}_0 \in \{\widehat{z}_0, \widehat{\varepsilon}_0, \widehat{\zeta}_0, \widehat{\varsigma}_0\}$, for large enough t,*

$$sign\left(\frac{\partial \pi_t}{\partial \mathcal{S}_0}\right) = sign\left(\frac{\partial(\widehat{y}_t - \widehat{y}_t^n)}{\partial \mathcal{S}_0}\right) = -1 \times sign\left(\frac{\partial \pi_0}{\partial \mathcal{S}_0}\right)$$

Proof. We know that $\frac{\partial \pi_0}{\partial \mathcal{S}_0} = \mathcal{B}_\mathcal{S} = -\frac{1}{\overline{\varepsilon}}\mathcal{A}_\mathcal{S}$. Thus, we need to show that for large enough t, π_t and $\widehat{y}_t - \widehat{y}_t^n$ have the same sign as $\mathcal{A}_\mathcal{S} \times \mathcal{S}_0$. The dynamics of x_t and π_t in response to a shock $\mathcal{S}_0$ are given by the system of two equations:

$$x_t = \mathcal{A}_x x_{t-1} + \mathcal{A}_\mathcal{S}\mathcal{S}_t \qquad \text{and} \qquad \mathcal{S}_t = \varrho_\mathcal{S}\mathcal{S}_{t-1}$$

with $\mathcal{S}_0$ given. The solution of this system is given by:

$$x_t = \mathcal{A}_\mathcal{S}\frac{\varrho_\mathcal{S}^{t+1} - \mathcal{A}_x^{t+1}}{\varrho_\mathcal{S} - \mathcal{A}_x}\mathcal{S}_0$$

as long as $\varrho_\mathcal{S} \neq \mathcal{A}_x$. Using this in (36) in the paper, the dynamics of inflation can then be written as:

$$\pi_t = -\frac{\mathcal{A}_\mathcal{S}}{\overline{\varepsilon}}\left(\frac{\varrho_\mathcal{S}^{t+1} - \mathcal{A}_x^{t+1}}{\varrho_\mathcal{S} - \mathcal{A}_x} - \frac{\varrho_\mathcal{S}^{t} - \mathcal{A}_x^{t}}{\varrho_\mathcal{S} - \mathcal{A}_x}\right)\mathcal{S}_0$$

where $\overline{\varepsilon} > 0$, $\mathcal{A}_\mathcal{S} > 0$ and $0 < \mathcal{A}_x < 1$ are defined in Appendix F. For large enough $t > 0$, the dynamics of x_t and π_t are governed by the dominant eigenvalue $\max\{\mathcal{A}_x, \varrho_\mathcal{S}\}$. If $\varrho_\mathcal{S} < \mathcal{A}_\mathcal{S}$, dividing expression for π_t above by $\mathcal{A}_x^t$ and taking the limit $t \to \infty$, we have:

$$\lim_{t\to\infty} \mathcal{A}_x^{-t}\pi_t = \frac{1}{\overline{\varepsilon}}\left(\frac{\mathcal{A}_\mathcal{S}}{\mathcal{A}_x - \varrho_\mathcal{S}}\right)\underbrace{(1 - \mathcal{A}_x)}_{>0}\mathcal{S}_0$$

which has the same sign as $\mathcal{A}_\mathcal{S} \times \mathcal{S}_0$. Similarly, dividing the Phillips curve by $\mathcal{A}_x^t$ and taking limits as $t \to \infty$ yields:

$$\lim_{t\to\infty}(1 - \beta\mathcal{A}_x)\frac{\pi_t}{\mathcal{A}_x^t} = \kappa \lim_{t\to\infty}\left(\frac{\widehat{y}_t - \widehat{y}_t^n}{\mathcal{A}_x^t}\right)$$

This implies that $\widehat{y}_t - \widehat{y}_t^n$ has the same sign as $\mathcal{A}_\mathcal{S} \times \mathcal{S}_0$ for large t. Instead if $\varrho_\mathcal{S} > \mathcal{A}_x$, dividing by $\varrho_\mathcal{S}^t$ and taking limits, we have

$$\lim_{t\to\infty}\varrho_{\mathcal{S}}^{-t}\pi_t = \frac{1}{\bar{\varepsilon}}\underbrace{(1-\varrho_{\mathcal{S}})}_{>0}\frac{\mathcal{A}_{\mathcal{S}}}{\varrho_{\mathcal{S}}-\mathcal{A}_x}\mathcal{S}_0 \qquad \text{and} \qquad \lim_{t\to\infty}(1-\beta\varrho_{\mathcal{S}})\frac{\pi_t}{\varrho_{\mathcal{S}}^t} = \kappa\lim_{t\to\infty}\left(\frac{\widehat{y}_t-\widehat{y}_t^n}{\varrho_{\mathcal{S}}^t}\right)$$

This implies that again, π_t and $\widehat{y}_t-\widehat{y}_t^n$ have the same sign as $\mathcal{A}_{\mathcal{S}}\times\mathcal{S}_0$ for large t. Finally, in the special case where both eigenvalues are identical $\mathcal{A}_x=\varrho_{\mathcal{S}}$, the solution for x_t is instead given by:

$$x_t = (t+1)\,\mathcal{A}_{\mathcal{S}}\varrho_{\mathcal{S}}^t\mathcal{S}_0$$

and so the target criterion implies that the path of inflation can be written as:

$$\pi_t = -\frac{\mathcal{A}_{\mathcal{S}}}{\bar{\varepsilon}}\left((t+1)\,\varrho_{\mathcal{S}}^t - t\varrho_{\mathcal{S}}^{t-1}\right)\mathcal{S}_0$$

Divide this by $(t+1)\,\varrho_{\mathcal{S}}^t$ and take limits:

$$\lim_{t\to\infty}\frac{\pi_t}{(t+1)\,\varrho_{\mathcal{S}}^t} = \frac{\mathcal{A}_{\mathcal{S}}}{\bar{\varepsilon}}\left(\frac{1-\varrho_{\mathcal{S}}}{\varrho_{\mathcal{S}}}\right)\mathcal{S}_0$$

Following the same steps as above with the Phillips curve and taking limits yields:

$$\left(1-\frac{\rho_z}{R}\right)\lim_{t\to\infty}\frac{\pi_t}{(t+1)\,\varrho_{\mathcal{S}}^t} = \kappa\lim_{t\to\infty}\left(\frac{\widehat{y}_t-\widehat{y}_t^n}{(t+1)\,\varrho_{\mathcal{S}}^t}\right)$$

Thus, even in this case, the sign of $\widehat{y}_t-\widehat{y}_t^n$ and π_t is the same as that of $\mathcal{A}_{\mathcal{S}}\times\mathcal{S}_0$ for large t. □

F.4 Interest rate rules

We have already seen that under optimal policy, the dynamics of x_t and π_t can be written as functions of x_{t-1} and shocks – equations (F.3) and (F.4). Substituting (E.5) into the linearized IS equation (23) in the paper:

$$\widehat{y}_t = (1+\Omega)\,\widehat{y}_{t+1} - \frac{1}{\gamma y}(i_t-\pi_{t+1}) - \frac{\Lambda}{\gamma y}\Gamma_z\varrho_z\widehat{z}_t$$

We can use this equation along with equations (F.3) and (F.4) to express i_t in terms of x_{t-1} and the shocks:

$$\begin{aligned}
i_t &= \gamma y\,(1+\Omega)\,\widehat{y}_{t+1} - \gamma y\widehat{y}_t + \pi_{t+1} - \Lambda\Gamma_z\varrho_z\widehat{z}_t \\
&= \gamma y\,(1+\Omega)\,x_{t+1} - \gamma y x_t + \pi_{t+1} \\
&\quad + \left\{\gamma y\Omega\delta\,(\Omega)\,\frac{1+(\rho/y)}{1+\gamma\rho}\varrho_z - \Lambda\Gamma_z\varrho_z - \gamma y\delta\,(\Omega)\,\frac{1+(\rho/y)}{1+\gamma\rho}(1-\varrho_z)\right\}\widehat{z}_t \\
&= (\gamma y\,(1+\Omega)\,\mathcal{A}_x - \gamma y + \mathcal{B}_x)\,x_t \\
&\quad + \gamma y\left\{\Omega\delta\,(\Omega)\,\frac{1+(\rho/y)}{1+\gamma\rho}\varrho_z - \frac{\Lambda\Gamma_z\varrho_z}{\gamma y} - \delta\,(\Omega)\,\frac{1+(\rho/y)}{1+\gamma\rho}(1-\varrho_z) + (1+\Omega)\,\mathcal{A}_z\varrho_z + \frac{\mathcal{B}_z\varrho_z}{\gamma y}\right\}\widehat{z}_t \\
&\quad + (\gamma y\,(1+\Omega)\,\mathcal{A}_\varepsilon\varrho_\varepsilon\widehat{\varepsilon}_t + \mathcal{B}_\varepsilon\varrho_\varepsilon)\,\widehat{\varepsilon}_t \\
&= \Phi_x x_{t-1} + \Phi_z\widehat{z}_t + \Phi_\varepsilon\widehat{\varepsilon}_t
\end{aligned}$$

where

$$\begin{aligned}
\Phi_x &= \{\gamma y (1+\Omega)\mathcal{A}_x - \gamma y + \mathcal{B}_x\}\mathcal{A}_x \\
\Phi_z &= \{\gamma y (1+\Omega)\mathcal{A}_x - \gamma y + \mathcal{B}_x\}\mathcal{A}_z \\
&\quad +\gamma y\left\{\Omega\delta(\Omega)\frac{1+(\rho/y)}{1+\gamma\rho}\varrho_z - \frac{\Lambda\Gamma_z\varrho_z}{\gamma y} - \delta(\Omega)\frac{1+(\rho/y)}{1+\gamma\rho}(1-\varrho_z) + (1+\Omega)\mathcal{A}_z\varrho_z + \frac{\mathcal{B}_z\varrho_z}{\gamma y}\right\} \\
\Phi_\varepsilon &= \gamma y(1+\Omega)\mathcal{A}_\varepsilon\varrho_\varepsilon\widehat{\varepsilon}_t + \mathcal{B}_\varepsilon\varrho_\varepsilon + (\gamma y(1+\Omega)\mathcal{A}_x - \gamma y + \mathcal{B}_x)\mathcal{A}_\varepsilon
\end{aligned}$$

Next, we show that (38) in the paper implements the optimal allocations uniquely. First, note that first-differencing the target criterion (36) in the paper and multiplying by $\phi_x \equiv \phi\frac{\Upsilon(\Omega)}{\varepsilon}$ yields:

$$\phi\pi_t + \phi_x\Delta x_t = \phi\pi_t + \phi_{\text{gap}}\left(\Delta\widehat{y}_t - \Delta\widehat{y}_t^e\right) + \phi_y\Delta\widehat{y}_t = 0$$

where $x_t = \widehat{y}_t - \delta(\Omega)\frac{1+(\rho/y)}{1+\gamma\rho}\widehat{z}_t$, $\phi > 0$ is a constant, $\phi_{\text{gap}} = \phi\frac{\Upsilon(\Omega)}{\varepsilon}\delta(\Omega)$ is the weight on the change in output gap and $\phi_y = \phi\frac{\Upsilon(\Omega)}{\varepsilon}(1-\delta(\Omega))$. Here, instead of writing the rule in terms of $\widehat{y}_t$ and the output gap $\widehat{y}_t - \widehat{y}_t^e$, it is more convenient to write it in terms of a single variable x_t; the two formulations are equivalent. Since by definition, we have $i_t = i_t^\star$ under optimal policy, it follows that the rule (38) in the paper is satisfied at the optimal allocation. To see that this rule implements optimal allocations uniquely, it suffices to look at the determinacy properties of the system comprised by the IS curve, the Phillips curve and the interest rate rule absent shocks. This system can be written as:

$$\begin{aligned}
(\gamma y + \phi_x)x_t &= \gamma y(1+\Omega)x_{t+1} - \Phi_x x_{t-1} - \phi\pi_t + \phi_x x_{t-1} + \pi_{t+1} \\
\pi_t &= \beta\pi_{t+1} + \kappa x_t
\end{aligned}$$

In matrix-form, this can be written as:

$$\begin{bmatrix} x_{t+1} \\ \pi_{t+1} \\ Lx_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{\beta\gamma y+\beta\phi_x+\kappa}{\beta\gamma(1+\Omega)} & -\frac{1-\beta\phi}{\beta\gamma(1+\Omega)} & \frac{\Phi_x-\phi_x}{\gamma(1+\Omega)} \\ -\frac{\kappa}{\beta} & \frac{1}{\beta} & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \pi_t \\ Lx_t \end{bmatrix}$$

where $Lx_t \equiv x_{t-1}$. The characteristic polynomial of this system is given by

$$\begin{aligned}
\mathcal{P}(\aleph) &= -\left(\frac{1}{\beta} - \aleph\right)\left(\frac{\Phi_x - \phi_x}{\gamma y(1+\Omega)}\right) \\
&\quad -\aleph\left\{\left(\frac{\beta\gamma y + \beta\phi_x + \kappa}{\beta\gamma y(1+\Omega)} - \aleph\right)\left(\frac{1}{\beta} - \aleph\right) - \frac{1-\beta\phi}{\beta\gamma y(1+\Omega)}\frac{\kappa}{\beta}\right\}
\end{aligned}$$

Notice that

$$\mathcal{P}(0) = \frac{\phi\Upsilon(\Omega)/\varepsilon - \Phi_x}{\beta\gamma y(1+\Omega)} \qquad \mathcal{P}(1) = \frac{\frac{\kappa(1-\phi)}{\beta} - \left(\frac{1}{\beta} - 1\right)(\Phi_x - \gamma y\Omega)}{\gamma y(1+\Omega)}$$

Clearly, for large enough ϕ, we have $\mathcal{P}(0) > 0$ and $\mathcal{P}(1) < 0$, implying that there is at least one root inside the unit circle. Also, note that:

$$\frac{\partial \mathcal{P}(\aleph)}{\partial \phi} = \frac{1}{\beta\gamma y (1+\Omega)} \left[\frac{\Upsilon(\Omega)}{\varepsilon} (1-\beta\aleph)(1-\aleph) - \kappa\aleph \right]$$

which is positive for a finite $\overline{\aleph} > \beta^{-1} > 1$. It follows that for sufficiently large ϕ, $\mathcal{P}(\overline{\aleph}) > 0$. Finally,

$$\lim_{\aleph \to \infty} \mathcal{P}(\aleph) = -\infty$$

implying that for sufficiently large ϕ, there are two roots above 1. Thus, the system has one stable and two unstable eigenvalues as we have 2 jump variables (π_t and x_t) and one predetermined variable Lx_t.

