Federal Reserve Bank of New York
Staff Reports

Optimal Monetary Policy According to HANK

Sushant Acharya
Edouard Challe
Keshav Dogra

Staff Report No. 916
February 2020

Optimal Monetary Policy According to HANK
Sushant Acharya, Edouard Challe, and Keshav Dogra
Federal Reserve Bank of New York Staff Reports, no. 916
February 2020
JEL classification: E21, E30, E52, E62, E63

Abstract

We study optimal monetary policy in a heterogeneous agent new Keynesian economy. A utilitarian planner seeks to reduce consumption inequality, in addition to stabilizing output gaps and inflation. The planner does so both by reducing income risk faced by households, and by reducing the pass-through from income to consumption risk, trading off the benefits of lower inequality against productive inefficiency and higher inflation. When income risk is countercyclical, policy curtails the fall in output in recessions to mitigate the increase in inequality. We uncover a new form of time inconsistency of the Ramsey plan—the temptation to exploit households' unhedged interest rate exposure to lower inequality.

Key words: new Keynesian model, incomplete markets, optimal monetary policy

---

Acharya, Dogra: Federal Reserve Bank of New York (emails: sushant.acharya@ny.frb.org, keshav.dogra@ny.frb.org). Challe: CREST and Ecole Polytechnique (email: edouard.challe@gmail.com). The authors thank Florin Bilbiie, Christopher Carroll, Russell Cooper, Clodomiro Ferreira, Antoine Lepetit, Galo Nuño, Pedro Teles, Gianluca Violante, and Pierre-Olivier Weil for helpful discussions. They also received useful comments from seminar participants at HEC Paris, UT Austin, UC3M, EUI, Université Paris-Dauphine, Université Paris 8, Banque de France, and CREST, as well as from conference participants at the Barcelona GSE Summer Forum (Monetary Policy and Central Banking), the NBER Summer Institute (Micro Data and Macro Models), the Salento Macro Meetings, SED, and T2M. Edouard Challe acknowledges financial support from the French National Research Agency (Labex Ecodec/ANR-11-LABX-0047).

To view the authors’ disclosure statements, visit https://www.newyorkfed.org/research/staff_reports/sr916.html.

1 Introduction

It is increasingly recognized by researchers and policymakers that monetary policy can have important effects on inequality. Despite this, the study of how monetary policy should be conducted optimally has only recently begun to depart from a representative agent framework in which concerns about inequality are trivially absent. While the large recent heterogeneous agent New Keynesian (HANK) literature has shown that uninsurable idiosyncratic risk and inequality can dramatically change the positive effects of monetary policy on the macroeconomy (See for example Ravn and Sterk (2017, Forthcoming); Kaplan et al. (2018); den Haan et al. (2018); Auclert et al. (2018); Auclert (2019); Bilbiie (2019a) and many others), the normative implications of HANK and the reciprocal effects of monetary policy on risk and inequality, have been less well studied. This gap in the literature partly exists because characterizing optimal policy in HANK economies is technically difficult. Solving for the Ramsey optimal policy involves choosing the evolution of an infinite dimensional state variable (the wealth distribution), as well as infinite dimensional controls (the distribution of consumption and hours worked across agents) subject to an infinite number of constraints (each household's optimality condition and budget constraints).

One approach to solving optimal policy problems for HANK economies is computational, and researchers have recently started developing numerical algorithms to handle them (Bhandari et al., 2018). We instead take an analytical approach. We study a standard NK economy with nominal rigidities with the exception that households face uninsurable idiosyncratic risk. Markets are incomplete and agents can only self-insure by trading a riskless bond or by working longer hours. We assume that households have constant absolute risk (CARA) utility and idiosyncratic shocks are normally distributed. As in Acharya and Dogra (2018), these assumptions imply that the economy permits linear aggregation which in turn means that the infinite dimensional distributions of consumption, hours worked and wealth can be summarized by their cross-sectional averages: the *positive* behavior of macroeconomic aggregates can be described independently of the distribution of the wealth distribution. Of course, from a *normative* perspective, the dispersion of wealth does affect social welfare and hence the optimal conduct of monetary policy. Crucially, the effect of inequality on social welfare is also summarized by a finite dimensional sufficient statistic. This makes the optimal policy problem analytically tractable, allowing us to drill down and identify exactly the features that makes the optimal conduct of monetary policy different in HANK economies compared to the RANK benchmark.

As is commonly known, in RANK the planner seeks to stabilize prices and keep output at its productively efficient level. In HANK, the planner has an additional objective - to use monetary policy to reduce the cross-sectional consumption dispersion that results from the cumulated effects of uninsured idiosyncratic shocks. This incentive is shut down in RANK by construction. We show that there are three ways in which the planner can use monetary policy to achieve this purpose. First, the central bank may attempt to *reduce the amount of income risk* that households are exposed to (the *income risk channel*). How to achieve this reduction naturally depends on the cyclicality of income risk: if this risk is countercyclical, then the central bank has an incentive to raise output in order to lower risk, while the opposite is true if risk is procyclical. Either way, the central bank's willingness and ability to manipulate the amount of idiosyncratic risk that households face gives it an incentive to move output away from the level consistent with stable prices and productive efficiency.[1]

The second way in which the central bank may reduce consumption dispersion, independently of affecting the level of income risk, is by reducing the *pass-through from income risk to consumption risk*, that is, by lowering the marginal propensity to consume out of a change in individual income (the *self-insurance channel*). This pass-through ultimately reflects the ability of the households to self-insure against idiosyncratic risk through borrowing or working longer hours, and is thus affected by the paths of real interest rates and wages going forward. On the one hand, lower interest rates make it easier for households to borrow in response to an unfavorable shock, making individual consumption less response to changes in individual income; this ultimately reduces consumption dispersion at any level of income risk. On the other hand, higher wages going forward make it easier for households to buffer the impact of a fall in current income on current consumption by borrowing today and working longer hours in the future to repay the debt. When future wages are high, only a small increase in hours (and hence the incurred dis-utility) worked is required to repay this debt. This again makes individual consumption less responsive to changes in current income. It follows that the central bank has an incentive to commit to low interest rates and high wages –i.e., to be *expansionary*– going forward in order to reduce the pass-through from income risk to consumption risk. This incentive is of course absent in RANK.

Finally, the central bank can reduce consumption dispersion through unanticipated changes in the marginal propensity to consume out of wealth. Given a distribution of wealth, an unexpected fall in interest rates benefits poor debtors, reducing their interest payments and increasing their consumption (the *unhedged interest rate exposure (URE) channel*) (Auclert, 2019). Conversely, lower interest rates reduce the interest income of rich savers, reducing their consumption. Overall, lower rates reduce the marginal propensity to consume out of wealth, reducing consumption inequality. Importantly, this channel (unlike the previous two) only operates for unexpected changes in interest rates, as we explain in more detail in Section 4.

How does the presence of these three channels, through which the planner can affect inequality, change the optimal conduct of monetary policy? As is commonly known, in RANK, optimal monetary policy features *divine coincidence* in response to productivity shocks (Blanchard and Galí, 2007): it is both feasible and optimal to stabilize both the gap between output and its efficient level (output gap), and inflation. In our RANK economy, in the empirically relevant case where income risk is countercyclical, while it remains feasible to stabilize the output gap and inflation, it is no longer optimal to do so. But to understand the tradeoffs that lead the planner to deviate from divine coincidence in this case, it is instructive to start by examining a HANK economy where divine coincidence is optimal. This is the case when risk is mildly procyclical and there is no initial wealth inequality. In this knife-edge case, the two channels through which anticipated monetary policy affects consumption inequality exactly offset each other: expansionary policy raises output and hence income risk, but makes it easier for households to self-insure, leaving the consumption risk faced by households unchanged. In addition, the absence of wealth inequality at date 0 mutes the URE channel. Thus, while the planner would like to reduce consumption inequality, since it is not possible to do this with monetary policy, it remains optimal to stabilize both the

[1] Earlier work has stressed that the cyclicality of income risk affects the response of aggregate demand in HANK economies (Werning, 2015; Acharya and Dogra, 2018; Bilbiie, 2019a), and this in itself generically requires a benevolent central bank to implement a different path of the policy rate than in RANK (Challe, 2020). This does not, however, necessarily warrant departing from price stability.

output gap and inflation.

Away from this knife-edge case, monetary policy can, and does exploit the channels above to affect consumption inequality. Suppose first that risk remains mildly procyclical so that income risk and self-insurance channels offset each other, but there is wealth inequality at date 0. In this case, monetary policy cannot affect inequality through the first two channels, but it can through the URE channel. Consequently, the optimal plan features an unexpected cut in interest rates at date 0, which reduces consumption inequality at the cost of deviating from the efficient level of output and inflation. That is, optimal policy features higher output and inflation at date 0 than would be optimal in RANK. While monetary policy cannot affect inequality from date 1 onwards, output and inflation continue to deviate from RANK as policy seeks to smooth the transition back to steady state. If policy was somehow prevented from creating the boom to exploit the URE channel at date 0, it would not seek to deviate from RANK at any other date either.

This brings us to the empirically relevant case of countercyclical risk. In this case expansionary monetary policy can reduce inequality at all dates through the self-insurance and income risk channels: a lower path of interest rates makes it easier for households to insure themselves, and makes a boom in output which reduces the level of income risk households face. In addition, at date 0, a cut in rates delivers a further reduction in inequality through the URE channel. Consequently, the planner always trades off this benefit of more expansionary monetary policy - namely, that it reduces consumption inequality - against the costs of inefficiently high output and inflation. These tradeoffs change the optimal response to productivity shocks, relative to RANK. Following a negative productivity shock, the planner lets output decline as much as output in the flexible price case and implements zero inflation by raising nominal interest rates. However, in HANK, optimal policy implements a lower path of nominal interest rates, curtailing the fall in output in order to mitigate the increase in inequality. Even though this entails higher inflation and output above its efficient level, this is optimal because inequality is already higher in recessions and so is the benefit from a reduction in inequality.

So far we have only discussed one channel through which unexpected cuts in interest rates can lower inequality (the URE channel). Several authors have highlighted another way in which unexpected changes in monetary policy can lower inequality, namely the *Fisher channel*: unexpected inflation redistributes from savers who hold nominal assets to debtors who hold nominal liabilities. We deliberately abstract from this channel in our baseline model (by letting households trade inflation-indexed debt) in order to emphasize that the URE channel does not depend on the ability of monetary policy to redistribute wealth through an inflation surprise. While conceptually distinct from the URE channel, the Fisher channel provides another avenue through which expansionary monetary policy can reduce inequality. In Section 6 we show that when households trade nominally denominated assets, optimal monetary policy is more expansionary in recessions compared to RANK.

Related Literature The paper most closely related to ours is Bhandari et al. (2018), who also study optimal monetary policy in a HANK model. The main difference between our paper and theirs is methodological. Bhandari et al. (2018) propose a numerical algorithm to derive optimal monetary policy in HANK models, while we study a HANK economy with constant constant absolute risk aversion (CARA) preferences and normally distributed shocks which permits closed form solutions.[2] We see the two approaches as inherently complementary: the first allows more flexibility in the structure of preferences and idiosyncratic shocks, while the second better isolates the channels by which the central bank manipulates consumption dispersion along the optimal policy plan. Another closely related paper is Nuño and Thomas (2019) who study how URE and Fisher effects affect the optimal conduct of monetary policy in the presence of heterogeneity. Unlike us, they study a small open economy in which short-term real interest rates and output are unaffected by monetary policy. Thus, the classic output inflation trade-off which is central to New Keynesian economies is absent in their setting.

Several authors have studied optimal monetary policy in simple HANK economies with limited cross-sectional heterogeneity–see, e.g., Bilbiie (2008); Bilbiie and Ragot (2018); Bilbiie (2019a) and Challe (2020).[3] Most of these papers achieve tractability by imposing the *zero liquidity limit* (households cannot borrow and government debt is in zero net supply).[4] This assumption rules out the self-insurance channel because in equilibrium households do not borrow or lend and hence they spend all their income on consumption. Our analysis shows that this assumption rules out an important channel through which monetary policy affects inequality.

More generally, our paper belongs to the growing strand of literature that revisits the transmission and optimality of various economic policies within the HANK framework. This includes not only the work on conventional monetary policy discussed above but also that on unconventional monetary policy (McKay et al., 2016; Acharya and Dogra, 2018; Bilbiie, 2019a; Cui and Sterk, 2019), on unemployment-insurance and social-insurance policies (McKay and Reis, 2016, 2019; den Haan et al., 2018; Kekre, 2019), and on fiscal policy (Auclert et al., 2018; Bilbiie, 2019b).

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 characterizes equilibria, defining the implementability constraints the planner faces. Section 4 shows that the utilitarian planner's objective function can be written in terms of aggregate variables and a single sufficient statistic for the welfare-relevant measure of inequality. Section 5 characterizes optimal monetary policy. Section 6 introduces nominal bonds and shows how our results extend to that case. Section 7 concludes.

2 Environment

2.1 Households

We study a Bewley-Huggett economy in which households face uninsurable idiosyncratic shocks to their disutility from supplying labor. We abstract from aggregate risk but allow for a one time unanticipated shock at date 0, after which agents have perfect foresight. Our economy features a perpetual youth structure à la Blanchard-Yaari in which each individual faces a constant survival probability ϑ in any period. Population is fixed and normalized to 1. Consequently, the size of a newly born cohort at any date t is $1-\vartheta$ and the

[2] Caballero (1990), Calvet (2001), Wang (2003), Angeletos and Calvet (2006) have used similar modeling assumptions in real economies. Recently, Acharya and Dogra (2018) shows that these assumptions are very helpful in understanding the positive properties of HANK economies.

[3] See also Nisticò (2016), who generalizes the Two-Agent New Keynesian (TANK) model of Galí et al. (2007) and Bilbiie (2008) to the case of stochastic asset-market participation, and Debortoli and Galí (2018) on the comparison between the TANK model and a HANK model with homogeneous borrowing-constrained households and heterogeneous unconstrained households.

[4] Bilbiie and Ragot (2018) is an exception as it it allows agents to hold money in positive amounts for self-insurance purposes.

date t size of a cohort born at $s<t$ is $(1-\vartheta)\vartheta^{t-s}$. The date s problem of an individual i born at date s is:

$$\max_{\{c_t^s(i),\ell_t^s(i),a_t^s(i)\}} \sum_{t=s}^{\infty} (\beta\vartheta)^{t-s}\, u\Big(c_t^s(i), \ell_t^s(i); \xi_t^s(i)\Big)$$

s.t.