G Unequal distribution of profits

The date s problem of an individual i who is a stockholder (d) or nonstockholder (nd) born at date s can be written as:

$$\max_{\{c_t^s(i), \ell_t^s(i), a_{t+1}^s(i)\}} -\mathbb{E}_s \sum_{t=s}^{\infty} (\beta\vartheta)^{t-s} \left(\prod_{k=s}^{t-1} \zeta_k \right) \left\{ \frac{1}{\gamma} e^{-\gamma c_t^s(i)} + \rho e^{\frac{1}{\rho}[\ell_t^s(i) - \xi_t^s(i)]} \right\}$$

s.t.

$$c_t^s(i) + q_t a_{t+1}^s(i) \quad = \quad w_t \ell_t^s(i) + (1-\tau_t^a) a_t^s(i) + \mathbb{T}_t(i) \tag{G.1}$$

where $a_s^s(i) = 0$ and $w_t = (1-\tau^w)\widetilde{w}_t$ and $\tau_t^a = 0$ for $t > 0$. For a stockholder i, $\mathbb{T}_t(i) = \frac{D_t}{\eta^d} - T_t - J$ where J is the lump sum tax on stockholders and D_t/η^d is the dividend received by each of the η^d stockholders. For a nonstockholder $\mathbb{T}_t(i) = -T_t + \frac{\eta^d}{1-\eta^d} J$. The individual decision problem then is the same as in Appendix A replacing $D_t - T_t$ with $\mathbb{T}_t(i)$. Thus, following the steps in Appendix A, it is easy to see that the consumption function for stockholders can be written as:

$$c_t^s(i; d) = \mathcal{C}_t^d + \mu_t x_t^s(i; d)$$

and for nonstockholders:

$$c_t^s(i; nd) = \mathcal{C}_t^{nd} + \mu_t x_t^s(i; nd)$$

where the definition of $x = a + w\left(\xi - \bar{\xi}\right)$ is the same as in the baseline model.

$$\mathcal{C}_t^d \quad = \quad -\frac{\vartheta\mu_t}{\mu_{t+1}R_t}\frac{1}{\gamma}\ln\beta R_t + \frac{\vartheta\mu_t}{\mu_{t+1}R_t}\mathcal{C}_{t+1}^d + \mu_t\left[w_t\left(\rho\ln w_t + \bar{\xi}\right) + \frac{D_t}{\eta^d} - T_t - J\right] - \frac{\vartheta}{R_t}\frac{\mu_t}{\mu_{t+1}}\frac{\gamma\mu_{t+1}^2 w_{t+1}^2 \sigma_{t+1}^2}{2} \tag{G.2}$$

$$\mathcal{C}_t^{nd} \quad = \quad -\frac{\vartheta\mu_t}{\mu_{t+1}R_t}\frac{1}{\gamma}\ln\beta R_t + \frac{\vartheta\mu_t}{\mu_{t+1}R_t}\mathcal{C}_{t+1}^{nd} + \mu_t\left[w_t\left(\rho\ln w_t + \bar{\xi}\right) - T_t + \frac{\eta^d}{1-\eta^d}J\right] - \frac{\vartheta}{R_t}\frac{\mu_t}{\mu_{t+1}}\frac{\gamma\mu_{t+1}^2 w_{t+1}^2 \sigma_{t+1}^2}{2} \tag{G.3}$$

$$\mu_t^{-1} \quad = \quad (1 + \rho\gamma w_t) + \frac{\vartheta}{R_t}\mu_{t+1}^{-1} \tag{G.4}$$

Since $x_t^s(i)$ has mean zero at any date and both types of households have the same μ_t, the goods market clearing condition can be written as:

$$\eta^d \mathcal{C}_t^d + (1-\eta^d)\mathcal{C}^{nd} = y_t$$

Multiplying (G.2) by η^d and (G.3) by $1-\eta^d$ and adding the two along with market clearing and rearranging yields the aggregate Euler equation which is the same as in the baseline model:

$$y_t = -\frac{1}{\gamma}\ln \beta R_t + y_{t+1} - \frac{\gamma}{2}\mu_{t+1}^2 w_{t+1}^2 \sigma_{t+1}^2 \tag{G.5}$$

Combining (G.2) and (G.5):

$$\left(\mathcal{C}_t^d - y_t\right) = \frac{\vartheta}{R_t}\frac{\mu_t}{\mu_{t+1}}\left(\mathcal{C}_{t+1}^d - y_{t+1}\right) + \mu_t\left(\frac{1-\eta^d}{\eta^d}d_t - J\right) \tag{G.6}$$

Iterating forwards:

$$\mathcal{V}_t \equiv \frac{\eta^d}{1-\eta^d}\left(\frac{\mathcal{C}_t^d - y_t}{\mu_t}\right) = \sum_{s=0}^{\infty}\frac{\vartheta^s}{\prod_{k=0}^{s-1} R_{t+s}}\left[D_{t+s} - \frac{\eta^d}{1-\eta^d}J\right]$$

In other words, we have $\mathcal{C}_t^d = y_t + \frac{1-\eta^d}{\eta^d}\mu_t\mathcal{V}_t$ as in the main text. Market clearing, then implies that $\mathcal{C}_t^{nd} = y_t - \mu_t\mathcal{V}_t$. As claimed in the main text, $J = \frac{1-\eta^d}{\eta^d}D$ implies that $\mathcal{V} = 0$ in steady state and average consumption of stockholders and nonstockholders is the same $\mathcal{C}^d = \mathcal{C}^{nd}$. Thus, as in the main text, we can rewrite the definition of $\mathcal{V}_t$ as:

$$\mathcal{V}_t = (D_t - D) + \frac{\vartheta}{R_t}\mathcal{V}_{t+1} \tag{G.7}$$

Since aggregate dividends $D_t = y_t - (1-\tau^\star)\widetilde{w}_t n_t$ can be written as:

$$D_t = y_t - \frac{(\varepsilon-1)w_t}{\varepsilon(1-\tau^w)}\frac{y_t}{z_t}\left[1 + \frac{\Psi}{2}(\Pi_t - 1)^2\right] = \left(1 - \frac{\varepsilon-1}{\varepsilon\,(1-\tau^w)}\frac{w_t}{z_t}\right)y_t - \frac{\varepsilon-1}{\varepsilon\,(1-\tau^w)}\frac{w_t}{z_t}\frac{\Psi}{2}\left(\Pi_t - 1\right)^2 y_t,$$

we can write the level-deviation $\widehat{D}_t$ as:

$$\frac{\widehat{D}_t}{y} = \underbrace{\left[\frac{1}{\varepsilon} - \left(\frac{\varepsilon-1}{\varepsilon}\right)\frac{1+\gamma\rho}{\rho/y}\right]}_{=\mathcal{D}_y}\widehat{y}_t + \underbrace{\left[\left(\frac{\varepsilon-1}{\varepsilon}\right)\frac{1+\rho/y}{\rho/y}\right]}_{=\mathcal{D}_z}\widehat{z}_t \tag{G.8}$$

Using this it is straightforward to derive $\widetilde{V}_t = \mathcal{D}_y\widehat{y}_t + \mathcal{D}_z\widehat{z}_t + \widetilde{\beta}\widetilde{V}_{t+1}$, where $\widetilde{V}_t = \widehat{\mathcal{V}}_t/y$ and $\widehat{\mathcal{V}}_t$ denotes the level deviation of $\mathcal{V}_t$ from its steady state value of 0.

G.1 Derivation of the Σ recursion

Even in this case, the objective function of the planner can be written as:

$$\mathbb{W}_0 = \sum_{t=0}^{\infty} \beta^t u\left(c_t, n_t; \overline{\xi}\right) \Sigma_t$$

where, as before, Σ_t is defined by:

$$\Sigma_t = (1-\vartheta) \sum_{s=-\infty}^{t} \int \vartheta^{t-s} e^{-\gamma\left(c_t^s(i) - c_t\right)} di$$

Since we have stockholders and nonstockholders, this can be further expanded:

$$\begin{aligned}
\Sigma_t &= (1-\vartheta)\left\{ \sum_{s=-\infty}^{t-1} \int \vartheta^{t-s} e^{-\gamma\left(c_t^s(i) - y_t\right)} di + \int e^{-\gamma\left(c_t^t(i) - y_t\right)} di \right\} \\
&= (1-\vartheta)\left\{ \sum_{s=-\infty}^{t-1} \int \vartheta^{t-s} e^{-\gamma\left(c_t^s(i) - y_t\right)} di + \eta^d \int e^{-\gamma\left(c_t^t(i;d) - y_t\right)} di + \left(1-\eta^d\right) \int e^{-\gamma\left(c_t^t(i;nd) - y_t\right)} di \right\}
\end{aligned}$$

Since $x_t^t(i) = w_t\left(\xi_t^t(i) - \overline{\xi}\right)$, we have:

$$\begin{aligned}
\Sigma_t &= (1-\vartheta) \sum_{s=-\infty}^{t-1} \int \vartheta^{t-s} e^{-\gamma\left(c_t^s(i) - y_t\right)} di \\
&\quad + (1-\vartheta)\left\{ \eta^d \int e^{-\gamma\left(\mathcal{C}_t^d - y_t + \mu_t w_t\left(\xi_t^t(i) - \overline{\xi}\right)\right)} di + \left(1-\eta^d\right) \int e^{-\gamma\left(\mathcal{C}_t^{nd} - y_t + \mu_t w_t\left(\xi_t^t(i) - \overline{\xi}\right)\right)} di \right\}
\end{aligned}$$

For dates $t > 0$, we can additionally write Σ_t as:

$$\begin{aligned}
\Sigma_t &= (1-\vartheta) \sum_{s=-\infty}^{t-1} \int \vartheta^{t-s} e^{-\gamma\left(c_{t-1}^s(i) - y_{t-1}\right)} e^{-\gamma\left(c_t^s(i) - c_{t-1}^s(i) - y_t + y_{t-1}\right)} di \\
&\quad + (1-\vartheta)\left\{ \eta^d \int e^{-\gamma\left(\mathcal{C}_t^d - y_t + \mu_t w_t\left(\xi_t^t(i) - \overline{\xi}\right)\right)} di + \left(1-\eta^d\right) \int e^{-\gamma\left(\mathcal{C}_t^{nd} - y_t + \mu_t w_t\left(\xi_t^t(i) - \overline{\xi}\right)\right)} di \right\} \\
&= (1-\vartheta) \sum_{s=-\infty}^{t-1} \int \vartheta^{t-s} e^{-\gamma\left(c_{t-1}^s(i) - y_{t-1}\right)} e^{-\gamma \mu_t w_t\left(\xi_t^s(i) - \overline{\xi}\right)} di \\
&\quad + (1-\vartheta)\left\{ \eta^d \int e^{-\gamma\left(\mathcal{C}_t^d - y_t + \mu_t w_t\left(\xi_t^t(i) - \overline{\xi}\right)\right)} di + \left(1-\eta^d\right) \int e^{-\gamma\left(\mathcal{C}_t^{nd} - y_t + \mu_t w_t\left(\xi_t^t(i) - \overline{\xi}\right)\right)} di \right\} \\
&= \vartheta(1-\vartheta) \sum_{s=-\infty}^{t-1} e^{\frac{\gamma^2 \mu_t^2 w_t^2 \sigma_t^2}{2}} \int \vartheta^{t-1-s} e^{-\gamma\left(c_{t-1}^s(i) - y_{t-1}\right)} di \\
&\quad + (1-\vartheta) e^{\frac{\gamma^2 \mu_t^2 w_t^2 \sigma_t^2}{2}} \left[\eta^d e^{-\gamma\left(\mathcal{C}_t^d - y_t\right)} + \left(1-\eta^d\right) e^{-\gamma\left(\mathcal{C}_t^{nd} - y_t\right)} \right] \\
&= \left[\vartheta \Sigma_{t-1} + (1-\vartheta) \mathbb{B}_t\right] e^{\frac{\gamma^2 \mu_t^2 w_t^2 \sigma_t^2}{2}}
\end{aligned}$$

or

$$\ln \Sigma_t = \frac{\gamma^2 \mu_t^2 w_t^2 \sigma_t^2}{2} + \ln\left[\vartheta \Sigma_{t-1} + (1-\vartheta) \mathbb{B}_t\right] \qquad \text{for} \qquad t > 0$$

where $\mathbb{B}_t = \eta^d e^{-\gamma\left(\mathcal{C}_t^d - y_t\right)} + \left(1-\eta^d\right) e^{-\gamma\left(\mathcal{C}_t^{nd} - y_t\right)}$. Given the properties, of $\mathcal{C}_t^d$ and $\mathcal{C}_t^{nd}$, we have:

$$\mathbb{B}_t = \mathbb{B}\left(\mu_t \mathcal{V}_t\right) \equiv \eta^d e^{-\gamma\left(\frac{1-\eta^d}{\eta^d}\right)\mu_t \mathcal{V}_t} + \left(1-\eta^d\right) e^{\gamma \mu_t \mathcal{V}_t}$$

At date 0, since the utilitarian planner sets $\tau_0^a = 1$, there is no pre-existing wealth inequality and $x_0^s(i) = w_0\left(\xi_0^s(i) - \overline{\xi}\right)$ for stockholders and nonstockholders born at some date $s \leq 0$. Thus, we have:

$$\begin{aligned}
\Sigma_0 &= (1-\vartheta) \sum_{s=-\infty}^{0} \int \vartheta^{-s} e^{-\gamma\left(c_0^s(i) - y_0\right)} di \\
&= (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{-s} e^{\frac{1}{2}\gamma^2 \mu_0^2 w_0^2 \sigma_0^2} \left\{\eta^d e^{-\gamma\left(\mathcal{C}_0^d - y_0\right)} + \left(1-\eta^d\right) e^{-\gamma\left(\mathcal{C}_0^{nd} - y_0\right)}\right\} \\
&= e^{\frac{1}{2}\gamma^2 \mu_0^2 w_0^2 \sigma_0^2} \mathbb{B}_0
\end{aligned}$$

or

$$\ln \Sigma_0 = \frac{1}{2}\gamma^2 \mu_0^2 w_0^2 \sigma_0^2 + \ln \mathbb{B}\left(\mu_0 \mathcal{V}_0\right)$$

Note that $\mathbb{B}(0) = 1, \mathbb{B}'(0) = 0$ and $\mathbb{B}''(0) = \gamma^2 \left(\frac{1-\eta^d}{\eta^d}\right) > 0$

G.2 Planning problem

The planner maximizes

$$\mathbb{W}_0 = \sum_{t=0}^{\infty} \beta^t \left\{-\frac{1}{\gamma}(1+\gamma\rho w_t) e^{-\gamma y_t} \Sigma_t\right\}$$

s.t.

$$\begin{aligned}
\gamma y_t &= \gamma y_{t+1} - \ln \beta\vartheta + \ln \mu_{t+1} + \ln\left[\mu_t^{-1} - (1+\gamma\rho w_t)\right] - \frac{\gamma^2 \mu_{t+1}^2 w^2 \sigma^2}{2} e^{2\varphi(y_{t+1} - y)} \\
\left(\Pi_t - 1\right)\Pi_t &= \frac{\varepsilon_t}{\Psi}\left[1 - \frac{\varepsilon(\varepsilon_t - 1)}{(\varepsilon-1)\varepsilon_t} \frac{(1-\tau^w) z_t}{w_t}\right] + \beta\left(\frac{z_t y_{t+1} w_{t+1}}{z_{t+1} y_t w_t}\right)\left(\Pi_{t+1} - 1\right)\Pi_{t+1} \\
\ln \Sigma_0 &= \frac{\gamma^2 \mu_t^2 w^2 \sigma^2}{2} e^{2\varphi(y_t - y)} + \ln \mathbb{B}(\mu_0 \mathcal{V}_0) \qquad \text{for } t = 0 \\
\ln \Sigma_t &= \frac{\gamma^2 \mu_t^2 w^2 \sigma^2}{2} e^{2\varphi(y_t - y)} + \ln\left[(1-\vartheta)\mathbb{B}(\mu_t \mathcal{V}_t) + \vartheta \Sigma_{t-1}\right] \qquad \text{for } t > 0 \\
y_t &= z_t \frac{\rho \ln w_t + \overline{\xi}}{1 + \gamma\rho z_t + \frac{\Psi}{2}\left(\Pi_t - 1\right)^2} \\
\mathcal{V}_t &= \left[1 - \frac{\varepsilon - 1}{\varepsilon(1-\tau^w)} \frac{w_t}{z_t}\right] y_t - \frac{\varepsilon - 1}{\varepsilon(1-\tau^w)} \frac{w_t}{z_t} \frac{\Psi}{2}\left(\Pi_t - 1\right)^2 y_t - \frac{\eta}{1-\eta} J + \left[\frac{\mu_t^{-1} - 1 - \gamma\rho w_t}{\mu_{t+1}^{-1}}\right] \mathcal{V}_{t+1}
\end{aligned}$$

where $1 - \tau^w$ is determined by (29) in the paper. The first order condition for $\mathcal{V}_t$ for $t > 0$ is:

$$0 = M_{3,t} \frac{(1-\vartheta)\mu_t \mathbb{B}'\left(\mu_t \mathcal{V}_t\right)}{(1-\vartheta)\mathbb{B}\left(\mu_t \mathcal{V}_t\right) + \vartheta \Sigma_{t-1}} - M_{5,t} + \beta^{-1} \frac{\vartheta}{R_{t-1}} M_{5,t-1}$$

In steady state $\mathcal{V}_t = 0$, and thus we have $M_5 = 0$ in steady state since $\mathbb{B}'(0) = 0$, where $M_{5,t}$ denotes the multiplier on the $\mathcal{V}_t$ recursion. Taking the rest of the first order conditions and linearizing around the

steady state in which the average consumption of stockholders and nonstockholders is equal, we have the following.
FOC wrt w_t:

$$-\gamma y\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\widehat{y}_t+\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\widehat{\Sigma}_t-\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}}\right)\widehat{m}_{1,t}-m_1\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}}\right)^2\frac{\gamma\rho}{1+\gamma\rho}\widehat{w}_t$$
$$-\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}^2}\right)m_1\widehat{\mu}_t+\frac{\kappa}{\gamma}\widehat{m}_{2,t}-\frac{\widehat{m}_{4,t}}{\gamma}+\frac{m_4}{\gamma}\widehat{w}_t-\frac{m_4}{\gamma}\frac{1}{(1+\gamma\rho)}\widehat{z}_t-\frac{1+\gamma\rho}{\gamma\rho}\widehat{m}_{5,t}\frac{y\varepsilon}{\varepsilon-1}=0 \quad \text{(G.9)}$$

FOC wrt y_t:

$$-\frac{\gamma\rho\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)}{1+\gamma\rho}\widehat{w}_t+\gamma\left[1+2\frac{(1-\Theta)^2}{\Lambda}\left(m_3-\frac{m_1}{\beta}\right)\right]\widehat{y}_t-\frac{\widehat{\Sigma}_t}{\Sigma}-\widehat{m}_{1,t}+\frac{\Theta}{\beta}\widehat{m}_{1,t-1}$$
$$+2(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)\widehat{\mu}_t+(1-\Theta)\,\widehat{m}_{3,t}+\frac{\widehat{m}_{4,t}}{\gamma}+\frac{\widehat{m}_{5,t}}{\gamma}\frac{1}{\varepsilon}=0 \quad \text{(G.10)}$$

FOC wrt Σ_t:

$$\frac{\gamma\rho w}{1+\gamma\rho w}\widehat{w}_t-\gamma\widehat{y}_t-\widehat{m}_{3,t}+\widetilde{\beta}\widehat{m}_{3,t+1}+\frac{1-\beta^{-1}\widetilde{\beta}^2}{1-\widetilde{\beta}}\widehat{\Sigma}_t=0 \quad \text{(G.11)}$$

FOC wrt Π_t:

$$\widehat{m}_{2,t}=\frac{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)}{1-\beta^{-1}\widetilde{\beta}\,(1-\Lambda)}\left(1+\frac{\Omega}{1-\widetilde{\beta}}\right)\frac{(\gamma y)\,\Psi}{1+\gamma\rho}\widehat{p}_t \quad \text{(G.12)}$$

FOC wrt μ_t:

$$-\left(\frac{1-\widetilde{\beta}+\Omega}{\widetilde{\beta}^2}\right)\frac{\gamma\rho}{1+\gamma\rho}m_1\widehat{w}_t+\left[2\Lambda\left(m_3-\frac{m_1}{\beta}\right)-\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}m_1\right]\widehat{\mu}_t+\Lambda\widehat{m}_{3,t}$$
$$+2\gamma(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)\widehat{y}_t-\frac{1}{\widetilde{\beta}}\left(\widehat{m}_{1,t}-\frac{\widetilde{\beta}}{\beta}(1-\Lambda)\,\widehat{m}_{1,t-1}\right)=0 \quad \text{(G.13)}$$

FOC wrt $\mathcal{V}_t$:

$$\gamma^2\frac{\mu^2}{1-\widetilde{\beta}}\left(\frac{1-\eta^d}{\eta^d}\right)\widehat{\mathcal{V}}_0-\widehat{m}_{5,0}=0 \quad \text{for} \quad t=0$$
$$\gamma^2\frac{\mu^2}{1-\widetilde{\beta}}\frac{1-\vartheta}{1-\vartheta+\vartheta\Sigma}\left(\frac{1-\eta^d}{\eta^d}\right)\widehat{\mathcal{V}}_t-\widehat{m}_{5,t}+\beta^{-1}\widetilde{\beta}\widehat{m}_{5,t-1}=0 \quad \text{for} \quad t>0 \quad \text{(G.14)}$$

where $\widehat{m}_{5,t}=\widehat{M_{5,t}}/\mathbb{U}$. Following the same steps as in Appendix E.4.1, we can arrive at the following expression which is the analog of equations (E.37)-(E.38) in that Appendix:

$$\Upsilon(\Omega)\,x_t+\varepsilon\widehat{p}_t=-\frac{\rho}{m_4}\left(\frac{\partial D}{\partial y}\right)\widehat{m}_{5,t}$$

where $x_t = \widehat{y}_t - \delta(\Omega)\frac{y+\rho}{1+\gamma\rho}\widehat{z}_t$. Next, for $t = 0$, combining this expression with equation (G.14), one gets the target criterion for date $t = 0$:

$$\Upsilon\left(\Omega\right)x_0 + \varepsilon\widehat{p}_0 + \mathbb{K}(\eta^d)\left(\frac{\partial D}{\partial y}\right)\left(\frac{\widehat{\mathcal{V}}_0}{y}\right) = 0$$

where $\mathbb{K}(\eta^d) = \gamma\rho\frac{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}{\left(1-\beta^{-1}\widetilde{\beta}\right)(1-\Lambda)\left(1-\widetilde{\beta}+\Omega\right)}\left(\frac{1-\eta^d}{\eta^d}\right)\mu^2 \geq 0$. Similarly for dates $t > 0$ we have:

$$\Upsilon\left(\Omega\right)\left(x_t - \frac{\widetilde{\beta}}{\beta}x_{t-1}\right) + \varepsilon\left(\widehat{p}_t - \frac{\widetilde{\beta}}{\beta}\widehat{p}_{t-1}\right) + \mathbb{K}(\eta^d)\left(1 - \frac{\widetilde{\beta}}{\beta}\right)\left(\frac{\partial D}{\partial y}\right)\left(\frac{\widehat{\mathcal{V}}_t}{y}\right) = 0$$

which is the same as in Proposition 7 in the paper. Clearly, $\mathbb{K}(1) = 0$ and $\mathbb{K}'(\eta^d) = -\frac{\rho}{m_4}\frac{\gamma^2\mu^2}{1-\widetilde{\beta}}\left(\frac{1}{\eta^d}\right)^2 < 0$.

Finally, it is easy to see that with no idiosyncratic risk ($\sigma = 0 \Rightarrow \Omega = 0$), the target criterion becomes:

$$\begin{aligned} x_0 + \varepsilon\widehat{p}_0 + \mathbb{K}(\eta^d)\left(\frac{\partial D}{\partial y}\right)\left(\frac{\widehat{\mathcal{V}}_0}{y}\right) &= 0 \quad \text{for} \quad t = 0 \\ \left(x_t - \frac{\widetilde{\beta}}{\beta}x_{t-1}\right) + \varepsilon\left(\widehat{p}_t - \frac{\widetilde{\beta}}{\beta}\widehat{p}_{t-1}\right) + \mathbb{K}(\eta^d)\left(1 - \frac{\widetilde{\beta}}{\beta}\right)\left(\frac{\partial D}{\partial y}\right)\left(\frac{\widehat{\mathcal{V}}_t}{y}\right) &= 0 \quad \text{for} \quad t > 0 \end{aligned}$$

As is clear, even in this case, the target criterion is different from RANK and there is a motive to stabilize $\mathcal{V}_t$ since $\mathbb{K} \neq 0$.

G.3 LQ representation

Relative to the derivation of the LQ problem in our baseline model in Appendix E.2, the only difference is that unequally distributed profits introduce an additional term in the second-order Σ_t recursion, which can now be written as:

$$\begin{aligned} \mathrm{g}_t^{\Sigma} \approx\ & \Lambda\widehat{\mu}_t + \gamma y(1-\Theta)\widehat{y}_t + \beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1} - \widehat{\Sigma}_t + \frac{1}{2}\widehat{\Sigma}_t^2 - \frac{1}{2}\left(\beta^{-1}\widetilde{\beta}\right)^2\widehat{\Sigma}_{t-1}^2 + (\gamma y)^2\frac{(1-\Theta)^2}{\Lambda}\widehat{y}_t^2 \\ & + 2\gamma y\left(1-\Theta\right)\widehat{\mu}_t\widehat{y}_t + \frac{\Lambda}{2}\widehat{\mu}_t^2 + \frac{1}{2}\left(\gamma y\right)^2\left(\frac{1-\eta^d}{\eta^d}\right)\mu^2\left[\mathbb{I}(t=0)\widetilde{\mathcal{V}}_0^2 + \mathbb{I}(t>0)(1-\beta^{-1}\widetilde{\beta})\widetilde{\mathcal{V}}_t^2\right] \end{aligned} \tag{G.15}$$

The rest of the equations remain unchanged. Thus, the purely second-order approximation to the planner's objective is as described in (E.22) plus the additional terms involving $\widetilde{V}_t$ (multiplied by m_3). Thus, following the same steps above, we can arrive at the same expression as in Proposition 7 in the paper:

$$\frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left\{\Upsilon\left(\Omega\right)\left(\widehat{y}_t - \delta(\Omega)\widehat{y}_t^e\right)^2 + \frac{\varepsilon}{\kappa}\pi_t^2\right\} + \frac{\mathbb{K}\left(\eta^d\right)}{2}\left\{\widetilde{\mathcal{V}}_0^2 + \sum_{t=1}^{\infty}\beta^t\left(1-\beta^{-1}\widetilde{\beta}\right)\widetilde{\mathcal{V}}_t^2\right\} \tag{G.16}$$

The optimal policy problem can now simply be specified as minimizing (G.16) subject to the linearized Phillips curve (30) in the paper and valuation equation (40) in the paper. In Lagrangian form:

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left\{\Upsilon(\Omega)\left(\widehat{y}_t-\delta(\Omega)\widehat{y}^e_t\right)^2+\frac{\varepsilon}{\kappa}\pi_t^2\right\}+\frac{\mathbb{K}\left(\eta^d\right)}{2}\left\{\widetilde{\mathcal{V}}_0^2+\sum_{t=1}^{\infty}\beta^t\left(1-\beta^{-1}\widetilde{\beta}\right)\widetilde{\mathcal{V}}_t^2\right\}\\ &+\sum_{t=0}^{\infty}\beta^t F_{1,t}\left\{\beta\pi_{t+1}+\kappa\left[\widehat{y}_t-\widehat{y}^e_t+\frac{\rho/y}{1+\gamma\rho}\widehat{\varepsilon}_t\right]-\pi_t\right\}\\ &+\sum_{t=0}^{\infty}\beta^t F_{2,t}\left\{\mathcal{D}_y\widehat{y}_t+\frac{\varepsilon-1}{\varepsilon}\frac{1+\gamma\rho}{\rho/y}\widehat{y}^e_t+\widetilde{\beta}\widetilde{\mathcal{V}}_{t+1}-\widetilde{\mathcal{V}}_t\right\}\end{aligned}$$

The FOC w.r.t. $\widehat{y}_t$ can be written as:

$$\Upsilon(\Omega)\left(\widehat{y}_t-\delta(\Omega)\widehat{y}^e_t\right)+\kappa F_{1,t}+F_{2,t}\mathcal{D}_y=0$$

The FOC w.r.t. π_t can be written as:

$$\frac{\varepsilon}{\kappa}\pi_t-F_{1,t}+F_{1,t-1}=0 \qquad \Leftrightarrow \qquad F_{1,t}=\frac{\varepsilon}{\kappa}\widehat{p}_t$$

where $\kappa=\frac{\varepsilon}{\Psi}\frac{1+\gamma\rho}{\rho/y}$. Finally the FOC w.r.t. $\mathcal{V}_t$ can be written as:

$$\begin{aligned}\mathbb{K}(\eta^d)\widetilde{\mathcal{V}}_0-F_{2,0}&=0 &&\text{for } t=0\\ \left(1-\beta^{-1}\widetilde{\beta}\right)\mathbb{K}(\eta^d)\widetilde{\mathcal{V}}_0-F_{2,t}+\beta^{-1}\widetilde{\beta}F_{2,t-1}&=0 &&\text{for } t>0\end{aligned}$$

Combining these three FOCs, we can derive the target criterion in Proposition 7 in the paper:

$$\Upsilon(\Omega)\,x_0+\varepsilon\widehat{p}_0+\mathbb{K}(\eta^d)\mathcal{D}_y\widehat{\mathcal{V}}_0 \;=\; 0 \tag{G.17}$$

and for $t>0$:

$$\Upsilon(\Omega)\left(x_t-\beta^{-1}\widetilde{\beta}x_{t-1}\right)+\varepsilon\left(\widehat{p}_t-\beta^{-1}\widetilde{\beta}\widehat{p}_{t-1}\right)+\mathbb{K}(\eta^d)\mathcal{D}_y\left(1-\beta^{-1}\widetilde{\beta}\right)\widehat{\mathcal{V}}_t \;=\; 0 \tag{G.18}$$

where $x_t=\widehat{y}_t-\delta(\Omega)\widehat{y}^e_t$.