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

$$a_s^s(i) = 0 \tag{2}$$

Agents have CARA preferences over both consumption and (disutility of) labor:

$$u\Big(c_t^s(i), \ell_t^s(i); \xi_t^s(i)\Big) = -\frac{1}{\gamma} e^{-\gamma c_t^s(i)} - \rho e^{\frac{1}{\rho}[\ell_t^s(i)-\xi_t^s(i)]} \tag{3}$$

Each agent i saves in riskless real actuarial bonds, issued by financial intermediaries (described below), which trade at a price of q_t at date t and pay off one unit of the consumption good at $t+1$ if the agent survives.[5] Each agent can take unrestricted positive or negative positions in the bond and these choices are only disciplined by the transversality condition. T_t denotes lump-sum transfers net of taxes and dividends from the firms. For simplicity, we assume that dividends are equally distributed across agents.

A household supplies labor $\ell_t^s(i)$ at the pre-tax real wage w_t. The agent faces uninsurable shocks $\xi_t^s(i) \sim N\left(\overline{\xi}, \sigma_t^2\right)$ to the dis-utility of supplying labor. $\xi_t^s(i)$ is independent across time and across individuals. A larger realization of $\xi_t^s(i)$ reduces the dis-utility from work and, given wages, increases the household's labor supply. Equivalently, one may think of $\xi_t^s(i)$ as a shock to the household's endowment of time available to supply labor.[6] To see this, define leisure as $l_t^s(i) = \xi_t^s(i) - \ell_t^s(i)$. Then one can rewrite the period utility functional (3) as $-e^{-\gamma c_t^s(i)}/\gamma - \rho e^{-l_t^s(i)/\rho}$ and the budget constraint as:

$$c_t^s(i) + w_t l_t^s(i) + q_t a_{t+1}^s(i) = w_t \xi_t^s(i) + a_t^s(i) + T_t \tag{4}$$

The LHS of (4) denotes the purchases of consumption, leisure and bonds by the household while the RHS denotes the *notional cash-on-hand* - the value of the household's time endowment along with savings net of transfers. Henceforth, we will simply refer to this as *cash-on-hand*. We allow for the possibility that the variance of ξ, σ_t^2, vary endogenously with the level of economic activity as we discuss below.

2.2 Financial intermediaries

There is a competitive financial intermediation sector which trades actuarial bonds with households and trades government debt. An intermediary only needs to repay households that survive between t and $t+1$.

[5] In economies with a distribution of *nominal* debt, unexpected inflation redistributes wealth between creditors and debtors. Bhandari et al. (2018) and Nuño and Thomas (2019) discuss how optimal monetary policy takes this into account. Our benchmark economy deliberately abstracts from this channel in order to highlight the other ways in which optimal monetary policy differs in HANK and RANK economies. In Section 6, we replace real debt with nominal debt, bring this channel back into play and show how our results change.

[6] We thank Gianluca Violante for suggesting this interpretation.

The representative intermediary solves:

$$\max_{a_{t+1},B_{t+1}} -\vartheta a_{t+1} + B_{t+1} \text{ s.t. } -q_t a_{t+1} + \Pi_{t+1}\frac{B_{t+1}}{1+i_t} \leq 0$$

where B_t denotes government debt, a_t denotes claims held by households, $R_t = \frac{1+i_t}{\Pi_{t+1}}$ is the gross real return on government debt, i_t is the nominal interest rate which is set by the monetary authority and Π_{t+1} denotes gross inflation between t and $t+1$. Zero profits require that the intermediary sells/buy bonds from the households at a price $q_t = \frac{\vartheta}{R_t}$ and that $\vartheta a_{t+1} = B_{t+1}$.

2.3 Final goods producers

A representative competitive final goods firm transforms the differentiated intermediate goods y_t^j, $j \in [0,1]$ into the final good y according to the CES aggregator $y_t = \left[\int_0^1 y_t(j)^{\frac{1}{\lambda}} dj\right]^{\lambda}$. As is standard, the final good producer's demand for variety j is:

$$y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\frac{\lambda}{\lambda-1}} y_t \tag{5}$$

where $\frac{\lambda}{\lambda-1}$ is the elasticity of substitution between varieties.

2.4 Intermediate goods producers

There is a continuum of monopolistically competitive intermediate goods firms indexed by $j \in [0,1]$. Each firm faces a quadratic cost of changing the price of the variety it produces (Rotemberg, 1982). If firm j hires $n_t(j)$ units of labor, it can only sell to the final goods firm the quantity:

$$y_t(j) = z_t n_t(j) - \frac{\Psi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 y_t \tag{6}$$

where z_t denotes the level of aggregate productivity at date t. Firm j solves:

$$\max_{\{P_t^j, n_t^j, y_t^j\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} Q_{t|0}\left\{\frac{P_t(j)}{P_t} y_t(j) - (1-\tau) w_t n_t(j)\right\}$$

subject to (5) and (6) where $Q_{t|0} = \prod_{s=0}^{t} R_s^{-1}$. This yields the standard Phillips curve:

$$(\Pi_t - 1)\,\Pi_t = \frac{\lambda}{\Phi(\lambda-1)}\left[1 - \frac{z_t}{(1-\tau)\,\lambda w_t}\right] + \frac{1}{R_t}\left(\frac{y_{t+1} z_t w_t}{y_t z_{t+1} w_{t+1}}\right)(\Pi_{t+1} - 1)\,\Pi_{t+1} \tag{7}$$

2.5 Government

The monetary authority sets the interest rate on nominal government debt. The fiscal authority subsidizes the wage bill of firms at a rate τ and rebates lumpsum taxes/transfers to all households equally. The government budget constraint is given by:

$$\frac{B_{t+1}}{1+i_t} = P_t T_t + \tau P_t w_t \int_0^1 n_t(j) dj + B_t \tag{8}$$

We further assume that government debt is in zero net supply, $B_t = 0$ for all $t \geq 0$.

2.6 Market clearing

In equilibrium, the markets for the final good, labor and assets must clear:

$$\begin{aligned} y_t &= (1-\vartheta)\sum_{s=-\infty}^{t}\vartheta^{s-t}\int_i c_t^s(i)di \\ \int_0^1 n_t(j)dj &= (1-\vartheta)\sum_{s=-\infty}^{t}\vartheta^{s-t}\int_i \ell_t^s(i)di \\ 0 = a_t &= (1-\vartheta)\sum_{s=-\infty}^{t}\vartheta^{s-t}\int_i a_{t+1}^s(i)di \end{aligned}$$

where the last equation holds because $B_t = 0$ for all t.

2.7 Shocks

As mentioned previously, we abstract from aggregate risk but allow for a one time unanticipated aggregate shock to the level of labor productivity z_0 at date 0. We assume that the shock decays geometrically: $\ln z_t = \varrho_z^t \ln z_0$ for $\varrho_z \in [0, 1)$.

3 Characterizing equilibria

As in Acharya and Dogra (2018), CARA preferences and normally distributed shocks imply that the model aggregates linearly and the distribution of wealth does not directly affect the dynamics of aggregate variables. We begin by describing optimal household decisions in equilibrium.

Proposition 1 (Household Decision Rules)**.** *In equilibrium, the optimal date t consumption and labor supply decisions of a household i born at date s are:*

$$c_t^s(i) = \mathcal{C}_t + \mu_t x_t^s(i) \tag{9}$$

$$\ell_t^s(i) = \rho \ln w_t - \gamma\rho c_t^s(i) + \xi_t^s(i) \tag{10}$$

where $x_t^s(i) = a_t^s(i) + w_t\left(\xi_t^s(i) - \overline{\xi}\right)$ is demeaned cash-on-hand*, $\mathcal{C}_t$ denotes aggregate consumption and μ_t is the "marginal propensity to consume" (MPC) out of cash-on-hand. These evolve according to:*

$$\mathcal{C}_t = -\frac{1}{\gamma}\ln\beta R_t + \mathcal{C}_{t+1} - \frac{\gamma\mu_{t+1}^2 w_{t+1}^2\sigma_{t+1}^2}{2} \tag{11}$$

$$\mu_t^{-1} = 1 + \gamma\rho w_t + \frac{\vartheta}{R_t}\mu_{t+1}^{-1} \tag{12}$$

Proof. See Appendix A ∎

To understand the role of market incompleteness in explaining the behavior of consumption and labor supply it is useful to compare equations (9) and (10) to their counterparts under complete markets.

Under complete markets, all households are insured against dis-utility shocks, i.e. the marginal utility of consumption $e^{-\gamma c_t^s(i)}$ and the marginal dis-utility of work $e^{\frac{1}{\rho}(\ell_t^s(i)-\xi_t^s(i))}$ are equalized across all states and so: $\frac{\partial c_t^s(i)}{\partial \xi_t^s(i)} = 0$ and $\frac{\partial \ell_t^s(i)}{\partial \xi_t^s(i)} = 1$. Thus, since households' consumption is insured, a household which draws a temporarily higher dis-utility from working can reduce hours without experiencing a drop in consumption. Instead, when markets are incomplete (9) and (10) imply that:

$$\frac{\partial c_t^s(i)}{\partial \xi_t^s(i)} = \mu_t w_t > 0 \qquad \text{and} \qquad \frac{\partial \ell_t^s(i)}{\partial \xi_t^s(i)} = 1 - \gamma\rho\mu_t w_t < 1$$

For example, after a negative ξ shock (i.e., greater dis-utility from working), consumption declines instead of remaining constant while labor supply does not fall quite as much as under complete markets. While households use credit and labor market to insure themselves to some extent, these are not perfect substitutes for Arrow securities, so agents are only able to *partially* insulate themselves from the shock. When the dis-utility of labor rises, households would like to work less, but reducing hours as much as under complete markets would cause consumption to drop too much. The optimal response to the shock is to use labor supply for self-insurance, i.e. to work longer hours than under complete markets.[7]

Proposition 1 also states that the MPC out of cash-on-hand is the same across individuals; (12) describes its evolution over time. Intuitively, consider a household i that receives an additional dollar at date t. They will optimally choose to spend $dc_t^s(i) = \mu_t$ of the dollar in the current period. Since consumption and leisure are normal goods, they also reduce hours worked by $\gamma\rho\mu_t$, resulting in $\gamma\rho w_t \mu_t$ less income. Saving the remaining $1 - \mu_t(1 + \gamma\rho w_t)$, they find themselves with $da_{t+1}^s(i) = \frac{R_t}{\vartheta}\left[1 - \mu_t(1 + \gamma\rho w_t)\right]$ next period out of which they will consume $dc_{t+1}^s(i) = \mu_{t+1} da_{t+1}^s(i)$. Finally, it is optimal to smooth consumption so that $dc_t^s(i) = dc_{t+1}^s(i)$ which yields $\mu_t = \mu_{t+1}\frac{R_t}{\vartheta}\left[1 - \mu_t(1 + \gamma\rho w_t)\right]$. Rearranging this expression yields equation (12). Iterating forwards yields:

$$\mu_t = \frac{1}{\sum_{s=0}^{\infty} Q_{t+s|t}(1 + \gamma\rho w_{t+s})}$$

The MPC μ_t, which measures the pass-through from a fall in cash-on-hand to consumption, is increasing in current and future interest rates and decreasing in current and future wages. This is because interest rates and wages affect households' ability to use the bond and labor markets, respectively, for self-insurance. Consider a household who receives an unfavorable shock $\xi < 0$. The household responds by working less today, borrowing in order to mitigate the decline in consumption, and working longer hours in the future. A lower path of interest rates reduces the cost of borrowing, making it easy to self insure using the bond market and lowering the responsiveness of consumption to changes in cash-on-hand. Similarly, higher future wages reduce the (disutility) cost of working more hours in the future since even a small increase in hours worked suffices to repay the same debt. This too lowers the sensitivity μ_t of consumption to cash-on-hand.[8]

While the sensitivity of household consumption to shocks (μ_t) depends on the factors we have just described, the average level of consumption in the economy $\mathcal{C}_t$ depends on interest rates relative to impa-

[7] As already mentioned, the model can be re-interpreted as one with an idiosyncratic time-endowment shock and utility from leisure time. In this interpretation, both consumption and leisure time stay constant under complete markets, while both co-vary with the idiosyncratic shock under incomplete markets.

[8] While Acharya and Dogra (2018) already discuss how the MPC responds to future real interest rates, the path of wages has no effect on the MPC in their paper because their environment features inelastic labor supply. In this model, however, since households can choose how much labor to supply, they use this additional margin to self-insure.

tience and households' desire for precautionary savings, as shown in equation (11). Absent idiosyncratic risk, $\sigma_t = 0$, (11) is a standard Euler Equation; higher real interest rates relative to household impatience raise consumption growth. The last term in (11) reflects precautionary savings. Given (9), the conditional variance of date $t+1$ consumption of household i is $\mathbb{V}_t\left(c^s_{t+1}(i)\right) = \mu^2_{t+1} w^2_{t+1} \sigma^2_{t+1}$. To the extent that consumption risk is positive and households are prudent ($\gamma > 0$), households save more than in a riskless economy for the same interest rate, i.e. they choose a steeper path of consumption growth. The variance of consumption, in turn, depends on both the variance of cash-on-hand $\mathbb{V}_t\left(x^s_{t+1}(i)\right) = w^2_{t+1}\sigma^2_{t+1}$, and the pass-through of cash-on-hand risk into consumption risk measured by the (squared) MPC μ^2_{t+1}.

Determination of y_t In a symmetric equilibrium, aggregating (6) across firms, we have:

$$y_t = z_t n_t - \frac{\Psi}{2}\left(\Pi_t - 1\right)^2 y_t \tag{13}$$

Aggregating labor supply (10) across currently alive households and using goods and labor market clearing:

$$n_t = \rho \ln w_t - \gamma \rho y_t + \bar{\xi} \tag{14}$$

Combining the two equations above, we have:

$$y_t = z_t \frac{\rho \ln w_t + \bar{\xi}}{1 + \gamma \rho z_t + \frac{\Psi}{2}\left(\Pi_t - 1\right)^2} \tag{15}$$

where $\frac{\Psi}{2}\left(\Pi_t - 1\right)^2$ denotes the resource cost of inflation - higher inflation reduces output.