H Hand to Mouth households

Our baseline model deliberately abstracts from MPC heterogeneity and shows that even absent such heterogeneity, optimal policy sharply differs from RANK. We now study how MPC heterogeneity, a feature of quantitative HANK models that has received much attention since Kaplan et al. (2018), affects optimal monetary policy. We do so by introducing a fraction η^h of hand-to-mouth (HtM) households who cannot trade bonds and consume their after tax-income. These households are otherwise identical to the remaining $1-\eta^h$ unconstrained households who trade bonds as in the baseline – in particular, both groups draw idiosyncratic shocks from the same distribution and receive the same dividends and transfers per capita.

While the MPC of unconstrained households μ_t is still described by (18) in the paper, the MPC of constrained households is $\widetilde{\mu}_t = (1 + \gamma\rho w_t)^{-1}$. These households can still self-insure to some extent by adjusting hours worked, implying that $\widetilde{\mu}_t < 1$. However, since they cannot insure using the bond market, their MPC is higher than that of the unconstrained households, i.e. $\widetilde{\mu}_t > \mu_t$.

Appendix H.1 and H.2 show that the presence of HtM households does not change the dynamics of aggregate variables, given a path of interest rates. These dynamics are still given by (23)-(25) in the paper – in equilibrium, since HtM households consume their income, and aggregate consumption equals aggregate income, the average consumption of unconstrained households must equal aggregate income as in our baseline.[2] However, introducing HtM households does affect social welfare, and therefore optimal policy. While the period t felicity function of the utilitarian planner can still be written as $\mathbb{U}_t = u(c_t, n_t; \overline{\xi}) \times \Sigma_t$, the welfare relevant measure of consumption inequality is now $\Sigma_t = (1-\eta^h)\Sigma_t^{nh} + \eta^h \Sigma_t^h$ where Σ_t^{nh} denotes consumption inequality among unconstrained households and evolves according to (27) in the paper, while Σ_t^h denotes consumption inequality among HtM households, and equals $\Sigma_t^h = \frac{1}{2}\gamma^2 \widetilde{\mu}_t^2 w_t^2 \sigma_t^2$. Since there is no wealth inequality among HtM households, unlike Σ_t^{nh}, Σ_t^h depends only on *current* consumption risk. However, since $\widetilde{\mu}_t > \mu_t$, consumption inequality moves more for this group in response to changes in income risk. While the tradeoffs facing the planner are qualitatively the same as in our baseline economy, quantitatively, monetary policy has even larger effects on Σ_t in the presence of HtM households:

Lemma 4. *The effect of a one-time increase in output engineered by monetary policy reduces inequality Σ_t by a larger amount, the larger the fraction of HtM households η^h: $\frac{\partial^2 \widehat{\Sigma}_t}{\partial \widehat{y}_t \partial \eta^h} < 0$ when income risk is acyclical or countercyclical.*

Proof. See Appendix H.3. □

Since the main differences in optimal policy in HANK relative to RANK arise because monetary policy can affect inequality, a higher sensitivity of inequality to monetary policy magnifies these differences.

Productivity Shock Figure H.1 shows the dynamics under optimal policy following a negative productivity shock in RANK (dashed red curves), HANK with no HtMs (solid blue curves) and HANK with 30% HtMs (dot-dashed magenta curves).[3] In our baseline ($\eta^h = 0$), monetary policy already prevents output from falling as much as $\widehat{y}_t^n$ on impact, permitting some inflation. With $\eta^h > 0$, policy cushions the fall in output even more (see panel a), resulting in even higher inflation responses initially (see panel b). Quantitatively, the impact response of the output gap is about twice as large with HtM households, and that of inflation about two and a half times as large. Intuitively, a fall in output is more costly with $\eta^h > 0$ because it increases consumption inequality more for HtMs who cannot self-insure using the bond market. This can be seen by comparing the dot-dashed magenta curves in panel c), which plots consumption inequality amongst unconstrained households, with panel d) which plots inequality among the HtMs. At its peak, the percentage increase in Σ_t^h is around ten times the increase in Σ_t^{nh}. Thus, the benefit of mitigating the fall in output, in terms of the effect on Σ_t, is much higher in the economy with HtMs. To see this, compare the dot-dashed magenta curve in panel e), which plots inequality under optimal policy with 30% HtMs, to the dotted-black curve, which plots inequality if monetary policy uses the target criterion which

[2] This is for the same reasons as in Bilbiie (2008); Werning (2015); Acharya and Dogra (2020).

[3] $\eta^h = 0.3$ is in line with Kaplan et al. (2014) who find that approximately 30% of U.S. households are hand-to-mouth. Given our calibration, this implies an average MPC of around 17% (around 40% for HtMs and 7% for unconstrained households), which is in line with the range of MPCs reported in the empirical literature.

would be optimal in an economy with no HtMs. The difference between these curves – the reduction in overall inequality due to a higher path of output – is much larger than the reduction in inequality amongst the unconstrained households, shown by the difference between the curves in panel c). Since inequality is more sensitive to the level of output in the presence of HtMs, the planner tolerates larger deviations from productive efficiency and price stability to mitigate the rise in inequality following an adverse shock.

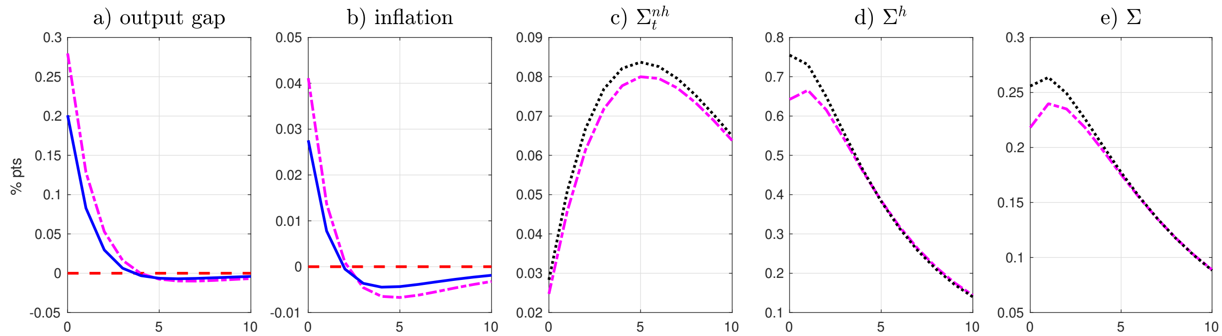

Figure H.1: **Optimal policy in response to productivity shocks** In panels a and b, solid blue curves depicts dynamics in HANK with $\Omega > 0$ and no HtM agents; red-dashed curves depict dynamics in RANK; and dot-dashed magenta lines depict the optimal response of an economy with 30% HtM households following a negative productivity shock. In panels c,d and e, the dot-dashed magenta line presents the evolution of $\Sigma_t^{nh}, \Sigma_t^h$ and Σ_t resp. under optimal policy in the economy with 30% HtMs, while the dotted-black line depicts the evolution of these variables in the economy with 30% HtMs if monetary policy implements the target criterion which would be optimal in an economy with no HtMs. All panels plot log-deviations from steady state $\times 100$.

Markup Shocks Similarly, when studying markup shocks in our HANK economy with HtMs, the difference between optimal policy in HANK and RANK is qualitatively the same as in our baseline, but quantitatively amplified. To mitigate the increase in inequality, particularly amongst HtMs, monetary policy stabilizes output more (dot-dashed magenta curve relative to solid blue curve in panel a), Figure H.2) at the cost of higher inflation (dot-dashed magenta curve relative to solid blue curve in panel b)). Quantitatively, in the presence of HtMs, optimal policy shaves off around half the initial fall in output in RANK while optimal policy only shaves off about a quarter in our baseline (absent HtMs). Similarly, the increase in inflation is larger with HtMs.

Overall, introducing MPC heterogeneity does not qualitatively change the tradeoffs analyzed in our baseline. In fact, it accentuates the differences relative to RANK: with higher MPCs, i.e., higher passthrough from income to consumption risk, consumption inequality is even more sensitive to monetary policy. Consequently, policy deviates even further from RANK to stabilize inequality. This suggests that the tradeoffs we study analytically would be even more important in quantitative HANK economies with a substantial fraction of high MPC households.

H.1 Decision problem of HtM households

A HtM agent's problem at any date t can be written as:

$$\max_{c_t^s(i;h),\ell_t^s(i;h)} -\frac{1}{\gamma}e^{-\gamma c_t^s(i;h)} - \rho e^{\rho(\ell_t^s(i)-\xi_t^s(i))}$$

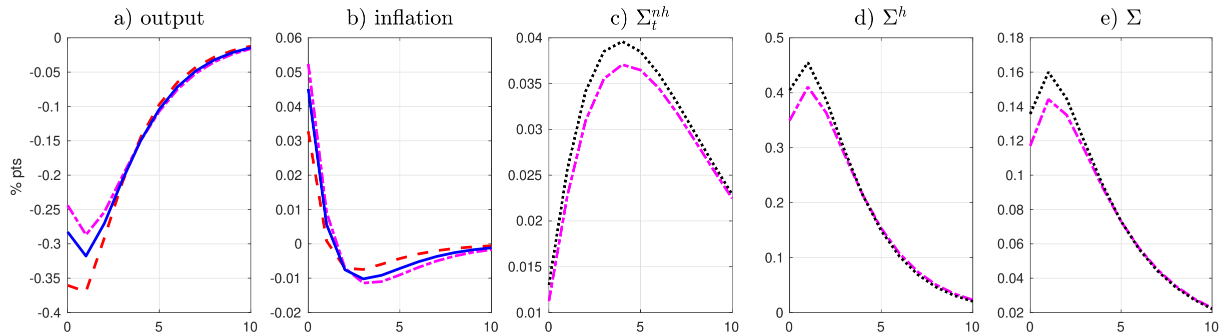

Figure H.2: **Optimal policy in response to markup shocks** In panels a and b, solid blue curves depicts dynamics in HANK with $\Omega > 0$ and no HtM agents; red-dashed curves depict dynamics in RANK; and dot-dashed magenta lines depict the optimal response of an economy with 30% HtM households following a positive markup shock. In panels c,d and e, the dot-dashed magenta line presents the evolution of $\Sigma_t^{nh}, \Sigma_t^h$ and Σ_t resp. under optimal policy in the economy with 30% HtMs, while the dotted-black line depicts the evolution of these variables in the economy with 30% HtMs if monetary policy implements the target criterion which would be optimal in an economy with no HtMs. All panels plot log-deviations from steady state $\times 100$.

s.t.

$$c_t^s(i;h) = w_t \ell_t^s(i;h) + D_t - T_t$$

The optimal labor supply can be written as:

$$\ell_t^s(i;h) = \rho \ln w_t - \gamma\rho c_t^s(i;h) + \xi_t^s(i;h) \tag{H.1}$$

which is the same as that for the non-HtM households (16) in the paper. Aggregating the individual labor supply across all HtM and non-HtM households, multiplying by w_t and adding $D_t - T_t$:

$$w_t \ell_t + D_t - T_t = w_t \ln w_t - \gamma\rho w_t y_t + w_t \overline{\xi} + D_t - T_t$$

The LHS of this expression is simply y_t, so we have

$$y_t = \frac{w_t \left(\ln w_t + \overline{\xi}\right) + D_t - T_t}{1 + \gamma\rho w_t}$$

Using this and the individual labor supply in the budget constraint for HtM households yields:

$$c_t^s(i;h) = y_t + \widetilde{\mu}_t x_t^s(i;h)$$

where $x_t^s(i;h) = w_t\left(\xi_t^s(i) - \overline{\xi}\right)$ and $\widetilde{\mu}_t = (1 + \gamma\rho w_t)^{-1}$.

Since the average consumption of HtM households is y_t, market clearing implies that the average consumption of unconstrained households is also $\mathcal{C}_t^{nh} = y_t$. Thus, it follows that the same aggregate Euler equation as in the baseline still holds with a fraction $\eta^h > 0$ of HtM households.