Deriving the aggregate IS equation Imposing goods market clearing in (11) yields the aggregate IS equation which describes the relation between output today and tomorrow:

$$y_t = y_{t+1} - \frac{1}{\gamma} \ln \beta \left(\frac{i + i_t}{\Pi_{t+1}}\right) - \frac{\gamma \mu^2_{t+1} w^2_{t+1} \sigma^2_{t+1}}{2} \tag{16}$$

Time varying σ_t Following McKay and Reis (2019), we allow for the variance of ξ shocks to vary endogenously with economic activity so that the model generates cyclical changes in the distribution of earnings risks. In particular, we assume that $\sigma^2_t w^2_t = \sigma^2 w^2 e^{2\phi(y_t - y)}$ where y denotes steady state output and $\phi = \frac{\partial \ln \mathbb{V}(x)}{\partial y}$ is the constant semi-elasticity of the variance of cash-on-hand w.r.t output. This flexible specification allows the variance of cash-on-hand $\mathbb{V}_t(x)$ to be either increasing in y_t (*procyclical* risk), when $\phi > 0$; decreasing in y_t (*countercyclical* risk), when $\phi < 0$; or independent of the level of y_t (*acyclical* risk) when $\phi = 0$.

3.1 Steady state

We now characterize allocations in the zero inflation steady state. We normalize the level of productivity $z = 1$ in steady state. Imposing $\Pi_t = \Pi_{t+1} = 1$ in (7) requires that steady state wages $w = \frac{1}{\lambda(1-\tau)}$. Given this wage, steady state output is $y = \frac{\rho \ln w + \xi}{1+\gamma\rho}$. Imposing steady state in (16) and (12) yields:

$$R = \beta^{-1} e^{-\frac{\Lambda}{2}} \qquad \text{and} \qquad \mu = \frac{1-\widetilde{\beta}}{1+\gamma\rho w} \tag{17}$$

where $\Lambda = \gamma^2\mu^2 w^2 \sigma^2$ denotes the consumption risk faced by households in steady state (scaled by the coefficient of prudence) and $\widetilde{\beta} = \frac{\vartheta}{R}$ is the steady state price of an actuarial bond. Observe that the presence of uninsurable risk ($\Lambda > 0$) implies that the equilibrium real interest rate $R < \beta^{-1}$. Furthermore, the steady state distribution of cash-on-hand x in the population is given by:

$$F(x) = (1-\vartheta) \sum_{s=0}^{\infty} \vartheta^s \Phi\left(\frac{x}{w\sigma\sqrt{s+1}}\right) \tag{18}$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution. This follows since, conditional on survival, x is a random walk with no drift and a variance of $w^2\sigma^2$ in steady state.

3.2 Linearized demand block

The demand block of the economy, given a path of interest rates, can be described by the IS equation (16), the MPC recursion (12) and the definition of GDP (15). Before analyzing optimal policy, it is useful to compare the dynamics of this HANK economy to its RANK counterpart. It is easiest to compare the first-order Taylor expansion of the equations describing aggregate dynamics in the neighborhood of the zero inflation steady state, which are:[9]

$$\widehat{y}_t = \Theta \widehat{y}_{t+1} - \frac{1}{\gamma}(i_t - \pi_{t+1}) - \frac{1}{\gamma}\Lambda\widehat{\mu}_{t+1} \tag{19}$$

$$\widehat{\mu}_t = -(1-\widetilde{\beta})\frac{\gamma\rho w}{1+\gamma\rho w}\frac{\widehat{w}_t}{w} + \widetilde{\beta}\left(\widehat{\mu}_{t+1} + i_t - \pi_{t+1}\right) \tag{20}$$

$$\widehat{y}_t = \frac{\rho}{1+\gamma\rho}\frac{\widehat{w}_t}{w} + \frac{y}{1+\gamma\rho}\widehat{z}_t \tag{21}$$

where $\Theta = 1 - \frac{\Lambda\phi}{\gamma}$. In RANK, there is no idiosyncratic risk, i.e. $\sigma^2 = 0$ which implies $\Theta = 1$ and $\Lambda = 0$, so that (19) becomes the standard RANK IS curve. As discussed in Acharya and Dogra (2018), uninsurable idiosyncratic risk changes the IS equation in two ways. First, it can introduce either *discounting* ($\Theta < 1$) or *compounding* ($\Theta > 1$) depending on the cyclicality of risk. If risk is acyclical, $\phi = 0$, then $\Theta = 1$ as in the RANK IS equation. If risk is procyclical $\phi > 0$ then $\Theta < 1$ and there is *discounting* in the IS equation. This is because in this situation low future output implies low idiosyncratic income risk, hence (all else equal) a fall in precautionary savings that mutes down the effect of the forthcoming recession on current aggregate demand. The opposite occurs when risk is countercyclical ($\phi < 0$), in which case $\Theta > 1$ –that is, we have *compounding* in the IS equation–, because then the impact of a future recession on current aggregate demand is magnified by the rise in precautionary savings. Second, the strength of the precautionary motive depends not only on the level of income risk but also on its pass-through to consumption risk, i.e. $\widehat{\mu}_{t+1}$, which in turn depends of the future path of real interest rates and wages. By committing to a lower path of real rates or a higher path of wages, monetary policy can lower the strength

[9]We linearize y_t and w_t in levels while all other variables are log-linearized.

of the precautionary savings motive at any level of income risk.

Equations (7), (12), (15) and (16) summarize the optimality conditions of the private sector and define the implementability constraints faced by the central bank. We may now turn to its objective function.

4 Objective function of the planner

We assume that the planner maximizes a utilitarian criterion; at any date t the planner assigns equal weights to the welfare of all households currently alive and a weight of β^{s-t} on the welfare of cohorts who will be born at dates $s > t$. Given these Pareto weights, the planner's objective can be written as maximizing $\sum_{t=0}^{\infty} \beta^t \mathbb{U}_t$ where $\mathbb{U}_t$, the period t felicity function of the planner is simply the average utility of all surviving agents:[10]

$$\mathbb{U}_t = -(1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{-s} \int u\Big(c_t^s(i), \ell_t^s(i); \xi_t^s(i)\Big) di,$$

This expression for $\mathbb{U}_t$ feature the full cross-sectional distribution of agents and is in general not tractable. Fortunately, it can be greatly simplified by exploiting our CARA-Normal structure again.

Proposition 2 (Social Welfare Function)**.** *The period felicity function $\mathbb{U}_t$ can be written as*

$$\mathbb{U}_t = u(y_t, n_t; \bar{\xi})\Sigma_t$$

where

$$\Sigma_t = (1-\vartheta) \sum_{s=-\infty}^{0} \vartheta^{-s} e^{\frac{1}{2}\gamma^2 \sigma_c^2(t,s)}$$

and $\sigma_c^2(t,s)$ is the date-t cross-sectional dispersion of consumption amongst the surviving households from the cohort born at date $s \leq t$, i.e., $c_t^s(i) \sim N\left(y_t, \sigma_c^2(t,s)\right)$.

Proof. See Appendix B.2. □

Intuitively, $u(y_t, n_t; \bar{\xi})$ is the notional flow utility of the representative agent, i.e., the period utility functional (3) evaluated at aggregate consumption y_t, aggregate labor supply n_t, and mean labor disutility $\bar{\xi}$. Σ_t can be thought of as the welfare cost of inequality, and is increasing in the within cohort dispersion of consumption. In a riskless economy, there would be no consumption dispersion and hence $\Sigma_t = 1$ at all dates. However, in the presence of risk, $\Sigma_t > 1$, reducing welfare relative to this representative agent benchmark. Recall that $u(\cdot) < 0$ and so higher Σ_t reduces welfare.

[10] Note that the planner discounts felicity at the same rate as the households would themselves. Consider a change in allocations which reduces the date t felicity of cohort s by du_t and increases their date $t+1$ felicity by du_{t+1}, while keeping the felicity at all other dates and for all other agents the same. A cohort s individual will be indifferent regarding this change if $du_t = \beta\vartheta du_{t+1}$. From the planner's perspective this changes aggregate welfare by $-\vartheta^{s-t} du_t + \beta\vartheta^{s+1-t} du_{t+1}$. Thus, the planner will be indifferent about this change if and only if the individuals themselves are indifferent. As discussed by Calvo and Obstfeld (1988), assuming that the planner and the households share the same rate of time preference ensures that social preferences are time-consistent, so that the first-best intertemporal allocation of consumption across cohorts does not change over time. This does not prevent other from of time inconsistencies from arising in decentralized equilibrium (as shown below), but these are unrelated to the form of social preferences.

Appendix B.2.1 shows that the evolution of Σ_t for $t > 0$ can be written as:

$$\ln \Sigma_t = \frac{1}{2}\gamma^2 \mu_t^2 w_t^2 \sigma_t^2 + \ln\left[1 - \vartheta + \vartheta \Sigma_{t-1}\right] \tag{22}$$

with

$$\ln \Sigma_0 = \frac{1}{2}\gamma^2 \mu_0^2 w_0^2 \sigma_0^2 + \ln\left[1 - \vartheta + \vartheta \Sigma_{-1}\right] + \underbrace{\ln\left(\frac{1 - \vartheta e^{\frac{\Lambda}{2}}}{1 - \vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mathbb{E}_{-1}\mu_0}\right)^2}}\right)}_{\text{effect of date 0 surprise / URE}} \tag{23}$$

where $\Sigma_{-1} = \Sigma = \frac{(1-\vartheta)e^{\frac{\Lambda}{2}}}{1-\vartheta e^{\frac{\Lambda}{2}}}$ is the steady state Σ.[11] Intuitively, higher cash-on-hand risk $w_t^2\sigma_t^2$ and a higher pass-through μ_t both tend to increase consumption inequality. In addition, consumption inequality inherits the slow moving dynamics of wealth inequality, as can be seen from the presence of Σ_{t-1} in (22).[12]

Equation (23) shows that the relation between μ_0 and Σ_0 is different than the relation between μ_t and Σ_t at all other dates. This can be explained intuitively as follows. At the beginning of date 0, the distribution of wealth a is at its steady state level: some households have positive net wealth and some are debtors. Since savers and debtors have different unhedged interest rate exposures (UREs) in the sense of Auclert (2019), an unanticipated change in interest rates affects consumption inequality. Suppose that at date 0, the central bank chooses a policy path that implements a transitory drop in the real interest rate. The lower interest rates benefit poor debtors, reducing their interest payments and allowing them to increase their consumption. By the same token, a lower path of rates reduces the interest income of rich savers, causing them to reduce consumption. In other words, lower rates reduce the MPC out of wealth ($\mu_0 \downarrow$) which reduces consumption inequality and hence Σ_0.

Importantly, an *anticipated* cut in rates would not reduce inequality as much as this unanticipated cut. If wealthy agents at date $t-1$ anticipated lower rates at date 0, they would save more in order to insure a higher level of consumption at date 0. Equally, the poor debtors would borrow more at date -1 knowing that their debt would be less costly to repay. For this reason, what reduces Σ_0 through this channel is not a fall in μ_0 per se but a fall in μ_0 relative to its expected value $\mathbb{E}_{-1}\mu_0$, as can be seen from the last term in (23). To be clear, anticipated cuts in rates *do* reduce inequality as discussed earlier: lower μ_t reduces Σ_t in equation (22). But there is an additional effect that comes from a surprise fall in rates. In our environment, since we do not have aggregate shocks (except for the unanticipated shock at date 0) and the fact that the Ramsey planner is only allowed to re-optimize at date 0 imply that this additional affect of an unanticipated change in μ can only occur at date 0. More generally, in an environment with aggregate shocks, surprise changes in μ would have this effect on any date, for example when there is an aggregate productivity shock and $z_t \neq \mathbb{E}_{t-1}z_t$.

Of course, this one-off redistribution would not operate in the absence of wealth inequalities at time 0. If the economy were starting with equal (zero) wealth for all households, instead of starting from the invariant distribution, then only the first effect would play out and the equilibrium value of consumption

[11] We are assuming that the economy is in steady state at date -1.

[12] Note that within-cohort consumption dispersion $\sigma_c^2(t,s)$ in general rises without bounds as the cohort ages (i.e., as $t - s \to \infty$), due to the cumulated effect of idiosyncratic shocks on the distribution of cash-on-hand. However, since every cohort gradually shrinks in size, while newborn cohorts have little consumption dispersion (i.e., $\sigma_c^2(t,t) = \mu_t^2 w_t^2 \sigma_t^2$), Σ_t does not necessarily blow up. In fact, provided that the survival rate $\vartheta < e^{-\Lambda/2}$, Σ_t is stationary.

dispersion at time zero would simply be:

$$\ln \Sigma_0 = \frac{1}{2}\gamma^2 \mu_0^2 w_0^2 \sigma_0^2. \tag{24}$$

5 Optimal monetary policy

The planner chooses sequences $\{w_t, \Pi_t, \mu_t, \Sigma_t, i_t, n_t\}_{t=0}^{\infty}$ to maximize $\sum_{t=0}^{\infty} \beta^t u\left(y_t, n_t; \overline{\xi}\right) \Sigma_t$ subject to the aggregate Euler equation (16), agggregate labor supply (14), the evolution of μ_t (12), the Phillips curve (7), the evolution of Σ_t (22)-(23) and the relationship between GDP and wages (15). In the RANK version of our economy, $\sigma = 0$ and (22) is replaced by $\Sigma_t = 1$ for all t. Appendix D presents the Lagrangian associated with this problem along with the first order necessary conditions for optimality.

5.1 Long run outcomes under the optimal Ramsey plan and the payroll subsidy

To proceed further, we need to take a stand on the value of the payroll subsidy τ. In RANK, if we imposed a production subsidy to eliminate the distortions caused by market power, zero inflation is optimal in the long run in the absence of aggregate shocks. This need not be true in our HANK economy where $\sigma > 0$ and so Σ_t is endogenous. As in the standard NK model, the Phillips curve (7) implies a long-run trade-off between inflation and economic activity:

$$(\Pi - 1)\Pi = \frac{\lambda}{(1 - \widetilde{\beta}\vartheta^{-1})(1-\lambda)\Psi}\left[1 - \frac{1}{(1-\tau)\lambda w}\right] \tag{25}$$

The presence of a long-run tradeoff implies that the policymaker can move wages (or equivalently output) above or below its flexible-price level by persistently deviating from price stability. While it is not optimal to do so in RANK, it may in fact be optimal in HANK because the level of economic activity affects both the amount of income risk households face $w_t^2\sigma_t^2$ and their ability to self-insure against this risk, μ_t. For example, if income risk is countercyclical, $\Theta > 1$ the planner may want to create inflation to raise wages (and output) above the level consistent with productive efficiency ($w > 1$), and thereby reduce income risk.

To make our results as comparable as possible to the classic NK literature on optimal monetary policy, we eliminate this motive for deviating from price stability by introducing an appropriately chosen payroll subsidy τ. To see how this works, consider the case with countercyclical risk in which the planner wants to implement a high after-tax wage $w > 1$ in the long run in order to reduce inequality. From (25) it is easy to see that if $\tau = \frac{\lambda-1}{\lambda}$ (the standard subsidy used in the RANK literature, such a high level of wages entails a marginal cost greater than 1 which implies positive long run inflation $\Pi > 1$. However, if the payroll subsidy is larger, the steady state marginal cost can be brought down to 1, consistent with zero inflation in the long run.

More generally, Appendix D.1 shows that in the presence of an appropriately chosen payroll subsidy to ensure zero inflation $\Pi = 1$, the planner's first-order condition for wages - which states that the net-benefit of higher economic activity must be zero at an optimum - becomes:[13]

$$\underbrace{\Omega}_{\substack{\text{benefit from}\\ \text{reduction of}\\ \text{inequality}}} \equiv \underbrace{\frac{\Lambda}{(1-\widetilde{\beta})(1-\Lambda)}}_{\substack{\text{reduction of inequality}\\ \text{due to low interest rates}}} + \underbrace{\frac{\Theta - 1}{(1-\widetilde{\beta})(1-\Lambda)}}_{\substack{\text{reduction of inequality}\\ \text{due to high output}}} = \underbrace{\frac{w-1}{1+\gamma\rho w}}_{\substack{\text{cost of deviating}\\ \text{from productive efficiency}}} \tag{26}$$

Equation (26) implies that the steady state wage and payroll subsidy consistent with the optimality of zero long-run inflation satisfy:

$$w = \frac{1+\Omega}{1-\gamma\rho\Omega} \qquad \text{and} \qquad \tau = \frac{\lambda-1}{\lambda} + \frac{1+\gamma\rho}{\lambda}\frac{\Omega}{\Omega+1} \tag{27}$$

Ω summarizes the benefit from a reduction in consumption inequality due to higher economic activity. The planner has the option to reduce interest rates and raise output above the flexible price level. Equation (27) states that at an optimum, the marginal benefit of lower inequality due to lower rates and higher output, Ω, must equal the marginal cost of distorting productive efficiency by raising output (and wages) above the flexible price level which is proportional to $w-1$, the negative of the *labor-wedge*.

Absent uninsurable risk ($\Lambda = 0, \Theta = 1$), there is no inequality and so there is no benefit from higher economic activity in terms of reducing inequality, $\Omega = 0$. Consequently, in RANK, optimal policy equates the cost of deviating from productive efficiency in steady state to 0 and hence $w = 1$ in the long run. In this case, $\Pi = 1$ in steady state can be implemented with the standard RANK subsidy $\tau = \frac{\lambda-1}{\lambda}$ (from eq. (27)) which removes the monopolistic distortion.

In the presence of risk ($\Lambda > 0$), optimal monetary policy seeks to reduce consumption inequality. This can be accomplished both by reducing labor income risk $w_t^2\sigma_t^2$ and by making it easier for households to self-insure against this risk (by reducing μ_t). Thus the policymaker may want to deviate from productive efficiency and price stability, both to facilitate self-insurance and to reduce income risk; equation (26) shows that Λ and $\Theta - 1$ represent the strength of these two motives respectively.

Consider first the *self-insurance channel*. When $\Theta = 1$, income risk is acyclical: the level of economic activity does not affect household income risk. In this case, deviating from price stability and productive efficiency (say by implementing a wage $w > z = 1$) delivers no benefits in terms of lower income inequality (second term on the LHS of (26) is zero). However, lowering real interest rates still makes it easier for households to smooth consumption by borrowing and reduces the pass-through from income shocks into consumption, measured by the first term of the LHS, reducing consumption inequality. Lower interest rates and the associated higher level of economic activity also have a cost, since they distort output and employment above the productively efficient level. Optimal policy equates these costs and benefits. So, even with acyclical risk, $\Omega > 0$, and output is optimally above productive efficiency in the long run. Since higher wages increase the firms marginal cost, it takes a higher payroll subsidy than in RANK $\tau > \frac{\lambda-1}{\lambda}$ to implement $\Pi = 1$ in steady state.

Next, consider the *income risk channel*. When $\Theta > 1$, income risk is countercyclical: higher economic activity lowers income risk. In this case, stimulating output above its productively efficient level lowers consumption inequality even for a fixed μ; in addition, the lower interest rates necessary to implement

[13] Note that here there is no term representing the cost of inflation, precisely because we assume that whatever the steady state level of wages, the appropriate payroll subsidy is chosen so that firms' marginal costs are consistent with zero inflation in the long-run, i.e., (25) holds with $\Pi = 1$.

higher output reduce μ further reducing consumption inequality. Thus, the benefit from higher output is even larger than if $\Theta = 1$ - both LHS components in (26) are positive and Ω is larger - and so output (and wages) must be even further above the productive efficient level than if $\Theta = 1$. Again, it takes a higher payroll subsidy $\tau > \frac{\lambda-1}{\lambda}$ to implement $\Pi = 1$ in steady state.

In contrast, when risk is procyclical ($\Theta < 1$), the effect of higher economic activity and lower rates on consumption inequality are ambiguous. While lower rates decrease the passthrough from income shocks to consumption ($\Lambda > 0$), they also raise output which now increases income risk ($\Theta - 1 < 0$). For sufficiently procyclical risk, the second effect dominates, $\Omega < 0$ and the optimal steady state level of output (and wages) is below the productively efficient level. In this case, it takes a lower payroll subsidy than $\frac{\lambda-1}{\lambda}$ to implement $\Pi = 1$ in steady state. For mildly procylical risk, the self-insurance channel dominates and $\Omega > 0$ with $w > 1$ in steady state. The self-insurance channel is perfectly balanced by the income risk channel if $1 - \Theta = \Lambda$ in which case $\Omega = 0$; higher economic activity and low interest rates have no first order effect on consumption inequality and in this case, the planner does not wish to distort productive efficiency in steady state, setting $w = 1$. The $\Omega = 0$ case will be a useful benchmark in what follows.

Remark 1 (Comparison with Zero-Liquidity Limits)**.** *The self-insurance channel is absent in models which feature incomplete markets but impose the zero liquidity limit such as Bilbiie (2019a,b); Ravn and Sterk (2017); Challe (2020) and others. Zero liquidity implies that in equilibrium, the passthrough from income risk to consumption risk is invariant to policy since it always equals 1. Instead, our analysis emphasizes that in HANK economies interest rate policy affects welfare not just via its effect on the level of economic activity, as in RANK, but also by affecting the ease with which households can self-insure.*

As is standard in the NK literature, a useful benchmark is the level of output under flexible prices. In a flexible-price version of the HANK economy, we would have $w_t = wz_t$ at all times, and output would be

$$y_t^n = z_t \frac{\rho\left(\ln w + \ln z_t\right) + \bar{\xi}}{1 + \gamma\rho z_t}$$

while the efficient level of output is:

$$y_t^e = z_t \frac{\rho \ln z_t + \bar{\xi}}{1 + \gamma\rho z_t}$$

In RANK, the flexible-price and efficient levels of output coincide: $y_t^e = y_t^n$. This is also true in HANK with $\Omega = 0$. But in general, when $\Omega \neq 0$, the flexible-price and efficient levels of output no longer coincide. With strongly procyclical risk $\Omega < 0$, the flexible-price level of output y_t^n is always below its efficient level y_t^e, while with mildly procyclical or countercyclical risk $\Omega > 0$, y_t^n is always above y_t^e. Using these definitions, we can express the linearized version of the Phillips curve (7) as:

$$\pi_t \quad = \quad \frac{\widetilde{\beta}}{\vartheta}\pi_{t+1} + \kappa\left(\widehat{y}_t - \widehat{y}_t^n\right) \tag{28}$$

where $\kappa = \left(\frac{1+\gamma\rho}{\rho\Psi}\right)\frac{\lambda}{\lambda-1}$, $\widehat{y}_t^e = \frac{\rho+1}{1+\gamma\rho}\widehat{z}_t$ and $\widehat{y}_t^n = \frac{\rho+y}{1+\gamma\rho}\widehat{z}_t$.

5.2 Calibration

While our results are primarily analytical, when plotting IRFs we parameterize the model as follows. We choose $\bar{\xi}$ to normalize aggregate steady state output y^* to 1 in the HANK economy with $\Omega = 0$

(equivalently, in the RANK economy). We calibrate the model to an annual frequency and choose the standard deviation of $\xi_t^s(i)$, σ, so that the standard deviation of income in steady state equals 0.5.[14] This is in line with Guvenen et al. (2014) who using administrative data find the standard deviation of 1 year log earnings growth rate to be slightly above 0.5. We set the slope of the Phillips curve $\kappa = 0.01$, and the elasticity of substitution between varieties $\frac{\lambda}{\lambda-1}$ to 10, implying a 10 percent steady state markup, $\lambda = 1.1$. We set the coefficient of relative prudence for the median household, $-\frac{cu'''(c)}{u''(c)} = \gamma$, to be 3, within the range of estimates in the literature (see e.g. Cagetti (2003); Fagereng et al. (2017); Christelis et al. (2015)). We set $r = 4\%$. ρ can be interpreted as the Frisch elasticity of labor supply for the median household; we set it equal to 1/3, within the range of estimates from the micro literature. We set the persistence of the shock $\varrho_z = 0.8$. Following Nisticò (2016), we set $\vartheta = 0.85$. Finally, we consider two values for the cyclicality of income risk: $\phi = \gamma$ (which implies $\Omega = 0$), and $\phi = -3$ (which implies $\Omega > 0$ and countercyclical risk). Finally, we set $\overline{\xi} = 1 + \gamma\rho$ to normalize the efficient level of output in steady state to 1.

5.3 Dynamics under optimal monetary policy under RANK

As is common in the NK literature, we characterize optimal policy by linearizing the first order conditions arising from the planner's Lagrangian (presented in Appendix E). It is useful to compare this characterization to optimal policy in a RANK version of our economy.

Lemma 1 (Optimal monetary policy in RANK). *In RANK ($\Lambda = 0, \Theta = 1$), output and inflation $\{\widehat{y}_t, \pi_t\}_{t=0}^{\infty}$ under optimal policy satisfy*

$$\eta_t = \widehat{y}_t - \widehat{y}_t^n \tag{29}$$

$$\eta_t = \eta_{t-1} - \frac{\lambda}{\lambda - 1}\pi_t \tag{30}$$

$$\pi_t = \frac{\widetilde{\beta}}{\vartheta}\pi_{t+1} + \kappa\left(\widehat{y}_t - \widehat{y}_t^n\right) \tag{31}$$

where η_t is the (normalized) multiplier on the Phillips curve (7) and is defined in Appendix E.1.

Proof. See Appendix E.1. □

Combining (29)-(30), we see that optimal policy in RANK satisfies the standard *target criterion*:[15]

$$\left(\widehat{y}_t - \widehat{y}_t^n\right) - \left(\widehat{y}_{t-1} - \widehat{y}_{t-1}^n\right) + \frac{\lambda}{\lambda - 1}\pi_t = 0 \tag{32}$$

where $\widehat{y}_{-1} = \widehat{y}_{-1}^e = 0$. Combining this with (31) and using the fact that $\widehat{y}_t^n = \widehat{y}_t^e$ in RANK, we see that the economy features a *divine coincidence*: it is both feasible and optimal for monetary policy to set $\widehat{y}_t = \widehat{y}_t^e$ and $\pi_t = 0$ at all dates and states. Given the appropriate steady state subsidy $\tau = \frac{\lambda-1}{\lambda}$, the flexible-price level of output, which is also consistent with zero inflation, maximizes social welfare –there is no tradeoff between implementing the efficient level of output and price stability.

[14] The standard deviation of income is given by $(1 - \gamma\rho\mu w)w\sigma$. We calibrate all parameters except the cyclicality of income risk to an economy with $\Omega = 0$, which implies $w = 1$.

[15] See for example, chapter 5 in Galí (2015).

5.4 Dynamics of monetary policy under HANK

In HANK, the planner has an additional objective relative to RANK: in addition to stabilizing inflation and keeping output close to its efficient level, she wants to keep inequality Σ_t as low as possible. The innovations to inequality depend on consumption risk $\mu_t^2\sigma_t^2$ which in turn depends on both income risk σ_t^2 and the pass-through μ_t^2. This can be seen from the linearized version of (22) which is given by:

$$\frac{\widehat{\Sigma}_t}{\Sigma} = \begin{cases} \underbrace{\gamma\left(1-\Theta\right)\widehat{y}_t + \Lambda\widehat{\mu}_t}_{\substack{\text{consumption}\\ \text{risk}}} + \frac{\vartheta}{\beta R}\frac{\widehat{\Sigma}_{t-1}}{\Sigma} & \text{for } t>0 \\ \\ \gamma\left(1-\Theta\right)\widehat{y}_0 + \Lambda\widehat{\mu}_0 + \frac{\vartheta\Lambda\Sigma}{1-\vartheta}\widehat{\mu}_0 & \text{for } t=0 \end{cases} \tag{33}$$

One way to reduce consumption inequality is to affect the level of output: with procyclical risk ($\Theta < 1$) lower output directly reduces income risk ($\Theta > 1$) faced by households while with countercyclical risk, a higher level of output is necessary to reduce income risk. An alternative path to lower consumption inequality is to commit to a lower path of interest rates which reduces pass-through from income to consumption risk.

However, the planner only has one instrument - the nominal interest rate. Lowering the nominal interest rate lowers the pass-through from income risk to consumption risk (measured by μ_t) but increases output. If risk is countercyclical $\Theta > 1$, then this too reduces consumption risk. However, if risk is procyclical $\Theta < 1$, then it increases income risk, leaving the overall effect on consumption risk unclear. To see what combinations of $\{\widehat{y}_t, \widehat{\mu}_t\}$ that the planner can implement with some path of nominal interest rates, combine the IS equation (19), μ recursion (20) using the definition of GDP (21) and solve forwards:

$$\gamma\left[1+\left(1-\widetilde{\beta}\right)\Omega\right]\widehat{y}_t + \widehat{\mu}_t = \left(1-\widetilde{\beta}\right)\left(1+\Omega\right)\frac{\gamma y}{y+\rho}\sum_{s=0}^{\infty}\widetilde{\beta}^s\left(1-\Lambda\right)^s\widehat{y}_{t+s}^n \equiv \Gamma_t \tag{34}$$

where Γ_t is an exogenous process driven by the sequence $\{y_t^n\}$, which in turn depends solely on $\{\widehat{z}_t\}$.