H.2 Deriving the Σ recursion

Even in this case, the objective function of the planner can be written as:

$$\mathbb{W}_0 = \sum_{t=0}^{\infty} \beta^t u\left(c_t, n_t; \overline{\xi}\right) \Sigma_t$$

where, as before, Σ_t is defined by:

$$\Sigma_t = (1-\vartheta) \sum_{s=-\infty}^{t} \int \vartheta^{t-s} e^{-\gamma(c_t^s(i) - c_t)} di$$

Since we have HtM and non-HtM households, this can be further expanded:

$$\Sigma_t = (1-\eta^h) \underbrace{(1-\vartheta) \sum_{s=-\infty}^{t} \int \vartheta^{t-s} e^{-\gamma(c_t^s(i;nh) - y_t)} di}_{\Sigma_t^{nh}} + \eta^h \underbrace{(1-\vartheta) \sum_{s=-\infty}^{t} \int \vartheta^{t-s} e^{-\gamma(c_t^s(i;h) - y_t)} di}_{\Sigma_t^{h}}$$

Since $c_s^t(i;h) = y_t + \widetilde{\mu}_t w_t \left(\xi_t^t(i) - \overline{\xi}\right)$, we have Σ_t^h:

$$\Sigma_t^h = (1-\vartheta) \sum_{s=-\infty}^{t} \vartheta^{t-s} \int e^{-\gamma \widetilde{\mu}_t w_t (\xi_t^s(i;h) - \overline{\xi})} = e^{\frac{1}{2}\gamma^2 \widetilde{\mu}_t^2 w_t^2 \sigma_t^2}$$

Since the consumption function of unconstrained households is the same as in the baseline model, it follows that Σ_t^{nh} evolves as:

$$\ln \Sigma_t^{nh} = \frac{\gamma^2 \mu_t^2 w_t^2 \sigma_t^2}{2} + \ln[1 - \vartheta + \vartheta \Sigma_{t-1}^{nh}]$$

H.3 Sensitivity of inequality w.r.t. monetary policy with HTMs

In the presence of HTMs, the welfare relevant measure of inequality at any date t (up to first order) is given by:

$$\widehat{\Sigma}_t = (1-\eta^h) \frac{\Sigma^{nh}}{\Sigma} \widehat{\Sigma}_t^{nh} + \eta^h \frac{\Sigma^h}{\Sigma} \widehat{\Sigma}_t^h$$

where (31) in the paper describes the evolution of $\widehat{\Sigma}_t^{nh}$. Up to first order, the relationship between $\Sigma_t^h = \frac{1}{2}\left(\frac{\gamma w_t \sigma_t}{1+\gamma \rho w_t}\right)^2$ and y_t can be expressed as:

$$\widehat{\Sigma}_t^h = -\frac{\gamma y}{\left(1 - \widetilde{\beta}\right)^2} \left[(\Theta - 1 + \Lambda) + \Lambda \left(\frac{w-1}{1+\gamma \rho w}\right)\right] \widehat{y}_t$$

where we have used the equilibrium relationship between wages and output (E.3) (we have also set all shocks to zero without loss of generality). Thus, we have:

$$\begin{aligned}\widehat{\Sigma}_t &= -\frac{\gamma y}{\left(1-\widetilde{\beta}\right)^2}\left\{\left(1-\widetilde{\beta}\right)^2\left(1-\eta^h\right)\frac{\Sigma^{nh}}{\Sigma}(\Theta-1)+\eta^h\frac{\Sigma^h}{\Sigma}\left[(\Theta-1+\Lambda)+\Lambda\left(\frac{w-1}{1+\gamma\rho w}\right)\right]\right\}\widehat{y}_t \\ &\quad +\left(1-\eta^h\right)\frac{\Sigma^{nh}}{\Sigma}\Lambda\widehat{\mu}_t+\left(1-\eta^h\right)\frac{\Sigma^{nh}}{\Sigma}\beta^{-1}\widetilde{\beta}\widehat{\Sigma}_t^{nh}\end{aligned}$$

We consider a one-time change in $\widehat{y}_t > 0$ engineered by monetary policy. Since equations (23)-(25) in the paper which describe the evolution of macroeconomic aggregates are purely forward looking, monetary policy can implement this with a change in the nominal interest rate only at date t without affecting the trajectory of macroeconomic aggregates in the future. The change in nominal rates which implement this one time increase in date t output can be derived by setting all $t+1$ variables (and all shocks) in (23)-(24) in the paper to zero:

$$\begin{aligned}\widehat{y}_t &= -\frac{1}{\gamma y}i_t \\ \widehat{\mu}_t &= -\gamma\mu w y(1+\gamma\rho)\widehat{y}_t+\widetilde{\beta}i_t\end{aligned}$$

where the first equation is (23) and the second is (24) in the paper. Combining the three equations and eliminating $\widehat{w}_t$ yields

$$\widehat{\mu}_t = -\gamma y\left[1+\left(1-\widetilde{\beta}\right)\left(\frac{w-1}{1+\gamma\rho w}\right)\right]\widehat{y}_t$$

Using this in the expression for $\widehat{\Sigma}_t$ yields

$$\begin{aligned}\widehat{\Sigma}_t &= -\gamma y\left(1-\eta^h\right)\frac{\Sigma^{nh}}{\Sigma}\left[(\Theta-1+\Lambda)+\Lambda\left(1-\widetilde{\beta}\right)\left(\frac{w-1}{1+\gamma\rho w}\right)\right]\widehat{y}_t \\ &\quad -\gamma y\eta^h\frac{\Sigma^h}{\Sigma}\frac{1}{\left(1-\widetilde{\beta}\right)^2}\left[(\Theta-1+\Lambda)+\Lambda\left(\frac{w-1}{1+\gamma\rho w}\right)\right]\widehat{y}_t+\left(1-\eta^h\right)\frac{\Sigma^{nh}}{\Sigma}\beta^{-1}\widetilde{\beta}\widehat{\Sigma}_t^{nh},\end{aligned}$$

Taking the derivative w.r.t η^h, we get:

$$\frac{\partial^2\widehat{\Sigma}_t}{\partial\eta^h\partial\widehat{y}_t} = -\gamma y\left\{\left[\Theta-1+\Lambda+\Lambda\left(\frac{w-1}{1+\gamma\rho w}\right)\right]\frac{1}{\Sigma}\left(\frac{\Sigma^h}{\left(1-\widetilde{\beta}\right)^2}-\Sigma^{nh}\right)+\Lambda\left(\frac{w-1}{1+\gamma\rho w}\right)\widetilde{\beta}\frac{\Sigma^{nh}}{\Sigma}\right\},$$

which is negative for countercyclical and acyclical risk ($\Theta \geq 1$) for β sufficiently close to 1.[4] Thus, a higher fraction of HTMs (η^h) implies that Σ_t falls more in response to the same increase in output.

[4] To see this, note that $\Sigma^h\frac{1}{\left(1-\widetilde{\beta}\right)^2}-\Sigma^{nh} = \frac{e^{\frac{\Lambda}{2}}\left(1-\beta\vartheta e^{\frac{\Lambda}{2}}\right)^{-2}}{\left(1-\beta\vartheta e^{\frac{\Lambda}{2}}\right)^2}-\frac{1-\vartheta}{1-\vartheta e^{\frac{\Lambda}{2}}}e^{\frac{\Lambda}{2}}$ is increasing in β, negative at $\beta=0$ and positive at $\beta=1$ for any ϑ,Λ satisfying $\vartheta e^{\Lambda/2}<1$.

H.4 Planning Problem

The utilitarian planner maximizes:

$$\mathbb{W}_0 = \sum_{t=0}^{\infty} \beta^t \left\{ -\frac{1}{\gamma}(1+\gamma\rho w_t) e^{-\gamma y_t} \left[(1-\eta^h)\Sigma_t^{nh} + \eta^h \Sigma_t^h \right] \right\}$$

s.t.

$$\begin{aligned}
\gamma y_t &= \gamma y_{t+1} - \ln \beta\vartheta + \ln \mu_{t+1} + \ln\left[\mu_t^{-1} - (1+\gamma\rho w_t)\right] - \frac{\gamma^2 \mu_{t+1}^2 w^2 \sigma^2}{2} e^{2\varphi(y_{t+1}-y)} \\
(\Pi_t - 1)\,\Pi_t &= \frac{\varepsilon_t}{\Psi}\left[1 - \frac{\varepsilon(\varepsilon_t - 1)}{(\varepsilon-1)\varepsilon_t}\frac{(1-\tau^w)z_t}{w_t}\right] + \beta\left(\frac{z_t y_{t+1} w_{t+1}}{z_{t+1} y_t w_t}\right)(\Pi_{t+1}-1)\,\Pi_{t+1} \\
\ln \Sigma_t^{nh} &= \frac{\gamma^2 \mu_t^2 w^2 \sigma^2}{2} e^{2\varphi(y_t - y)} + \ln\left[(1-\vartheta) + \vartheta \Sigma_{t-1}^{nh}\right] \\
\ln \Sigma_t^{h} &= \frac{\gamma^2 (1+\gamma\rho w_t)^{-2} w^2 \sigma^2}{2} e^{2\varphi(y_t - y)} \\
y_t &= z_t \frac{\rho \ln w_t + \overline{\xi}}{1 + \gamma\rho z_t + \frac{\Psi}{2}(\Pi_t - 1)^2}
\end{aligned}$$

Fiscal policy sets τ^w such that the planner finds it optimal to implement $\Pi = 1$ in steady state, as in our baseline. To plot Figures H.1 and H.2, we first solve for τ^w numerically, then we linearize the first order conditions and compute the optimal dynamics to shocks numerically.

I Persistent income risk

Our baseline model described in the main paper featured i.i.d. idiosyncratic income risk, whereas empirical studies find that idiosyncratic income risk is highly persistent (Heathcote et al., 2010; Guvenen et al., 2021). We now relax this assumption by allowing for persistent idiosyncratic disutility shocks. Specifically, we assume that

$$\xi_t^s(i) - \overline{\xi} = \sigma_t e_t^s(i) \qquad \text{where} \qquad e_t^s(i) = \varrho_\xi e_{t-1}^s(i) + v_t^s(i), \qquad v_t^s(i) \sim i.i.d.\ N(0,1), \qquad e_{s-1}^s(i) = 0 \tag{I.1}$$

We allow for $0 \leq \varrho_\xi \leq 1$. Setting $\varrho_\xi = 0$ corresponds to the baseline model. As in the baseline model, we allow for a flexible specification for the cyclicality of income risk by assuming that $w_t\sigma_t = w\sigma e^{\varphi(y_t - y)}$. Appendix I.1 shows that the optimal consumption decision rule of a household is described by

$$c_t^s(i) = \mathcal{C}_t + \mu_t\Big(a_t^s(i) + h_t^s(i)\Big) \tag{I.2}$$

and the aggregate Euler equation is now given by

$$\mathcal{C}_t = -\frac{1}{\gamma}\ln \beta R_t + \mathcal{C}_{t+1} - \frac{\gamma \mu_{t+1}^2 \sigma_{h,t+1}^2}{2} \tag{I.3}$$

where $a_t^s(i)$ is the household's financial wealth and $h_t^s(i) \equiv \sigma_{h,t} e_t^s(i)$ denotes the household's human wealth, defined as the expected present-discounted value of their labor endowment

$$h_t^s(i) = \mathbb{E}_t \sum_{\tau=0}^{\infty} Q_{t+\tau|t} w_{t+\tau} \left(\xi_{t+\tau}^s(i) - \overline{\xi}\right) = \underbrace{\left[\sum_{\tau=0}^{\infty} Q_{t+\tau|t} w_{t+\tau} \sigma_{t+\tau} \varrho_\xi^\tau\right]}_{\sigma_{h,t}} e_t^s(i) \tag{I.4}$$

where $Q_{t+\tau|t} = \prod_{k=0}^{\tau-1} \frac{\vartheta}{R_{t+k}}$. As in the baseline model, the MPC out of household financial and human wealth, μ_t is still given by (18). The consumption risk faced by households, the last term in (I.3) depends on the passthrough from human wealth to consumption (measured by μ_{t+1}^2) and the variance of shocks to human wealth $\sigma_{h,t+1}^2$. In our baseline model ($\varrho_\xi = 0$), human wealth $h_t^s(i)$ is simply $w_t(\xi_t^s - \overline{\xi})$, making (I.2) identical to (15) in the paper, and the variance of shocks to human wealth is simply $\sigma_{h,t+1}^2 = w_{t+1}^2 \sigma_{t+1}^2$. However, with persistent idiosyncratic income, a positive shock to the household's current labor endowment also increases the expected value of their endowment in the future. This is reflected in the fact that $\sigma_{h,t}$ depends on not just $w_t \sigma_t$, but the whole future path $\{w_{t+k}\sigma_{t+k}\}_{k=0}^{\infty}$.

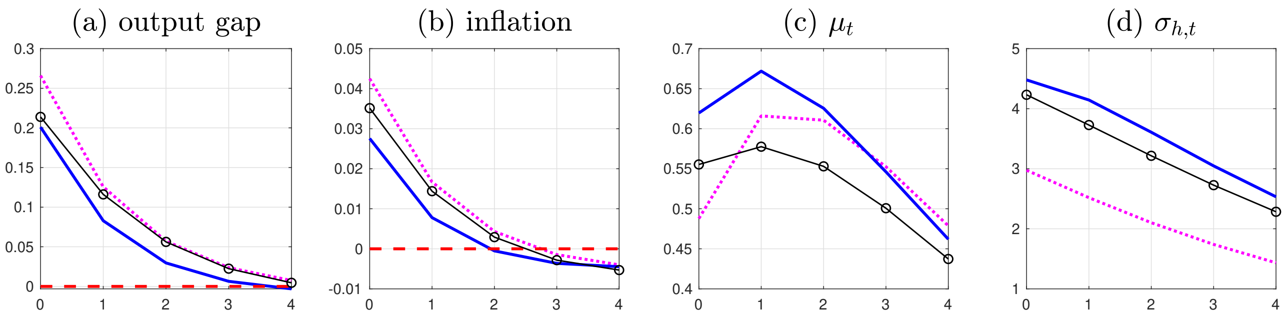

Figure I.1: **Optimal policy in response to productivity shocks** In all panels, red-dashed curves depict dynamics in RANK; solid blue curves depicts dynamics in HANK with $\varrho_\xi = 0$; black lines with circle markers depicts dynamics in HANK with $\varrho_\xi = 0.5$ and the magenta dotted line depicts dynamics in HANK with $\varrho_\xi = 1$. All panels plot log-deviations from steady state $\times 100$.

Appendix I.2 shows that, as in our baseline, a utilitarian planner's felicity function can be decomposed into the flow utility of a notional representative agent and a welfare-relevant measure of consumption inequality Σ_t, which now evolves according to

$$\ln \Sigma_t = \frac{\gamma^2 \mu_t^2 \sigma_{h,t}^2}{2} + \ln\left[1 - \vartheta + \vartheta \Sigma_{t-1}\right] \tag{I.5}$$

It is worth nothing that with $\varrho_\xi > 0$, our economy features not one, but two dimensions of persistent wealth inequality: financial and human wealth inequality. In principle, this means that the planner must forecast the evolution of the joint distribution of financial and human wealth, not just the distribution of financial wealth as in the baseline. However, as (I.5) indicates, the evolution of this joint distribution can still be summarized by a single scalar Σ_t which depends on its own lagged value. This highlights the analytical tractability of our framework.