5.4.1 HANK with $\Omega = 0$

To understand the trade-offs facing the planner, it is useful to consider the special case in which $\Omega = 0$ (or equivalently $1-\Theta = \Lambda$). Recall from (26) that this is the case in which the zero inflation steady state features productive efficiency ($\Pi = 1, w = 1$). This benchmark features mildly procyclical risk: while this may not be the empirically relevant case, it is a useful benchmark because in this case, the constraint on the planner's ability to affect consumption inequality is particularly severe. Recall that when risk is procyclical, the effect of expansionary monetary policy on consumption inequality is generally ambiguous: higher output reduces the level of income risk, but lower interest rates reduces the passthrough from income to consumption risk. When $\Omega = 0$, both these effects exactly cancel each other out and consumption is invariant to monetary policy to first-order. To see this, note that (33) becomes:

$$\frac{\widehat{\Sigma}_t}{\Sigma} = \begin{cases} \Lambda(\gamma\widehat{y}_t + \widehat{\mu}_t) + \frac{\vartheta}{\beta R}\frac{\widehat{\Sigma}_{t-1}}{\Sigma} & \text{for } t>0 \\ \\ \Lambda(\gamma\widehat{y}_0 + \widehat{\mu}_0) + \frac{\vartheta\Lambda\Sigma}{1-\vartheta}\widehat{\mu}_0 & \text{for } t=0 \end{cases} \tag{35}$$

while (34) becomes:

$$\gamma \widehat{y}_t + \widehat{\mu}_t = \Gamma_t \tag{36}$$

Clearly, in this case, the planner cannot affect the evolution of consumption risk for dates $t > 0$ which is solely driven by exogenous shocks $\{y_t^n\}_{t\geq 0}$ denoted Γ_t in (36). Plugging in (36) into (35) shows that the evolution of inequality after date 0 is governed completely by the exogenous sequence $\{\Gamma_t\}$:

$$\frac{\widehat{\Sigma}_t}{\Sigma} = \begin{cases} \Lambda\Gamma_t + \frac{\vartheta}{\beta R}\frac{\widehat{\Sigma}_{t-1}}{\Sigma} & \text{for } t > 0 \\ \\ \Lambda\Gamma_0 + \frac{\vartheta\Lambda\Sigma}{1-\vartheta}\widehat{\mu}_0 & \text{for } t = 0 \end{cases} \tag{37}$$

Again, while a cut in interest rates lowers $\widehat{\mu}_t$, it increases output $\widehat{y}_t$ and hence income risk, leaving consumption risk unchanged. A higher path of aggregate productivity $\{\widehat{z}_t\}$ (which implies a higher path of $\{\widehat{y}_t^n\}$) increases consumption risk in this case and monetary policy cannot do anything to prevent it: higher productivity tends to increase output and hence the level of income risk that households face but tighter monetary policy, which would be needed to forestall the rise in output, tends to make μ higher which itself increases consumption risk.

Even though changes in interest rates (and hence μ) cannot affect consumption inequality after date 0, the planner can affect consumption inequality at date 0 (and thus at all subsequent dates, because $\widehat{\Sigma}_t$ depends on $\widehat{\Sigma}_{t-1}$). This is because monetary policy changes after date 0 are anticipated while the date 0 change in monetary policy is *unanticipated*. As described in section 4, an unanticipated cut in interest rates (and hence μ_0) effectively redistributes from savers to borrowers. To be clear, since we have an economy where agents hold real (and not nominal) claims, this is not because inflation redistributes date 0 real wealth from savers to borrowers: the date 0 distribution of real wealth is unaffected.[16] But the distribution of consumption is affected, as rich savers find that they receive a lower return on their bond holdings than they had anticipated while poor debtors find that their interest payments are smaller than they had expected.

In sum, while the planner seeks to reduce consumption inequality in addition to stabilizing prices and the gap output and its flexible-price level, this is not possible after date 0 since (to first order) the evolution of consumption inequality is unaffected by policy. Effectively, then after date 0, the planner faces the same trade-off between output and inflation as in the RANK economy. At date 0, it is *possible* to reduce consumption inequality via a surprise cut in interest rates which exploits households' unhedged interest rate exposure. Thus, the planner has an additional motive to cut rates at this date. This is reflected in the optimal design of monetary policy, as we now demonstrate.

Proposition 3 (Optimal monetary policy with $\Omega = 0$)**.** *Output and inflation* $\{\widehat{y}_t, \pi_t\}_{t=0}^{\infty}$ *under optimal*

[16] Section 6 discusses the case where households hold nominal debt.

policy satisfy

$$\eta_t = \begin{cases} \alpha\left(\widehat{y}_t - \delta\widehat{y}_t^n - \chi\right) & \textit{for } t = 0 \\ \widehat{y}_t - \widehat{y}_t^n & \textit{for } t \geq 1 \end{cases} \tag{38}$$

$$\eta_t = \frac{1}{\beta R}\eta_{t-1} - \frac{\lambda}{\lambda - 1}\pi_t \tag{39}$$

$$\pi_t = \frac{\widetilde{\beta}}{\vartheta}\pi_{t+1} + \kappa(\widehat{y}_t - \widehat{y}_t^n) \tag{40}$$

where $\alpha > 1, 0 < \delta < 1$ and $\chi > 0$ are defined in Appendix E.1 and η_t is the (normalized) multiplier on the Phillips curve (40).

Proof. See Appendix E.1. □

Combining (38) and (39), we get the following target criterion for all dates $t > 1$:

$$(\widehat{y}_t - \widehat{y}_t^n) - \frac{1}{\beta R}\left(\widehat{y}_{t-1} - \widehat{y}_{t-1}^n\right) + \frac{\lambda}{\lambda - 1}\pi_t = 0 \tag{41}$$

Comparing equations (32) and (41) shows that when $\Omega = 0$ the target criteria in HANK and RANK for $t > 0$ are almost identical (under complete markets we have $\beta R = 1$ so the former collapses to the latter). This reflects that fact that monetary policy cannot affect consumption risk at dates $t > 0$, and thus faces the same tradeoff as RANK. But this is not true at $t = 0$, where the target criterion is:

$$\widehat{y}_t - \delta\widehat{y}_t^n + \frac{1}{\alpha}\frac{\lambda}{\lambda - 1}\pi_t = \chi \tag{42}$$

Equation (42) shows that at time 0 optimal policy in HANK deviates from that in RANK in three ways. First, it is optimal to create a boom at date 0. To see this most clearly, suppose that productivity is at its steady-state value, so $\widehat{y}_t^n = \widehat{z}_t = 0$ for all t. Even in this case, it is optimal to move away from $\widehat{y}_0 = \pi_0 = 0$ and implement an output boom $\widehat{y}_0 > 0$ which is accompanied by inflation $\pi_0 > 0$. This is because, as discussed earlier, the evolution of Σ_t is different at date 0, compared to all other dates. A surprise cut in interest rates at date 0 reduces consumption inequality–as is summarized by the last term of (37) for $t = 0$. The constant term $\chi > 0$ in the date-0 target criterion (42) reflects exactly this benefit from cutting rates and reducing inequality at date 0. This is *not* because it is infeasible to set $\widehat{y}_t = \widehat{y}_t^e = 0$ and $\pi_t = 0$ in the HANK economy; this remains feasible, and it remains costly for the planner to deviate from this benchmark. But the costs of doing so are balanced by benefits of reducing consumption inequality via a surprise interest rate cut. Note that the desirability of exploiting households' unhedged interest-rate exposure for redistribution (the URE channel) makes the Ramsey plan *time-inconsistent*. Suppose that the planner has been following a Ramsey plan since $t = -\infty$ and the economy has converged to steady state. Given the opportunity to deviate from this plan at date 0, the planner would do so, lowering interest rates temporarily - i.e. the continuation of a Ramsey plan is not a Ramsey plan.

Second, whereas in RANK, optimal policy seeks to move output $\widehat{y}_t$ one-for-one with its efficient level $\widehat{y}_t^e$ ($= \widehat{y}_t^n$ in RANK), which fluctuates due to productivity shocks, under HANK it is optimal to track the flexible-price level of output $\widehat{y}_t^n$ less than one-for-one. In other words, $0 < \delta < 1$ in (42). Figure 1 depicts the optimal level of date 0 output as a function of flexible-price level of output $\widehat{y}_t^n$ in HANK and RANK.

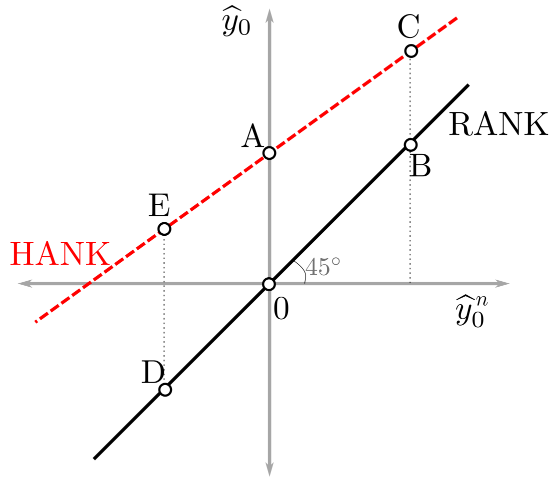

Figure 1. Optimal level of $\widehat{y}_0$ in HANK ($\Omega = 0$) and RANK absent markup shocks

First, suppose, $\widehat{y}_0^n = 0$. In RANK, it is optimal to track this and to set $\widehat{y}_0 = \widehat{y}_0^n = 0$. But in HANK there is a first order benefit from cutting interest rates to reduce inequality, creating a boom in output, until the marginal benefit of an additional reduction in nominal rates is outweighed by the cost of distorting output further above its efficient level (point A in Figure 1). Next, suppose that $\widehat{y}_0^n > 0$. Again, the RANK planner sets $\widehat{y}_0 = \widehat{y}_0^n$ (denoted by point B in the Figure). If the HANK planner were also to set $\widehat{y}_0 = \widehat{y}_0^n > 0$, this would already generate a surprise fall in interest rates, which would reduce inequality to some extent. The HANK planner still perceives some additional benefit to reducing inequality further but the marginal benefit is smaller since inequality has already been reduced. Consequently, it is not optimal to deviate as much from productive efficiency as in the case $\widehat{y}_0^n = 0$, and so the gap between points C and B is smaller than that between points A and the origin. Conversely, when $\widehat{y}_0^n < 0$, tracking the flexible price allocation would entail a surprise interest rate increase which would increase consumption inequality. The benefit of deviating from this RANK allocation is larger in this case and therefore it is optimal to tolerate a larger deviation from production efficiency: the gap between points D and E is larger than that between point A and the origin.

Finally, the date 0 target criterion puts less weight on inflation relative to output $\alpha > 1$. This is because, given the constraint on monetary policy imposed by (36), the evolution of consumption inequality depends directly on the sequences $\{\widehat{y}_t, \widehat{y}_t^n\}_{t=0}^{\infty}$ but not on inflation. Thus, the target criterion puts relatively less weight on inflation and more on output, which attains heightened importance due to its effects on inequality.

Figure 2 shows the dynamics of optimal policy absent shocks in the economy with $\Omega = 0$. The policy maker cuts nominal interest rates at date 0 (panel d), which generates a fall in real rates and in the passthrough from income risk to consumption risk μ_t (panel e). The fall in consumption risk in turn reduces consumption inequality on impact (panel c), after which it gradually returns to steady state. The fall in interest rates also generates a boom in output (panel a) and inflation (panel b). Since there are no shocks to the efficient level of output in this scenario, both the boom in output and increase in inflation are in themselves undesirable from an efficiency perspective, even if they are a price worth paying for a persistent reduction in inequality. In order to arrest the increase in inflation, the planner commits to tighten policy, generating a fall in output and deflation, from date 1 onwards. Since firms are forward looking, a commitment to lower inflation in the future mitigates the date 0 inflation caused by the boom

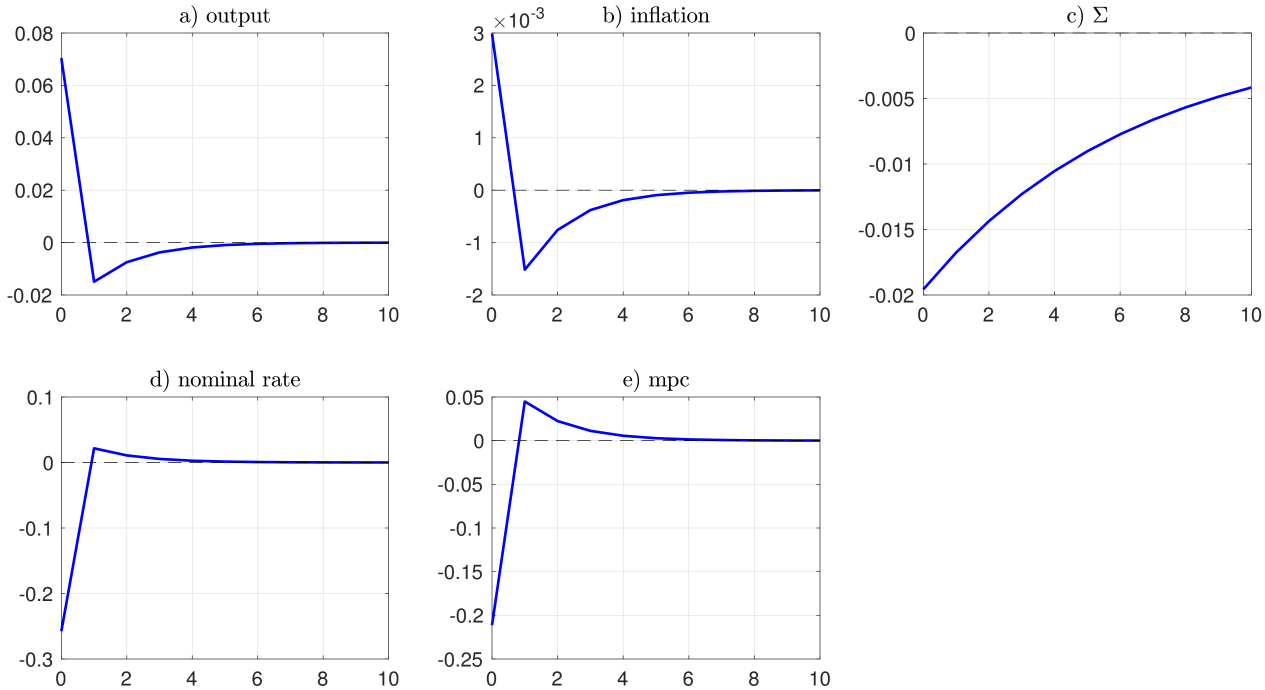

Figure 2. Time Inconsistency: Optimal dynamics absent shocks. All variables are plotted as percentage deviations from their steady state values.

in output.