Equation (I.5) along with the definition of σ_h in (I.4) reveals that persistence ($\varrho_\xi > 0$) modifies the effect of monetary policy on consumption inequality in two ways. First, lower real interest rates, holding the path of aggregate output and wages fixed, now tend to increase the variance of human wealth σ_h^2, putting more weight on the value of future labor endowments. Thus, while the effect of interest rates on passthrough μ_t remain unchanged (relative to the baseline i.i.d. case), $\varrho_\xi > 0$ tends to weaken the overall effect of interest rates on consumption risk, given the level of output. But this is not the only effect of higher persistence. Lower real interest rates also increase output, which reduces human capital risk (in the countercyclcial income risk case) as in our baseline model. This effect becomes more pronounced, the higher the level of human capital risk σ_h. Higher ϱ_ξ tends to increase the level of human capital risk (for the same sequence of $\{w_{t+k}, \sigma_{t+k}\}_{k=0}^{\infty}$), since the same shock to current income has a larger effect on lifetime income: $\sigma_{h,t}(\varrho_\xi > 0) > \sigma_{h,t}(\varrho_\xi = 0)$. Thus, higher ϱ_ξ amplifies the effect of monetary policy on Σ_t via the level of output. Overall, this second effect dominates and higher persistence increases the sensitivity of Σ_t to changes in output induced by monetary policy. This effect is itself long-lived– $\frac{\partial \widehat{\Sigma}_{t+k}}{\partial \widehat{y}_t}$ is larger in absolute value at all horizons $k > 0$ when ϱ_ξ is higher– because consumption inequality is only slow to revert to its mean value following an increase in consumption risk (cf. equation (I.5)).

Lemma 5. *The effect of a one-time increase in output engineered by monetary policy reduces inequality Σ_{t+k} at all horizons $k \geq 0$ by a larger amount, the larger the persistence of idiosyncratic income ϱ_ξ:*

$$\frac{\partial}{\partial \varrho_\xi}\left(\frac{\partial \widehat{\Sigma}_t}{\partial \widehat{y}_t}\right) < 0 \qquad \textit{with acyclical/countercyclical income risk, } \Theta \geq 1$$

and

$$\frac{\partial}{\partial \varrho_\xi}\left(\frac{\partial \widehat{\Sigma}_{t+k}}{\partial \widehat{y}_t}\right) = \left(\beta^{-1}\widetilde{\beta}\right)^k \frac{\partial}{\partial \varrho_\xi}\left(\frac{\partial \widehat{\Sigma}_t}{\partial \widehat{y}_t}\right) \qquad \forall k > 0$$

Proof. See Appendix I.4. □

Consequently, since the sensitivity of consumption risk to monetary policy is the main force leading optimal monetary policy to differ in HANK and RANK, introducing persistent idiosyncratic income risk magnifies these differences. Figure I.1 shows the dynamics under optimal policy following a negative productivity shock in RANK (dashed red curves), and HANK with $\varrho_\xi = 0$ (blue line), $\varrho_\xi = 0.5$ (black line with circle markers) and $\varrho_\xi = 1$ (magenta dotted line). Recall that in our baseline with $\varrho_\xi = 0$, the HANK planner already cushions the fall in output relative to the RANK planner, resulting in higher inflation on impact. The black line with circle markers and magenta dotted line indicate that higher ϱ_ξ leads the HANK planner to cushion the fall in output even more, leading to higher inflation on impact. To understand why, note that the steady state *level* of human capital risk σ_h is the highest for the economy with $\varrho_\xi = 1$ and the lowest when $\varrho_\xi = 0$. Thus, panel (d) shows that by curtailing the fall in output, the HANK planner permits a smaller *proportional* increase in $\sigma_{h,t}$ when the level of σ_h is already high, i.e., in the economy with $\varrho_\xi = 1$ (compare the magenta dotted and blue lines). The planner does not allow a large increase in the level of $\sigma_{h,t}$, even temporarily, since doing so would persistently increase consumption inequality Σ_t (cf. Lemma 5). This more moderate decline in output (and smaller proportional increase in $\sigma_{h,t}$) also results in a smaller increase in passthrough μ_t (panel (c)). The case with $\varrho_\xi = 0.5$ lies between the i.i.d and random walk extremes. A higher ϱ_ξ modifies the optimal response to a markup shock in a similar fashion; we omit the results for the sake of brevity.

Overall, persistent income risk, like MPC heterogeneity, does not change the tradeoff facing the planner qualitatively. In fact, it also accentuates the difference relative to RANK, compared to the case with i.i.d.

income risk. Again this suggests that the tradeoffs we study analytically would be even more important in quantitative HANK models with realistic income processes.

I.1 Derivation of household decision rules

The date s problem of an individual i born at date s is now

$$\max_{c_t^s(i),\ell_t^s(i),a_{t+1}^s(i)} -\mathbb{E}_s \sum_{t=s}^{\infty} (\beta\vartheta)^{t-s} \left\{ \frac{1}{\gamma} e^{-\gamma c_t^s(i)} + \rho e^{\frac{1}{\rho}(\ell_t^s(i)-\xi_t^s(i))} \right\}$$

subject to

$$c_t^s(i) + q_t a_{t+1}^s(i) = w_t \ell_t^s(i) + (1-\tau_t^a) a_t^s(i) + D_t - T_t$$

where $\xi_t^s(i) - \overline{\xi} = \sigma_t e_t^s(i)$ and

$$e_t^s(i) = \varrho_\xi e_{t-1}^s(i) + \upsilon_t^s(i) \qquad \upsilon_{i,t} \sim N(0,1)$$

The derivation of the consumption function follows that in Appendix A. Guess that the consumption function takes the form:

$$c_t^s(i) = \mathcal{C}_t + \mu_t \left(a_t^s(i) + h_t^s(i)\right)$$

where $h_t^s(i)$ denotes the expected present-discounted value of the household's labor endowment:

$$h_t^s(i) = \mathbb{E}_t \sum_{k=0}^{\infty} Q_{t+k|t} w_{t+k} \left(\xi_{t+k}^s(i) - \overline{\xi}\right) \equiv \sigma_{h,t} e_t^s(i) \qquad Q_{t+k|t} = \prod_{k=0}^{\tau} \frac{\vartheta}{R_{t+k}}$$

Using the budget constraint, labor supply and the household's Euler equation, we have:

$$\begin{aligned} y_t + \left\{\mu_t - \mu_{t+1}\frac{R_t}{\vartheta}\left[1-(1+\rho\gamma w_t)\mu_t\right]\right\} a_t^s(i) &= -\frac{1}{\gamma}\ln\beta R_t + y_{t+1} \\ &+ \left\{\mu_{t+1}\frac{R_t}{\vartheta}\left[\sigma_t w_t - (1+\rho\gamma w_t)\mu_t\sigma_{h,t}\right] + \varrho_\xi \mu_{t+1}\sigma_{h,t+1} - \mu_t \sigma_{h,t}\right\} e_t^s(i) \\ &- \frac{\gamma}{2}\mu_{t+1}^2 \sigma_{h,t+1}^2 \end{aligned}$$

Matching coefficients yields the standard μ_t recursion:

$$\mu_t^{-1} = 1 + \gamma\rho w_t + \frac{\vartheta}{R_t}\mu_{t+1}^{-1} \tag{I.6}$$

In addition, we have the following equation describing $\sigma_{h,t}$

$$\sigma_{h,t} = \sigma_t w_t + \frac{\vartheta}{R_t}\varrho_\xi \sigma_{h,t+1} \tag{I.7}$$

and the aggregate Euler equation is now given by

$$y_t = y_{t+1} - \frac{1}{\gamma} \ln \beta R_t - \frac{\gamma \mu_{t+1}^2 \sigma_{h,t+1}^2}{2} \tag{I.8}$$

where we have used $\mathcal{C}_t = y_t$ from market clearing.

I.2 Deriving the Σ recursion

As in our baseline, we assume that the planner is utilitarian and puts identical weight (equal to 1) on the welfare of all individuals on individual i both at date $s \leq 0$ and β^s on the welfare of individuals who will be born at date $s > 0$. Recall that in our baseline we allow the planner to set a date 0 tax on financial wealth to focus on the role of monetary policy in providing insurance, rather than redistribution between borrowers and lenders. But when $\varrho_\xi > 0$, households alive at the beginning of date 0 differ not only in financial wealth but also in terms of human wealth. To remove the planner's incentive to use monetary policy to redistribute between individuals with high and low human wealth, we allow the planner to tax indivduals on their total wealth at the beginning of date 0. At the beginning of date 0, when the date 0 idiosyncratic shock $\upsilon_0^s(i)$ to household i's time endowment has not yet been realized, the household's total wealth is given by $a_0^s(i) + \sigma_{h,0}\varrho_\xi e_{-1}^s(i)$ where we have used the fact that

$$h_0^s(i) \quad = \quad \sigma_{h,0} e_0^s(i) = \sigma_{h,0}\varrho_\xi e_{-1}^s(i) + \sigma_{h,0}\upsilon_0^s(i)$$

The planner levies a tax τ_0^aon this total amount implying that the hosuehold's post-tax human wealth after the realization of their date 0 idiosyncratic shock $\upsilon_0^s(i)$ is given by $(1-\tau_0^a)\left[a_0^s(i) + \sigma_{h,0}\varrho_\xi e_{-1}^s(i)\right] + \sigma_{h,0}\upsilon_0^s(i)$. This also implies that the date 0 tax on financial wealth $\tau_0^a = 1$. However, δ now measures the extent to which the planner is willing to tolerate pre-existing human wealth inequality. $\delta = 0$ implies that the planner is also utilitarian towards human wealth inequality at date 0 while a higher δ implies that the planner assigns higher weights to the welfare of those with higher human wealth as of date -1. As in the baseline, the planner's objective function can be written as:

$$\mathbb{W}_0 = \sum_{t=0}^{\infty} \beta^t u\left(c, n; \overline{\xi}\right) \Sigma_t$$

where Σ_t is now defined as

$$\Sigma_t \quad = \quad (1-\vartheta) \sum_{s=-\infty}^{t} \vartheta^{t-s} \int e^{-\gamma\left(c_t^s(i) - c_t\right)} di$$

Next, subtracting the aggregate Euler equation from a household's Euler equation for all dates $t \geq 0$, we get

$$c_{t+1}^s(i) - c_{t+1} = c_t^s(i) - c_t + \mu_{t+1}\sigma_{h,t+1}\upsilon_{t+1}^s(i)$$

Using this in the definition of Σ_t for $t \geq 1$

$$\begin{aligned}
\Sigma_t &= (1-\vartheta) \sum_{s=-\infty}^{t-1} \vartheta^{t-s} \int e^{-\gamma\left(c_{t-1}^s(i) - c_{t-1} + \mu_t \sigma_{h,t} \upsilon_t^s(i)\right)} di + (1-\vartheta) \int e^{-\gamma\left(\mu_t \sigma_{h,t} \upsilon_t^t(i)\right)} di \\
&= \vartheta e^{\frac{1}{2}\gamma^2 \mu_t^2 \sigma_{h,t}^2} \left\{ (1-\vartheta) \sum_{s=-\infty}^{t-1} \vartheta^{t-1-s} \int e^{-\gamma\left(c_{t-1}^s(i) - c_{t-1}\right)} di \right\} + (1-\vartheta)\, e^{\frac{1}{2}\gamma^2 \mu_t^2 \sigma_{h,t}^2} \\
&= e^{\frac{1}{2}\gamma^2 \mu_t^2 \sigma_{h,t}^2} \left[1 - \vartheta + \vartheta \Sigma_{t-1}\right]
\end{aligned}$$

Taking logs, we get

$$\ln \Sigma_t = \frac{\gamma^2 \mu_t^2 \sigma_{h,t}^2}{2} + \ln\left[1 - \vartheta + \vartheta \Sigma_{t-1}\right]$$

Next, for $t = 0$, we have

$$\begin{aligned}
\Sigma_0 &= (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{-s} \int e^{-\gamma\left(c_0^s(i) - c_0\right)} di \\
&= (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{-s} \int e^{-\gamma \mu_0 \left(1-\tau_0^a\right)\left(a_0^s(i) + \sigma_{h,0} \varrho_\xi e_{-1}^s(i)\right) - \gamma \mu_0 \sigma_{h,0} \upsilon_0^s(i)} di \\
&= e^{\frac{1}{2}\gamma^2 \mu_0^2 \sigma_{h,0}^2} \left\{ (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{-s} \int e^{-\gamma \mu_0 \left(1-\tau_0^a\right)\left(a_0^s(i) + \sigma_{h,0} \varrho_\xi e_{-1}^s(i)\right)} di \right\}
\end{aligned}$$

Clearly, since $a_0^s(i) + \sigma_{h,0} \varrho_\xi e_{-1}^s(i)$ has zero mean, the planner chooses $\tau_0^a = 1$ to minimize this expression, implying that the date 0 Σ recursion is the same as at all future dates:

$$\ln \Sigma_0 = \frac{1}{2}\gamma^2 \mu_0^2 \sigma_{h,0}^2 + \ln\left[1 - \vartheta + \vartheta \Sigma_{-1}\right] \qquad \text{where} \qquad \Sigma_{-1} = 1$$

I.3 Planning Problem

The planning problem can be written as:

$$\max \sum_{t=0}^{\infty} \beta^t \left\{ -\frac{1}{\gamma} \left(1 + \gamma \rho w_t\right) e^{-\gamma y_t} \Sigma_t \right\}$$

s.t.

$$\gamma y_t = \gamma y_{t+1} - \ln \beta\vartheta + \ln \mu_{t+1} + \ln\left[\mu_t^{-1} - \left(1 + \gamma \rho w_t\right)\right] - \frac{\gamma^2 \mu_{t+1}^2 \sigma_{h,t+1}^2}{2} \tag{I.9}$$

$$\left(\Pi_t - 1\right) \Pi_t = \frac{\varepsilon_t}{\Psi} \left[1 - \frac{\varepsilon_t - 1}{\varepsilon_t} \frac{\left(1 - \tau^w\right) z_t}{\left(1 - \tau^\star\right) w_t}\right] + \beta \left(\frac{z_t y_{t+1} w_{t+1}}{z_{t+1} y_t w_t}\right) \left(\Pi_{t+1} - 1\right) \Pi_{t+1} \tag{I.10}$$

$$\ln \Sigma_t = \frac{\gamma^2 \mu_t^2 \sigma_{h,t}^2}{2} + \ln\left[1 - \vartheta + \vartheta \Sigma_{t-1}\right] \tag{I.11}$$

$$y_t = z_t \frac{\rho \ln w_t + \bar{\xi}}{1 + \rho\gamma z_t + \frac{\Psi}{2}(\Pi_t - 1)^2} \tag{I.12}$$

$$\sigma_{h,t} = \sigma_t w_t + \varrho_\xi \mu_{t+1}\left[\mu_t^{-1} - 1 - \gamma\rho w_t\right]\sigma_{h,t+1} \tag{I.13}$$

$$\Sigma_{-1} = 1 \tag{I.14}$$

This can be expressed as a Lagrangian:

$$\begin{aligned}
\mathcal{L} \quad = \quad & \sum_{t=0}^{\infty}\beta^t\left\{-\frac{1}{\gamma}(1+\gamma\rho w_t)e^{-\gamma y_t}\Sigma_t\right\} \\
&+\sum_{t=0}^{\infty}\beta^t M_{1,t}\left\{\gamma y_{t+1} - \ln\beta\vartheta + \ln\mu_{t+1} + \ln\left[\mu_t^{-1} - (1+\gamma\rho w_t)\right] - \frac{\gamma^2\mu_{t+1}^2\sigma_{h,t+1}^2}{2} - \gamma y_t\right\} \\
&+\sum_{t=0}^{\infty}\beta^t M_{2,t}\left\{\frac{\varepsilon_t}{\Psi}\left[1 - \frac{\varepsilon_t - 1}{\varepsilon_t}\frac{(1-\tau^w)z_t}{(1-\tau^\star)w_t}\right] + \beta\left(\frac{z_t w_{t+1} y_{t+1}}{z_{t+1} w_t y_t}\right)(\Pi_{t+1}-1)\Pi_{t+1} - (\Pi_t - 1)\Pi_t\right\} \\
&+\sum_{t=0}^{\infty}\beta^t M_{3,t}\left\{\frac{\gamma^2\mu_t^2\sigma_{h,t}^2}{2} + \ln\left[1 - \vartheta + \vartheta\Sigma_{t-1}\right] - \ln\Sigma_t\right\} \\
&+\sum_{t=0}^{\infty}\beta^t M_{4,t}\left\{y_t - z_t\frac{\rho\ln w_t + \bar{\xi}}{1+\gamma\rho z_t + \frac{\Psi}{2}(\Pi_t - 1)^2}\right\} \\
&+\sum_{t=0}^{\infty}\beta^t M_{5,t}\left\{\sigma w e^{\varphi(y_t - y)} + \varrho_\xi\mu_{t+1}\left[\mu_t^{-1} - 1 - \gamma\rho w_t\right]\sigma_{h,t+1} - \sigma_{h,t}\right\}
\end{aligned}$$

FOC wrt y_t (equation is divided by $-\gamma$):

$$\begin{aligned}
&\mathbb{U}_t + M_{1,t} - \beta^{-1}M_{1,t-1} + \beta\gamma M_{2,t}\left(\frac{z_t w_{t+1} y_{t+1}}{z_{t+1} w_t y_t^2}\right)(\Pi_{t+1}-1)\Pi_{t+1} - \gamma M_{2,t-1}\left(\frac{z_{t-1} w_t}{z_t w_{t-1} y_{t-1}}\right)(\Pi_t - 1)\Pi_t \\
&-\frac{M_{4,t}}{\gamma} - \frac{\varphi}{\gamma}M_{5,t} w\sigma e^{\varphi(y_t - y)} = 0
\end{aligned}$$

FOC wrt w_t (equation is multiplied by w_t):

$$\begin{aligned}
0 \quad = \quad & \frac{\gamma\rho w_t}{1+\gamma\rho w_t}\mathbb{U}_t - M_{1,t}\frac{\gamma\rho w_t}{\mu_t^{-1} - (1+\gamma\rho w_t)} + M_{2,t}\frac{\varepsilon_t - 1}{\Psi}\frac{(1-\tau^w)z_t}{(1-\tau^\star)w_t} \\
&-\beta M_{2,t}\left(\frac{z_t w_{t+1} y_{t+1}}{z_{t+1} w_t y_t}\right)(\Pi_{t+1}-1)\Pi_{t+1} + M_{2,t-1}\left(\frac{z_{t-1} w_t y_t}{z_t w_{t-1} y_{t-1}}\right)(\Pi_t - 1)\Pi_t \\
&-\frac{M_{4,t}}{\gamma}\frac{\gamma\rho z_t}{1+\rho\gamma z_t + \frac{\Psi}{2}(\Pi_t - 1)^2} - M_{5,t}\varrho_\xi\mu_{t+1}\gamma\rho w_t\sigma_{h,t+1}
\end{aligned}$$

FOC wrt Σ_t (equation is multiplied by Σ_t):

$$M_{3,t} = -\frac{1}{\gamma}\left(1+\gamma\rho w_t\right)e^{-\gamma y_t}\Sigma_t + \beta\frac{\vartheta\Sigma_t}{1-\vartheta+\vartheta\Sigma_t}M_{3,t+1}$$

FOC wrt μ_t (equation is multiplied by μ_t):

$$-M_{1,t}\frac{\mu_t^{-1}}{\mu_t^{-1}-(1+\gamma\rho w_t)} + \beta^{-1}M_{1,t-1} - \left(\beta^{-1}M_{1,t-1}-M_{3,t}\right)\gamma^2\mu_t^2\sigma_{h,t}^2 - \varrho_\xi\frac{\mu_{t+1}}{\mu_t}M_{5,t}\sigma_{h,t+1}$$
$$+\beta^{-1}\varrho_\xi\mu_t\left[\mu_{t-1}^{-1}-1-\gamma\rho w_{t-1}\right]M_{5,t-1}\sigma_{h,t} = 0$$

FOC wrt $\sigma_{h,t}$ (equation is multiplied by $\sigma_{h,t}$):

$$0 = -\left(\beta^{-1}M_{1,t-1}-M_{3,t}\right)\gamma^2\mu_t^2\sigma_{h,t}^2 - M_{5,t}\sigma_{h,t} + \beta^{-1}\varrho_\xi\mu_t\left[\mu_{t-1}^{-1}-1-\gamma\rho w_{t-1}\right]M_{5,t-1}\sigma_{h,t}$$

FOC wrt Π_t:

$$\left[M_{2,t} - \left(\frac{z_{t-1}w_t y_t}{z_t w_{t-1} y_{t-1}}\right)M_{2,t-1}\right](2\Pi_t - 1) = M_{4,t}z_t\frac{y_t}{1+\gamma\rho z_t + \frac{\Psi}{2}(\Pi_t-1)^2}\Psi(\Pi_t-1)$$

Fiscal policy sets τ^w such that the planner finds it optimal to implement $\Pi = 1$ in steady state, as in our baseline. We solve this system numerically, taking a first-order approximation of the first-order conditions and the constraints (linearizing the multipliers and log-linearizing all other variables).

I.4 Proof of Lemma 5

We consider a one time change in output $\widehat{y}_t > 0$ engineered by monetary policy. Since the equations (I.9), (I.12) and (I.13) are forward looking, monetary policy can implement this with a change in nominal interest rates only at date t without affecting macroeconomic aggregates in the future. Thus, the response of the other variables to a one time change in $\widehat{y}_t$ are given by the solution to the following linearized equations, where we have imposed that all variables return to their steady state values at date $t+1$ (except for $\widehat{\Sigma}_{t+1}$):

$$\gamma y\widehat{y}_t = -\frac{\mu^{-1}}{\mu^{-1}-(1+\gamma\rho w)}\widehat{\mu}_t - \frac{\gamma\rho w}{\mu^{-1}-(1+\gamma\rho w)}\widehat{w}_t$$

$$\widehat{w}_t = \frac{1+\gamma\rho}{\rho/y}\widehat{y}_t$$

$$\widehat{\sigma}_{h,t} = \frac{\sigma w}{\sigma_h}\varphi y\widehat{y}_t - \varrho_\xi\widehat{\mu}_t - \varrho_\xi\mu\gamma\rho w\widehat{w}_t$$

Using the steady state relationships between these variables, we have:

$$\widehat{\mu}_t = -\gamma y\left[\widetilde{\beta} + \left(1-\widetilde{\beta}\right)\frac{(1+\gamma\rho)\,w}{1+\gamma\rho w}\right]\widehat{y}_t$$

$$\widehat{\sigma}_{h,t} = \gamma y\left[\left(1-\widetilde{\beta}\varrho_\xi\right)\frac{\varphi}{\gamma} + \varrho_\xi\left[\widetilde{\beta} + \left(1-\widetilde{\beta}\right)\frac{(1+\gamma\rho)\,w}{1+\gamma\rho w}\right] - \left(1-\widetilde{\beta}\right)\varrho_\xi\frac{(1+\gamma\rho)\,w}{1+\gamma\rho w}\right]\widehat{y}_t$$

Finally, log-linearizing (I.11)

$$\begin{aligned}\widehat{\Sigma}_t &= \gamma^2\mu^2\sigma_h^2\left(\widehat{\mu}_t+\widehat{\sigma}_{h,t}\right)+\beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1}\\ &= -\left(\gamma y\right)\gamma^2\mu^2w^2\sigma^2\left(\frac{1}{1-\widetilde{\beta}\varrho_\xi}\left(1-\frac{\varphi}{\gamma}\right)+\left(\frac{1-\widetilde{\beta}}{1-\widetilde{\beta}\varrho_\xi^2}\right)\frac{w-1}{1+\gamma\rho w}\right)\widehat{y}_t+\beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1}\end{aligned}$$

Thus, we have:

$$\frac{\partial}{\partial\varrho_\xi}\left(\frac{\partial\widehat{\Sigma}_t}{\partial\widehat{y}_t}\right)=-\widetilde{\beta}\gamma y\left[\left(\Theta-1+\Lambda\right)+2\left(\frac{\Lambda\Omega}{1-\widetilde{\beta}\varrho_\xi}\right)\right]$$

where $\Lambda=\frac{\gamma^2\mu^2w^2\sigma^2}{1-\widetilde{\beta}\varrho_\xi^2}$ and $\Theta=1-\frac{\varphi\Lambda}{\gamma}$ and $\Omega=(1-\widetilde{\beta})\frac{w-1}{1+\gamma\rho w}$. With countercyclical risk and $w>1$, this derivative is negative, implying that higher ϱ_ξ increases the sensitivity of $\widehat{\Sigma}_t$ to $\widehat{y}_t$ (in absolute value). Given that $\widehat{y}_{t+k}=0$ for $k>0$ in the experiment considered, we have

$$\begin{aligned}\frac{\partial}{\partial\varrho_\xi}\left(\frac{\partial\widehat{\Sigma}_{t+k}}{\partial\widehat{y}_t}\right) &= \frac{\partial}{\partial\varrho_\xi}\left(\frac{\partial\widehat{\Sigma}_{t+k}}{\partial\widehat{\Sigma}_t}\frac{\partial\widehat{\Sigma}_t}{\partial\widehat{y}_t}\right)\\ &= \left(\beta^{-1}\widetilde{\beta}\right)^k\left\{-\widetilde{\beta}\gamma y\left[\left(\Theta-1+\Lambda\right)+2\left(\frac{\Lambda\Omega}{1-\widetilde{\beta}\varrho_\xi}\right)\right]\right\}\end{aligned}$$

J Optimal response to demand shocks

In Section 4 of the paper, we focused on productivity and markup shocks, both of which affect the natural level of output y_t^n. The RANK literature also studies the optimal response to other shocks which do not affect y_t^n, e.g. changes in households' discount factor. Following the literature, we term these *demand* shocks. Since these shocks do not induce a tradeoff between productive efficiency and price stability, the RANK planner simply implements $\widehat{y}_t=\widehat{y}_t^n=\pi_t=0$ in response to these shocks by setting the interest rate equal to the *natural rate of interest* $r_t^\star$, i.e. the interest rate consistent with $y_t=y_t^n$ at all dates.

As shown in Section 4, the HANK planner generally does not implement $y_t=y_t^n$, even in response to productivity shocks which do not induce a tradeoff between productive efficiency and price stability. This is because responding one-for-one to fluctuations in the natural level of output would adversely affect inequality. Similarly, in response to demand shocks, setting $y_t=y_t^n$ is in general not optimal, because these shocks would affect inequality should monetary policy fully insulate output from them. Consequently, optimal policy lets output vary in order to offset these undesirable changes in inequality.

We study two demand shocks: (i) changes in households' discount factor and (ii) shocks to the variance of idiosyncratic shocks faced by households. We now assume that household preferences are given by:

$$\mathbb{E}_s\sum_{t=s}^{\infty}\left(\beta\vartheta\right)^{t-s}\Big(\prod_{k=s}^{t-1}\zeta_k\Big)u\Big(c_t^s(i),\ell_t^s(i);\xi_t^s(i)\Big)$$

where ζ_t is a shock to the individual's discount factor between dates t and $t+1$. Appendix A shows that Proposition 1 in the paper remains true except that the aggregate Euler equation (17) in the paper becomes:

$$\mathcal{C}_t=-\frac{1}{\gamma}\ln\beta\zeta_tR_t+\mathcal{C}_{t+1}-\frac{\gamma\mu_{t+1}^2w_{t+1}^2\sigma_{t+1}^2}{2}$$

The preference shock is internalised by the utilitarian planner who puts weight $\beta^s\left(\prod_{k=0}^{s-1}\zeta_k\right)$ on the lifetime utility of a household born at date $s > 0$.

We also introduce a shock to the variance of idiosyncratic risk faced by households (ξ) by assuming that this variance satisfies $\sigma_t^2 w_t^2 = \sigma^2 w^2 \exp\{2[\varphi(y_t - y) + \varsigma_t]\}$. Higher ς_t increases the cross-sectional variance of cash-on-hand at date t. To the extent that the shock is persistent ($\varrho_\varsigma > 0$), this can also be thought of as a *risk shock*: higher ς_{t+1} increases the uncertainty households face at date t about the realization of the shock to disutility (and hence to cash-on-hand) at date $t+1$. When plotting IRFs, following Bayer et al. (2020), we set the persistence and standard deviation of risk shocks and discount factor shocks to $\varrho_\varsigma = 0.68^4$, $\varrho_\zeta = 0.83^4$, $\sigma_\varsigma = 1.4$ and $\sigma_\zeta = 0.01$.

Both discount factor shocks and risk shocks affect the evolution of consumption inequality. This can be seen through the linearized Σ_t recursion (32) in the paper which now becomes:[5]

$$\widehat{\Sigma}_t = -\gamma y \Omega \widehat{y}_t - \frac{\widetilde{\beta}\Lambda}{1-\widetilde{\beta}\varrho_\zeta(1-\Lambda)}\widehat{\zeta}_t + \frac{(1-\widetilde{\beta}\varrho_\varsigma)\Lambda}{1-\widetilde{\beta}(1-\Lambda)\varrho_\varsigma}\widehat{\varsigma}_t + \beta^{-1}\widetilde{\beta}\widehat{\Sigma}_{t-1} \tag{J.1}$$

An increase in $\widehat{\varsigma}_t$ directly affects income risk and thus persistently affects consumption inequality. More subtly, a fall in households discount factor $\widehat{\zeta}_t < 0$ increase the natural rate of interest, which in our economy is given by $r_t^\star = -\frac{1-\widetilde{\beta}\varrho_\zeta}{1-\widetilde{\beta}(1-\Lambda)\varrho_\zeta}\widehat{\zeta}_t - \frac{(1-\widetilde{\beta}\varrho_\varsigma)\Lambda}{1-\widetilde{\beta}(1-\Lambda)\varrho_\varsigma}\widehat{\varsigma}_{t+1}$. Thus, if monetary policy keeps output unchanged in response to a fall in ζ_t, this entails a rise in interest rates which increases the passthrough μ_t. For a given level of income risk, higher passthrough increases consumption risk and hence the level of consumption inequality. A persistent increase in ς_t also reduces $r_t^\star$ as households attempt to increase their precautionary savings in response to the increase in risk. This decline in interest rates reduces μ_t somewhat, offsetting some of the direct effect of a higher ς_t on consumption risk. However, a higher ς_t still increases Σ_t on net.