Next, we discuss the optimal dynamic response of the economy to a negative productivity shock. Figure 3 shows the impulse responses to a negative productivity shock under optimal policy in the HANK and RANK models, defined as the difference between outcomes with and without the shock. That is, for any variable of interest x, we plot $x_t|_{z_0=-0.01} - x_t|_{z_0=0}$.[17] In the RANK economy (red lines), the planner allows output to fall (panel a) in line with the natural rate of output, keeping the output gap (panel f) and inflation (panel b) equal to zero. Implementing a fall in output requires an increase in nominal interest rates (panel d), but since agents do not face idiosyncratic risk, this has no effect on consumption inequality which is always zero.

In the HANK economy (blue lines), a sharp increase in interest rates would increase the passthrough from income to consumption risk and would persistently increase inequality. To avoid this, the planner actually cuts interest rates at date 0, dampening the increase in passthrough μ_t and the increase in inequality. Output does not fall as much as the flexible price level of output (as shown by the output gap in panel f), and as a result inflation increases at date 0 - even relative to the scenario absent shocks. This increase in inflation is not costless. To mitigate its impact, the planner commits to a slightly higher path of interest rates, and lower output and inflation, from date 1 onwards.

Understanding the forces generating time inconsistency The difference in the optimal target criterion than at all other dates arises only because there is existing wealth inequality at the start of date 0. Suppose that there is an equalization of asset positions across all households at the beginning of date 0.

[17] In the RANK economy, the impulse response is identical to the response of the economy to the shock, since in the absence of shocks, optimal policy keeps all variables at their steady state values and all deviations from steady state equal zero. This is not true in the HANK economy, where (as shown in Figure 2) the Ramsey planner deviates from steady state even absent shocks.

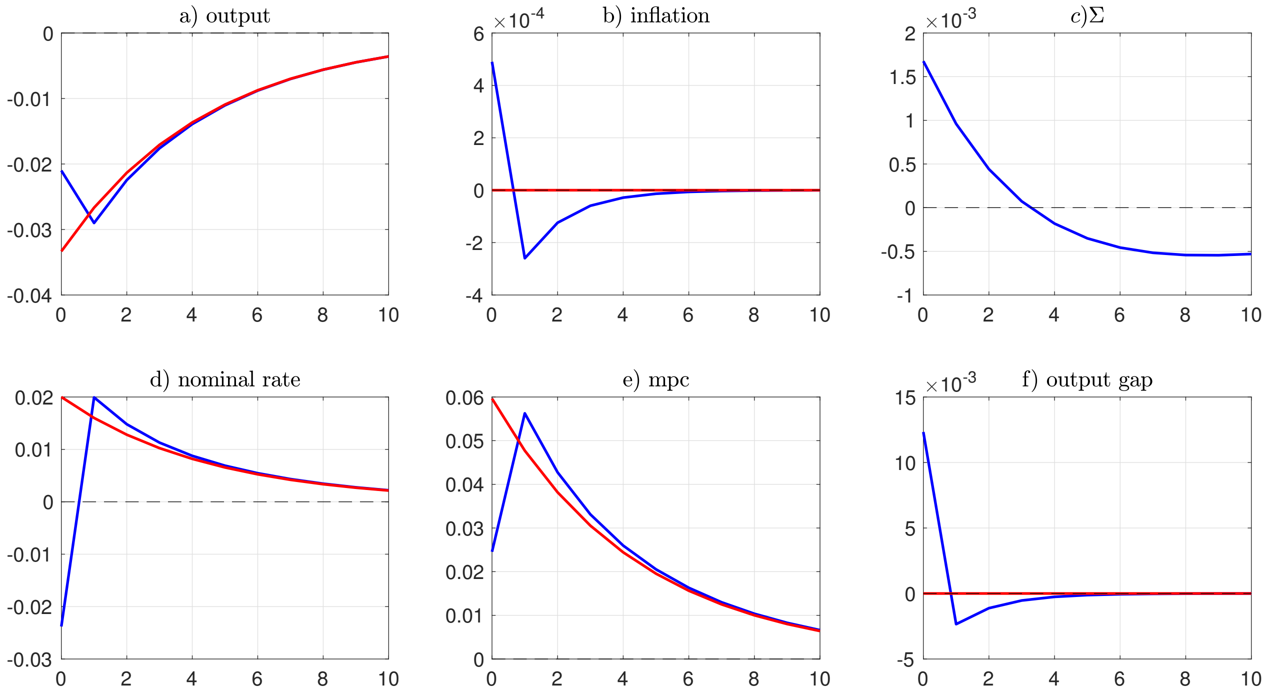

Figure 3. HANK ($\Omega = 0$) vs RANK optimal response to a negative productivity shock. Red line denotes RANK while blue line represents HANK ($\Omega = 0$). All variables are plotted as percentage deviations from their steady state values. In panel f), we plot output gap as the difference between y_t and y_t^n normalized by steady state y: $(\widehat{y}_t/y) - (\widehat{y}_t^n/y)$.

In this scenario, the URE channel is inoperative, as discussed in section 4 and surprise changes in μ_0 have no additional effect relative to anticipated changes. When there is existing wealth inequality, a surprise cut in the path of interest rates (lower μ_0) lowers the consumption of rich savers and raises the consumption of poor borrowers, reducing consumption inequality. Consequently, at $t = 0$, the planner deviates from the date $t > 1$ target criterion by cutting interest rates do engineer precisely this outcome. Absent existing wealth inequality at the beginning of date 0, clearly there are no borrowers and savers to redistribute between and this motive is absent. In fact, with $\Omega = 0$, absent initial wealth inequality, equation (37) becomes:

$$\frac{\widehat{\Sigma}_t}{\Sigma} = \Lambda\Gamma_t + \frac{\vartheta}{\beta R}\frac{\widehat{\Sigma}_{t-1}}{\Sigma} \qquad \forall t \geq 0 \tag{43}$$

with $\frac{\widehat{\Sigma}_{-1}}{\Sigma} = -\frac{\beta R}{\vartheta}\ln\left(\frac{1-\vartheta}{1-\vartheta e^{\frac{\Lambda}{2}}}\right) < 0$. Equation (43) shows that without initial wealth inequality, the planner is unable to affect Σ_t at any date (up to first-order). Unsurprisingly then, the planner faces the standard trade-off between price stability and productive inefficiency. Consequently, the optimal target criterion at all dates in this case simplifies to:

$$(\widehat{y}_t - \widehat{y}_t^n) - \frac{1}{\beta R}\left(\widehat{y}_{t-1} - \widehat{y}_{t-1}^n\right) + \frac{\lambda}{\lambda - 1}\pi_t = 0 \tag{44}$$

with $\widehat{y}_{-1} = \widehat{y}_{-1}^n = 0$ and so *divine coincidence* holds - optimal policy implements $\pi_t = 0$ and $\widehat{y}_t = \widehat{y}_t^n$ at all dates $t \geq 0$ in response to productivity shocks.

Remark 2 (The *interest rate exposure channel* versus the *Fisher channel*)**.** *Previous research has shown* *how the redistribution of wealth through inflation when assets and liabilities are nominal may create an inflationary bias on the part of the central bank. This "Fisher channel" is absent in our baseline economy where households only hold real assets: the planner is unable to redistribute real wealth via surprise inflation. Nonetheless, the planner can redistribute consumption via a surprise cut in the path of real rates via the* interest rate exposure channel *(Auclert, 2019). Auclert (2019) studies how both these channel mediate the effect of interest rate changes on aggregate consumption in an environment with MPC heterogeneity. In our economy, all agents have the same MPC, so monetary policy does not affect aggregate consumption through these channels. However, it does affect the distribution of consumption and in our baseline economy, optimal monetary policy deliberately utilizes the interest rate exposure channel to reduce consumption inequality. In Section 6, when we introduce nominal bonds into our model, optimal monetary policy exploits both channels to achieve the same ends.*

5.4.2 Optimal policy with countercyclical risk

While the case with $\Omega = 0$ is a useful benchmark to understand the forces driving optimal monetary policy, the empirically relevant case is one of countercyclical risk, which implies that $\Omega > 0$.[18] In this case a cut in interest rates, and the associated boom in output (and wages), reduce consumption risk through both the income risk and self-insurance channels. With $\Omega = 0$, risk was procyclical, and the benefit of lower interest rates (which lower μ and increase households' ability to self insure) was exactly balanced by the cost of the associated boom in output which increased the level of income risk faced by households. In contrast, when risk is countercyclical, the increase in output actually *reduces* the level of income risk, at the same time as lower interest rates improve households' ability to self-insure: expansionary policy unambiguously reduces consumption risk and inequality. This can be seen by substituting (34) into (33):

$$\frac{\widehat{\Sigma}_t}{\Sigma} = \begin{cases} -\gamma\left(1-\widetilde{\beta}\right)\Omega\widehat{y}_t + \Lambda\Gamma_t + \frac{\vartheta}{\beta R}\frac{\widehat{\Sigma}_{t-1}}{\Sigma} & \text{for } t>0 \\ \\ -\gamma\left(1-\widetilde{\beta}\right)\Omega\widehat{y}_0 + \Lambda\Gamma_0 + \frac{\vartheta\Lambda\Sigma}{1-\vartheta}\widehat{\mu}_0 & \text{for } t=0 \end{cases} \tag{45}$$

When $\Omega > 0$, an increase in output $\widehat{y}_t$ reduces the consumption risk faced by households at any date, not just at date 0: $\frac{\partial\widehat{\Sigma}_t}{\partial\widehat{y}_t} < 0$. In this environment, the planner would only refrain from increasing output if doing so distorts productive efficiency and creates inflation. Indeed recall from (26) that in steady state with $\Omega > 0$, the marginal benefit of reducing inequality by stimulating demand is exactly balanced by the cost of deviating further from production inefficiency. The reason a marginal increase in output has a first-order cost in terms of productive efficiency is that the steady state in this case features an inefficiently high level of output $y > 1$ (also reflected in $w > 1$). Out of steady state the planner faces a similar trade-off between the benefit of reducing inequality via higher output and the cost of further distorting productive efficiency, while also seeking to limit deviations from price stability. This tradeoff informs the optimal design of monetary policy, as we discuss next.

[18] More precisely, countercyclical risk ($\Theta > 1$) implies $\Omega > \frac{\Lambda}{(1-\widetilde{\beta})(1-\Lambda)} > 0$. When $0 < \Omega < \frac{\Lambda}{(1-\widetilde{\beta})(1-\Lambda)}$, risk is weakly procyclical so expansionary monetary policy reduces pass through than it increase income risk, reducing consumption risk in net.

Proposition 4 (Optimal monetary policy with countercyclical risk). *Output and inflation* $\{\widehat{y}_t, \pi_t\}_{t=0}^{\infty}$ *under optimal policy satisfy*

$$\eta_t = \begin{cases} \alpha_0(\Omega)\Big(\widehat{y}_t - \delta_0(\Omega)\widehat{y}_t^n - \chi(\Omega)\Big) & \text{for } t = 0 \\ \alpha(\Omega)\Big(\widehat{y}_t - \delta(\Omega)\widehat{y}_t^n\Big) & \text{for } t \geq 1 \end{cases} \tag{46}$$

$$\eta_t = \frac{1}{\beta R}\eta_{t-1} - \frac{\lambda}{\lambda - 1} y \pi_t \tag{47}$$

$$\pi_t = \frac{\widetilde{\beta}}{\vartheta}\pi_{t+1} + \kappa(\widehat{y}_t - \widehat{y}_t^e) + u_t \tag{48}$$

where $\alpha(\Omega), \delta(\Omega), \alpha_0(\Omega), \delta_0(\Omega)$ *and* $\chi(\Omega)$ *are defined in Appendix E.1 and satisfy* $\alpha_0(0) = \alpha$, $\delta_0(0) = \delta$, $\chi(0) = \chi$, $\alpha(0) = 1, \delta(0) = 1$. *Further, when risk is countercyclical* $\alpha(\Omega), \alpha_0(\Omega) > 1$ *and* $\delta(\Omega), \delta_0(\Omega) < 1$.

Proof. See Appendix E.1. □

By combining (46) with (48), one can derive the optimal target criterion for $t = 0$:

$$\Big(\widehat{y}_0 - \delta_0(\Omega)\widehat{y}_0^n\Big) + \frac{\lambda}{\lambda - 1}\frac{y}{\alpha_0(\Omega)}\pi_0 = \chi(\Omega) \qquad \text{for } t = 0 \tag{49}$$

and for dates $t > 1$:

$$\Big(\widehat{y}_t - \widehat{y}_{t-1}\Big) - \delta(\Omega)\Big(\widehat{y}_t^e - \widehat{y}_{t-1}^*\Big) + \frac{\lambda}{\lambda - 1}\frac{y}{\alpha(\Omega)}\pi_t = 0 \tag{50}$$

As in the case $\Omega = 0$, optimal policy deviates from RANK at date 0, creating a boom even absent exogenous shocks. However, unlike in the case $\Omega = 0$, with countercyclical risk optimal policy also deviates from the RANK target criterion at subsequent dates $t > 1$. This should not be not surprising: with $\Omega = 0$, monetary policy can only affect the evolution of consumption inequality (to first order) at date 0. However, with countercyclical risk, an expansion in output reduces inequality at *all* dates, so monetary policy faces a tradeoff between the benefits of lower inequality and the cost of distorting productive efficiency and price stability at all dates, and deviates from RANK as a result. In particular, $\delta(\Omega) < 1$ when risk is countercyclical: output tracks the flexible price level of output less than one for one at all dates. In addition, $\alpha(\Omega) > 1$: the planner puts more weight on stabilizing output relative to inflation at all dates.