Since demand shocks affect inequality, the planner generally deviates from keeping output equal to its natural level and implementing zero inflation (even though this remains *feasible*) in order to mitigate the impact on inequality. This is formalized in the following Proposition.[6]

Proposition J.1. *In response to demand shocks, the planner sets nominal interest rates so that the following target criterion holds at all dates $t \geq 0$:*

$$(\widehat{y}_t - y_t^\star) + \frac{\varepsilon}{\Upsilon(\Omega)}\widehat{p}_t = 0 \tag{J.2}$$

where $y_t^\star = -\chi(\Omega)\widehat{\zeta}_t + \Xi(\Omega)\widehat{\varsigma}_t$ is the desired level of output (in deviations from steady state). $\chi(\Omega)$ and $\Xi(\Omega)$ are defined in Appendix E.4.1 and satisfy $\chi(0) = \Xi(0) = 0$. $\Upsilon(\Omega)$ is the same as in Proposition 3. When risk is countercyclical ($\Theta > 1 \Rightarrow \Omega > \Omega^c$), $\chi(\Omega) > 0$ and $\Xi(\Omega) > 0$.

As described earlier, the target criterion (J.2) indicates that the planner seeks to minimize fluctuations of the price level while also keeping output close to its desired level $y_t^\star$. When risk is acyclical or countercyclical, demand shocks which tend to increase consumption inequality – higher ς_t or lower ζ_t – increase $y_t^\star$.

[5] See Appendix E.1 for a derivation. We have implicitly set $\widehat{z}_t = 0$ throughout this section.

[6] For this section, we do not derive a quadratic loss function but derive the target criterion by linearizing the non-linear first order conditions of the planner's problem. The target criterion in Proposition J.1 is a generalization of the target criterion in (36) in the paper to include demand shocks but abstracting from productivity shocks ($\widehat{z}_t = 0$). Appendix E.4.1 derives a general target criterion which is valid in the presence of all four shocks that we study.

That is, the planner targets a higher level of output because this tends to reduce consumption inequality when $\Omega \geq \Omega^c > 0$, mitigating the increase in inequality due to the shock. Since demand shocks keep y_t^n unchanged, adjusting output in response to these shocks entails some inflation; as discussed earlier, the HANK planner puts a smaller relative weight on price stability $\Upsilon(\Omega) > 1$ relative to the RANK planner.

Risk shocks We start by describing the dynamics under optimal policy in response to a risk shock $\widehat{\varsigma}_0 > 0$.

Proposition J.2. *Under optimal policy with acyclical or countercyclical income risk, following an increase in risk ($\widehat{\varsigma}_0 > 0$), $\widehat{y}_0$ and π_0 both increase. In addition, there exists $T > 0$ such that for all $t \in (T, \infty)$, $\pi_t < 0$ and $\widehat{y}_t < 0$. Following a decline in risk ($\widehat{\varsigma}_0 < 0$) all these signs are reversed.*

Figure J.1 plots the optimal response to a an increase in risk in RANK and HANK (with $\Omega \geq \Omega^c$). In RANK, since households can trade Arrow securities, an increase in the cross-sectional dispersion of income does not result in any increase in consumption inequality. Since risk shocks do not affect y_t^n, the RANK planner keeps output fixed at $\widehat{y}_t = \widehat{y}_t^n = 0$, implying zero inflation $\pi_t = 0$ (dashed red lines).

In contrast, in HANK with $\Omega \geq \Omega^c$, monetary policy cuts nominal interest rates on impact (panel e) to raise output above its natural level $\widehat{y}_0 > \widehat{y}_0^n = 0$ in response to a positive risk shock (panel a). In the acyclical or countercyclical case ($\Omega \geq \Omega^c > 0$), higher output tends to reduce consumption inequality, partially offsetting the effect of the risk shock (see equation (32) in the paper). Lower interest rates and higher output (which implies higher wages) also makes it easier for households to self insure, lowering the passthrough from income to consumption risk, i.e., $\widehat{\mu}_0 < 0$ (panel f). Monetary policy trades off the benefit from mitigating the increase in inequality against the cost of higher inflation (panel b) and productive inefficiency ($\widehat{y}_t \neq \widehat{y}_t^n$). To mitigate this inflation, the planner commits to mildly lower output and inflation in the future. If instead, monetary policy implements $\widehat{y}_t = \widehat{y}_t^n = 0$ and $\pi_t = 0$ (which was optimal under RANK), this would result in higher inequality (dotted black curve in panel c).

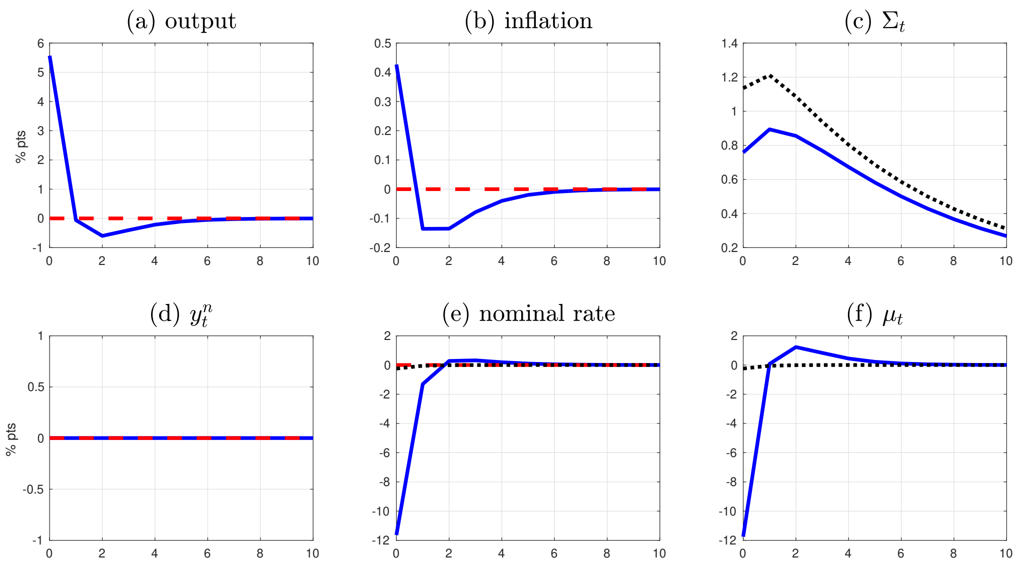

Figure J.1: **Optimal policy in response to risk shocks** in HANK with $\Omega > 0$ (solid blue curves) and RANK (dashed red curves). Black-dotted lines denote outcomes in HANK under non-optimal policy which sets $\widehat{y}_t = \widehat{y}_t^n = 0$, $\pi_t = 0$ $\forall t \geq 0$. All panels plot log-deviations from steady state $\times 100$.

Discount factor shock A decrease in households' discount factor ($\widehat{\zeta}_t < 0$) increases $r_t^\star$, the interest rate consistent with $\widehat{y}_t = \widehat{y}_t^n = 0$ and $\pi_t = 0$. Consequently, the RANK planner raises interest rates one-for-one with $r_t^\star$, keeping inflation and output unchanged. However, in HANK, this rise in interest rates would increase passthrough μ_t and hence consumption inequality. Thus, as with a positive risk shock, monetary policy deviates from the flexible-price allocation ($\widehat{y}_t = \widehat{y}_t^n = \pi_t = 0$) to mitigate this rise in inequality.

Proposition J.3. *Under optimal policy with acyclical or countercyclical income risk, following an decrease in households' discount factor ($\widehat{\zeta}_0 < 0$), $\widehat{y}_0$ and π_0 both increase. In addition, $\exists T > 0$ such that for all $t \in (T, \infty)$, $\pi_t < 0$ and $\widehat{y}_t < 0$. Following a rise in households' discount factor, all these signs are reversed.*

Figure J.2 plots the optimal dynamics following a negative discount factor shock. As in RANK, the HANK planner raises rates (panel e), increasing passthrough μ_t (panel f). This in turn tends to increase consumption inequality (panel c). However, the HANK planner does not increase rates one-for-one with $r_t^\star$ (panel d) as this would result in a larger increase in inequality (black-dotted line in panel c). This lower path of interest rates increases output on impact (panel a), reducing the level of risk faced by households (when risk is countercyclical) and further curtailing the increase in inequality. To mitigate the rise in date 0 inflation, the planner commits to lower output and inflation in the future (panel b). However, these differences relative to RANK are fairly small given our calibration.

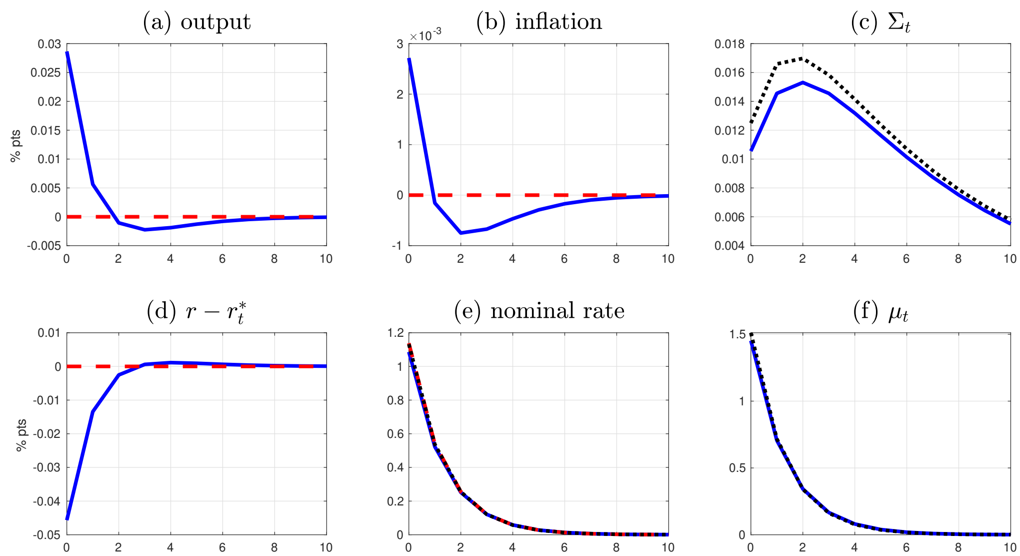

Figure J.2: **Optimal policy in response to discount factor shock** in HANK with $\Omega > 0$ (solid blue curves) and RANK (dashed red curves). Black-dotted lines denote outcomes in HANK under non-optimal policy which sets $\widehat{y}_t - \widehat{y}_t^n = \pi_t = 0 \ \forall t \geq 0$. All panels plot log-deviations from steady state $\times 100$.

Absence of self-insurance channel in zero-liquidity HANK models The optimal response to discount factor shocks highlights an important difference between our economy with $\Omega \geq \Omega^c > 0$ and zero-liquidity HANK economies (in which households cannot borrow and government debt is in zero net supply). In zero-liquidity models, interest rates do not affect households' ability to self-insure via the bond market, since they always consume their income in equilibrium. Thus, as in RANK, interest rates perform a single task in these economies: implementing the planner's desired path of output growth, which in turn affects inflation via the Phillips curve. Consequently, the planner can first choose output and inflation to maximize welfare subject to the Phillips curve, ignoring the IS curve. After this, the planner can use the IS equation to back out the interest rates implementing the desired path of output and inflation. Since discount factor shocks only affect the IS curve which can be dropped as a constraint, the planner in a zero-liquidity or RANK economy leaves output and inflation unchanged following such a shock, raising interest rates one-for-one with $r_t^{\star}$.

In our HANK economy, the IS curve cannot be dropped as a constraint since the interest rate performs two tasks: (i) it affects output via the IS curve (21) in the paper and (ii) it affects the passthrough from income to consumption risk μ_t through (18) in the paper. Formally, Appendix D.2 shows that the multiplier on the IS equation is non-zero in our HANK model but zero in RANK; it would also be 0 in a zero-liquidity HANK model. Our planner, therefore, faces a tradeoff absent in both RANK and zero-liquidity economies: when choosing what path of output to target, they must also consider how the interest rates which implement the desired path of output affect consumption inequality. Thus, in response to a negative discount factor shock, the HANK planner raises interest rates less than one-for-one with $r_t^{\star}$, tolerating higher output and inflation to curtail the rise in inequality. While this difference relative to zero-liquidity HANK models is easiest to see with discount factor shocks, the same difference is also present in response to other shocks as well. For example, one reason the planner does not let output fall as much as y_t^n following a negative productivity shock, is that this would require a steeper increase in interest rates, impairing households' ability to self-insure using the bond market.

References

Acharya, Sushant and Keshav Dogra, "Understanding HANK: Insights From a PRANK," *Econometrica*, May 2020, *88* (3), 1113–1158.

Bayer, Christian, Benjamin Born, and Ralph Luetticke, "Shocks, Frictions, and Inequality in US Business Cycles," CEPR Discussion Papers 14364 January 2020.

Benigno, Pierpaolo and Michael Woodford, "Inflation Stabilization and Welfare: The Case of a Distorted Steady State," *Journal of the European Economic Association*, 2005, *3* (6), 1185–1236.

Bilbiie, Florin O., "Limited Asset Markets Participation, Monetary Policy and (Inverted) Aggregate Demand Logic," *Journal of Economic Theory*, May 2008, *140* (1), 162–196.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song, "What do data on millions of US workers reveal about lifecycle earnings dynamics?," *Econometrica*, 2021, *89* (5), 2303–2339.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 2010, *118* (4), 681–722.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante, "Monetary Policy According to HANK," *American Economic Review*, March 2018, *108* (3), 697–743.

_ , Giovanni L. Violante, and Justin Weidner, "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, 2014, *45* (1 (Spring), 77–153.

Werning, Ivan, "Incomplete Markets and Aggregate Demand," Working Paper 21448, National Bureau of Economic Research August 2015.