To understand why output moves less than one-for-one with its flexible price level at all date, $\delta(\Omega) < 1$, recall the distinction between three different levels of output: the *productively efficient* level of output y_t^e, the *flexible price* level of output y_t^n, and the *equilibrium* level of output y_t. When $\Omega = 0$, $y_t^n = y_t^e$. In this case, the standard subsidy corrects for the distortion due to monopolistic competition, and productive efficiency is obtained when there is zero inflation at all dates, or when prices are flexible. In other words, the flexible price version of the economy with $\Omega = 0$ features a *zero labor wedge* at all dates, $\mathcal{T}_t \equiv \ln z_t - \ln w_t^n = 0$ where w_t^n denotes the wage consistent with the flexible price level of output.

However, with countercyclical risk, $\Omega > 0$, and the flexible price level of output y_t^n is higher than the efficient level of output y_t^e, whatever the level of productivity. In this case, firms enjoy a subsidy which is larger than the one necessary to eliminate the distortion due to monopolistic competition (see eq. (27)), and output would be inefficiently high under flexible prices. The flexible price version of this economy features a *negative* labor wedge - a net subsidy to output and employment - at all times. Consequently keeping

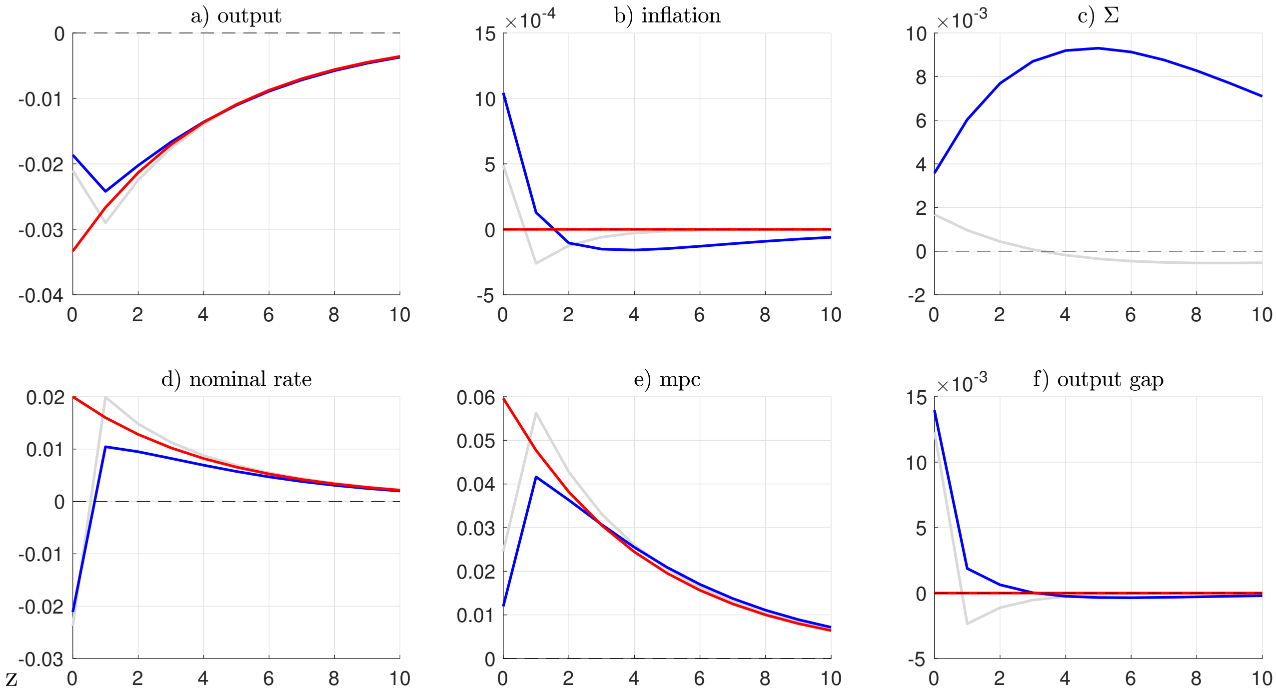

Figure 4. HANK ($\Omega > 0$) vs RANK optimal response to a negative productivity shock. Red line denotes RANK, blue line represents HANK with countercyclical risk ($\Omega > 0$), and gray line represents HANK with $\Omega = 0$. All variables are plotted as percentage deviations from their steady state values. In panel f), we plot output gap as the difference between y_t and y_t^n normalized by steady state y: $(\widehat{y}_t/y) - (\widehat{y}_t^n/y)$.

output equal to its flexible price level entails keeping output inefficiently high at all times. Conversely, in order to set output equal to its efficient level, and undo the effects of the net subsidy to output and employment, one would have to drive output y_t down below its flexible price level y_t^n.

Absent shocks, the planner *could* always implement the efficient level of output by raising interest rates above their steady state level, reducing output below its inefficiently high flexible price level, and eliminating the negative labor wedge. She chooses *not* to do this, however, because the cost in terms of increased consumption inequality is too high - so she sets output equal to its flexible price level.

Now consider a scenario in which productivity is higher than in steady state ($z > 1$). In this case, the planner could continue to set output equal to its flexible price level, keeping the labor wedge unchanged ($\mathcal{T}_t = \ln z_t - \ln w_t = -\ln w < 0$). This would entail increasing y_t one for one with y_t^n. But it is no longer optimal to do so. The only reason the planner does *not* eliminate the inefficient subsidy in steady state is that doing so would increase inequality too much. Now however, with higher output, consumption inequality is already lower, and the benefit of the inefficient subsidy is lower: even if the planner were to implement a lower level of output relative to the flexible price level, improving productive efficiency, inequality would still be relatively low. Optimal policy seeks to reduce the magnitude of the labor wedge by bringing the level of output y_t closer to the efficient level of output y_t^e. Since $1 = y^e < y^n = y$ in steady state, this entails $\widehat{y}_t < \widehat{y}_t^n$.

Conversely, suppose productivity falls relative to steady state, $z < 1$. Again, the planner could track the flexible price level of output y_t^n, keeping the labor wedge unchanged. Doing so would entail reducing output one for one with the flexible price level of output. But when output is lower, consumption inequality is higher. In this case, the rationale for keeping output above its efficient level y_t^e is not eliminated - it is actually strengthened. It is optimal to deviate even *further* from productive efficiency, increase the effective subsidy to output and employment, and reduce output less than one for one with the efficient level of output.

Figure 4 shows the impulse responses to a negative productivity shock under optimal policy in HANK with countercyclical risk (blue lines), with outcomes in RANK (red lines) and HANK with $\Omega = 0$ (gray lines) shown as benchmarks. As in the benchmark with $\Omega = 0$, the planner deviates from optimal policy in RANK - which tracks the flexible price level of output - by cutting nominal interest rates at date 0 (panel d) and mitigating the fall in output (panel a) on impact, at the cost of creating an increase in inflation (panel b). Recall that in the economy with $\Omega = 0$, the planner attempted to dampen this increase in inflation by committing to tighter policy starting at date 1. With countercyclical risk, however, any fall in output is costly because it directly increases inequality. Indeed, even under optimal policy, the path of inequality is higher with countercyclical risk (panel c). To prevent inequality from rising even more, the planner postpones the commitment to tighter policy in the future. While interest rates still rise at date 1 (blue line in panel d), the increase is smaller than in the RANK economy (red line) or the HANK economy with $\Omega = 0$ (gray line). As a result, the fall in inflation and output-gap are also smoothed out over time.

Remark 3 (The role of initial wealth inequality with countercyclical risk)**.** *With countercyclical risk ($\Omega > 0$), the optimal target criterion is different at date $t = 0$ and at all subsequent dates, as was the case when $\Omega = 0$. In particular, optimal policy features a boom at date 0 even absent exogenous shocks. As in the $\Omega = 0$ case, such a policy is optimal because given existing wealth inequality at date 0, a surprise cut in interest rates raises the consumption of poor borrowers and reduces the consumption of rich savers. To understand the role of wealth inequality in driving this result, suppose again that there is an equalization of household wealth at the beginning of date 0. In this case, the optimal target criterion* (50) *characterizes optimal policy at all dates (with $\widehat{y}_{-1} = \widehat{y}^*_{-1} = 0$). This implies that absent shocks, the planner keeps interest rates at their steady state value. However, unlike in the case with $\Omega = 0$ and wealth equalization, it is no longer optimal to implement zero inflation and replicate flexible price allocations in response to productivity shocks - even though it remains feasible to do so.*

Figure 5 plots dynamics under optimal policy in this scenario. Qualitatively, policy is similar to the case without wealth equalization in which there is a time inconsistency problem. Relative to RANK, the planner implements a lower increase in interest rates on impact (panel d), resulting in a lower increase in output (panel a) and the passthrough from income to consumption risk μ_t (panel e). This comes at the cost of a short-lived increase in inflation. Note that inequality (shown in the panel c) is below its steady state level, despite the increase in income risk caused by the fall in output, because the initial equalization of wealth reduces consumption inequality at the start of date 0.

6 Nominal debt

So far we have considered an economy in which households traded inflation indexed bonds. As mentioned in Remark 2, we made this assumption to distinguish between two ways in which monetary policy can affect the distribution of consumption in HANK economies. The first is the *interest rate exposure channel*: an unanticipated fall in real interest rates increases the consumption of poor debtors and reduces that of rich savers. The second is the *Fisher effect*: unanticipated inflation redistributes real wealth from savers

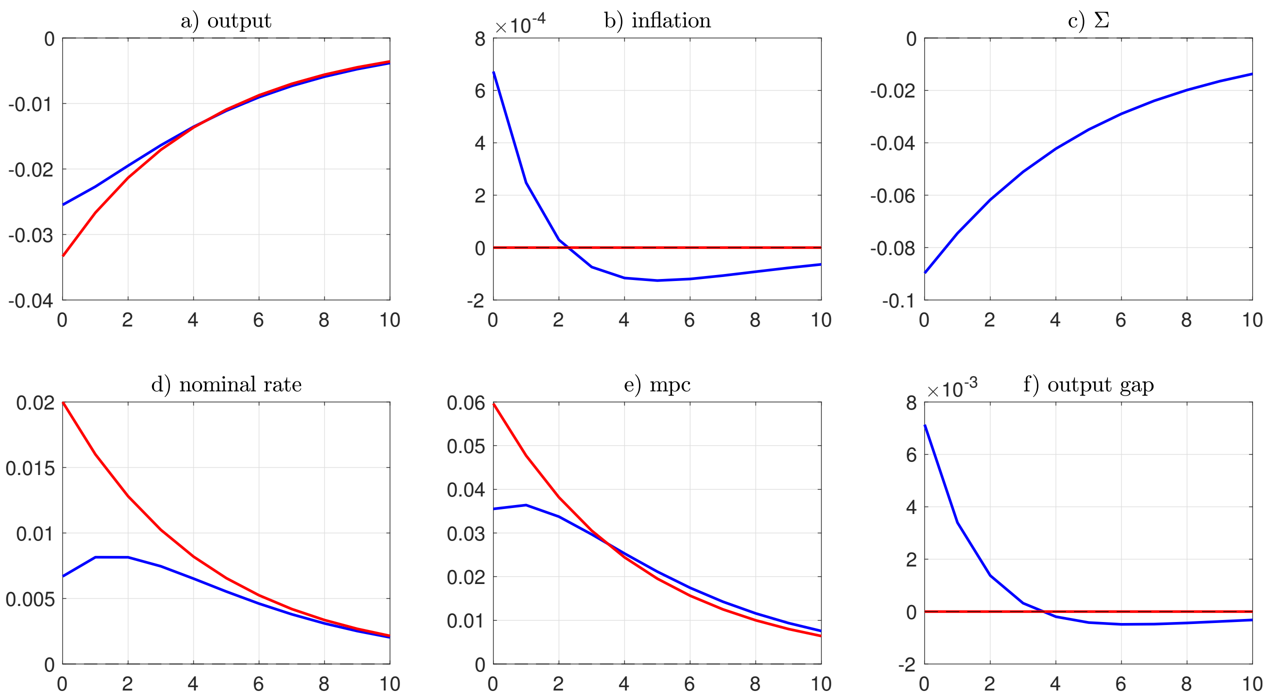

Figure 5. HANK ($\Omega > 0$) vs RANK optimal response to a negative productivity shock in the economy with wealth equalization at the beginning of date 0. Red line denotes RANK, blue line represents HANK with countercyclical risk ($\Omega > 0$). All variables are plotted as percentage deviations from their steady state values. In panel f), we plot output gap as the difference between y_t and y_t^n normalized by steady state y: $(\widehat{y}_t/y) - (\widehat{y}_t^n/y)$.

who hold nominal assets to debtors with nominal liabilities. Our baseline model with inflation indexed debt abstracts altogether from the second effect to focus on the first. In this section, we allow households to trade nominal debt and show how the the presence of the Fisher effect changes the optimal conduct of monetary policy.

Financial intermediaries now trade nominal claims at nominal price $\frac{\vartheta}{1+i_t}$ which pay a dollar tomorrow. Then the household's budget constraint can be written as:

$$P_t c_t^s\left(i\right) + \frac{\vartheta}{1+i_t} A_{t+1}^s\left(i\right) = P_t w_t \ell_t^s\left(i\right) + A_t^s\left(i\right) + P_t T_t$$

where $A_{t+1}^s\left(i\right)$ is the quantity of nominal claims purchased by the household at date t. The details of this extension are in Appendix F. As before the evolution of this economy is characterized by the aggregate IS equation (11), the evolution of μ(12), Phillips curve (7), the definition of GDP (15) and the evolution of Σ, replacing $R_t = \frac{1+i_t}{\Pi_{t+1}}$. Four of these five equations are unaffected by the introduction of nominal bonds - the exception is the evolution of Σ, which becomes:[19]

$$\ln \Sigma_t = \frac{1}{2}\gamma^2 \mu_t^2 w_t^2 \sigma_t^2 + \ln\left[1 - \vartheta + \vartheta \Sigma_{t-1}\right] + \mathbb{I}(t=0) \ln \left(\frac{1 - \vartheta e^{\frac{\Lambda}{2}}}{1 - \vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mathbb{E}_{-1}\mu_0}\frac{\mathbb{E}_{-1}\Pi_0}{\Pi_0}\right)^2}} \right) \tag{51}$$

(51) shows that the presence of nominal debt means that unanticipated higher inflation reduces consumption inequality. Since there are no aggregate shocks except at date 0 and the Ramsey planner can only re-optimize at date 0, actual and expected inflation coincide in equilibrium except at date 0. As Appendix

[19] See Appendix F for the derivation.

F details, the cohort born at date $s < 0$ enters date 0 with with a cross-sectional distribution of real wealth which is $N\left(0, -sw^2\sigma^2\left(\frac{\mathbb{E}_{-1}\Pi_0}{\Pi_0}\right)^2\right)$. Higher than expected inflation $\Pi_0 > \mathbb{E}_{-1}\Pi_0$ compresses the distribution of real wealth. Thus, by generating inflation at date 0, the planner can reduce wealth and hence consumption inequality. This reduction in consumption inequality is in addition to the reduction achieved by the surprise cut in interest rates at date 0.

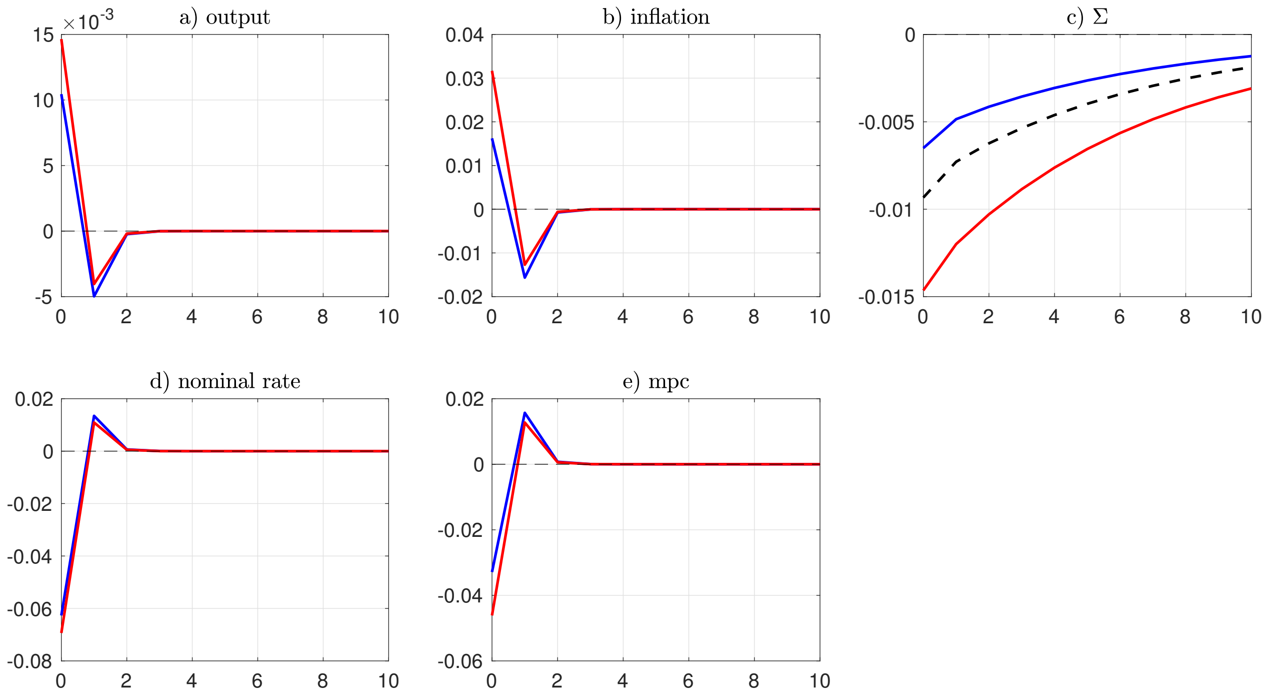

Figure 6. Time inconsistency with nominal debt: Optimal dynamics absent shocks in alternative calibration ($\kappa = 0.5$). Blue line denotes HANK with countercylcical risk and real debt, red line denotes HANK with countercyclical risk and nominal debt. All variables are plotted as percentage deviations from their steady state values.

Qualitatively, this increases the planner's incentive to create a boom at date 0 even absent aggregate shocks. Quantitatively, though, this effect is small in our baseline calibration, as Figure 8 in Appendix F shows - optimal policy in the two economies is essentially identical. This is primarily because in our baseline calibration, the Phillips curve is relatively flat ($\kappa = 0.01$) which means that even a large cut in real interest rates, and a large boom in output, generates only a small increase in inflation. In order to generate a large enough increase in inflation to effect significant redistribution, it would be necessary to engineer a massive deviation from the productively efficient level of output, and it is not optimal for the planner to do this.

To illustrate the qualitative effect of introducing inflation-indexed debt, Figure 6 shows optimal policy absent shocks with a steeper Phillips curve ($\kappa = 0.5$). As panel b shows, the planner creates a larger increase in inflation in the economy with nominal debt (red line) than in the economy with inflation-indexed debt (blue line). This results in a larger reduction in inequality (panel c). The black dashed line in this panel shows the effect of implementing the optimal policy from the economy with inflation-indexed debt in the economy with nominal debt. Even if the planner follows this policy, and does not actively exploit the Fisher effect, this already automatically generates a larger reduction in inequality than in the economy with inflation-indexed debt (the black dashed line is below the blue line) because the same increase in inflation now redistributes real wealth. The planner actually generates a further increase in inflation, and

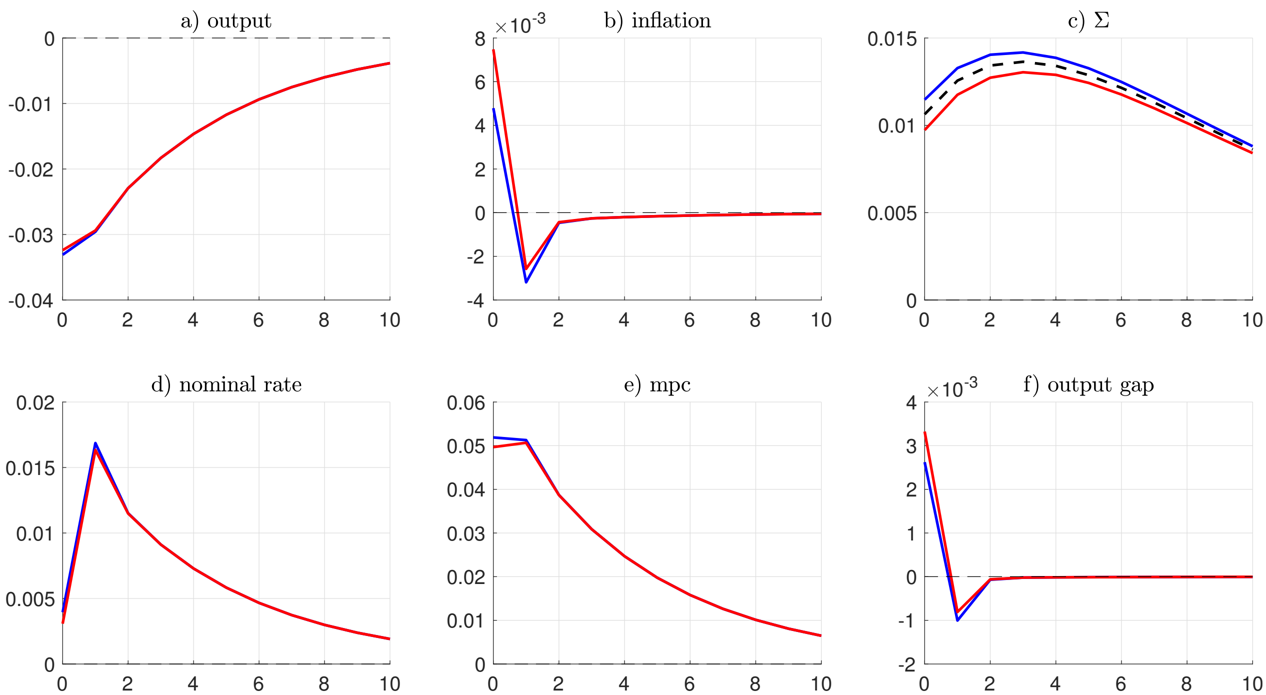

Figure 7. Nominal vs inflation indexed debt: optimal response to a negative productivity shock in alternative calibration ($\kappa = 0.5$). Blue line represents HANK with countercyclical risk and real debt, red line denotes HANK with countercyclical risk and nominal debt. All variables are plotted as percentage deviations from their steady state values. In panel f), we plot output gap as the difference between y_t and y_t^n normalized by steady state y: $(\widehat{y}_t/y) - (\widehat{y}_t^n/y)$.

so inequality falls even more (the red line is below the black dashed line).

The presence of nominal assets and liabilities also affects the optimal response to shocks. Figure 7 plots the impulse responses (defined as in the previous section) to a negative productivity shock under optimal policy in the alternative calibration with $\kappa = 0.5$. Allocations under optimal policy in the economy with nominal debt (red line) are similar to those with inflation-indexed debt (blue line), even with a steep Phillips curve. However, the level of output now deviates even more from y_t^n (panel f) in order to generate a larger surprise increase in inflation (panel b). This is optimal when risk is countercyclical because negative productivity shock tends to increase inequality (panel c). Creating more surprise inflation partially offsets this, yielding a lower increase in inequality (red line) than would obtain if the planner followed the same policy as in the economy with inflation-indexed debt (black dashed line). As in the previous figure, though, even following this same policy would automatically generate a larger reduction in inequality than in the economy with inflation-indexed debt (blue line).

7 Conclusion

We use a analytically tractable HANK model to study how monetary policy affects inequality, and the extent to which this warrants a change in the principles governing the design of monetary policy which have been developed in the Representative Agent New Keynesian literature. In a complete markets economy, monetary policy affects output and inflation but, trivially, has no effects on inequality (since households can use Arrow securities to insure themselves, eliminating consumption inequality). In our incomplete markets economy, monetary policy affects consumption risk and inequality through four channels. The first is the *income risk channel*: when the idiosyncratic income risk faced by households is countercyclical, expansionary monetary policy, by generating a boom in output, tends to reduce income risk and inequality. More subtly, lower interest rates make it easier for households to self-insure against income shocks, reducing consumption risk for a given level of income risk - the *self-insurance channel*. Finally, unexpected cuts in interest rates redistribute consumption through an *unhedged interest rate exposure channel*, and unexpected inflation redistributed real wealth through the *Fisher channel*. Thus, expansionary monetary policy can reduce consumption inequality through all four of these channels.

Given that monetary policy has this power to reduce inequality, how and when should it be used? A utilitarian planner trades off the benefits of lower inequality against the costs of pushing up output and inflation above their efficient levels. In recessions, inequality is already high - in the relevant case with countercyclical risk - so the marginal benefit of reducing inequality is particularly high. Consequently, optimal monetary policy is more accommodative in recessions relative in to a RANK benchmark: the planner prevents output from falling as much as the efficient level of output, even though this entails higher inflation, because curtailing the fall in output also curtails the rise in inequality.

References

Acharya, Sushant and Keshav Dogra, “Understanding HANK: insights from a PRANK,” Staff Reports 835, Federal Reserve Bank of New York February 2018.

Angeletos, George-Marios and Laurent Emmanuel Calvet, “Idiosyncratic Production Risk, Growth and the Business Cycle,” *Journal of Monetary Economics*, 2006, *53*, 1095 – 1115.

Auclert, Adrien, “Monetary Policy and the Redistribution Channel,” *American Economic Review*, 2019, *109* (6), 2333 – 2367.

_ **, Matt Rognlie, and Ludwig Straub**, “The Intertemporal Keynesian Cross,” Working Paper 25020 2018.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J Sargent, “Inequality, Business Cycles and Monetary-Fiscal- Policy,” Working Paper 24710, National Bureau of Economic Research June 2018.

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

_ , “Monetary Policy and Heterogeneity: An Analytical Framework,” Working Paper, November 2019.

_ , “The New Keynesian Cross: Understanding Monetary Policy with Hand-to-Mouth Households,” *Journal of Monetary Economics*, 2019, *Forthcoming.*

_ **and Xavier Ragot**, “Optimal Monetary Policy and Liquidity with Heterogeneous Households,” Working Papers 2018.

Blanchard, Olivier Jean and Jordi Galí, “Real Wage Rigidities and the New Keynesian Model,” *Journal of Money, Credit and Banking*, 2007, *39*, 35 – 65.

Caballero, Ricardo J., "Consumption Puzzles and Precautionary Savings," *Journal of Monetary Economics*, 1990, *25* (1), 113 – 136.

Cagetti, Marco, "Wealth Accumulation over the Life Cycle and Precautionary Savings," *Journal of Business & Economic Statistics*, 2003, *21* (3), 339–353.

Calvet, Laurent Emmanuel, "Incomplete Markets and Volatility," *Journal of Economic Theory*, 2001, *98*, 295 – 338.

Calvo, Guillermo A and Maurice Obstfeld, "Optimal Time-Consistent Fiscal Policy with Finite Lifetimes," *Econometrica*, 1988, *56* (2), 411–432.

Challe, Edouard, "Uninsured Unemployment Risk and Optimal Monetary Policy in a Zero-Liquidity Economy," *American Economic Journal: Macroeconomics*, April 2020.

Christelis, Dimitris, Dimitris Georgarakos, Tullio Jappelli, and Maarten van Rooij, "Consumption Uncertainty and Precautionary Saving," CSEF Working Papers 421, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy December 2015.

Cui, Wei and Vincent Sterk, "Quantitative Easing," Working Papers, UCL 2019.

Debortoli, Davide and Jordi Galí, "Monetary Policy with Heterogeneous Agents: Insights from TANK models," Working Paper 2018.

den Haan, Wouter, Pontus Rendahl, and Markus Riegler, "Unemployment (Fears) and Deflationary Spirals," *Journal of the European Economic Assocation*, 2018, *16(5)*, 1281 – 1349.

Fagereng, Andreas, Luigi Guiso, and Luigi Pistaferri, "Firm-Related Risk and Precautionary Saving Response," *American Economic Review*, May 2017, *107* (5), 393–97.

Galí, Jordi, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications, Second Edition*, Princeton University Press, 2015.

Galí, Jordi, David Lopez-Salido, and Javier Valles, "Understanding the Effects of Government Spending on Consumption," *Journal of the European Economic Association*, 2007, *5* (1), 227–270.

Guvenen, Fatih, Serdar Ozkan, and Jae Song, "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 2014, *122* (3), 621–660.

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

Kekre, Rohan, "Unemployment Insurance in Macroeconomic Stabilization," Working Paper 2019.

McKay, Alisdair and Ricardo Reis, "The Role of Automatic Stabilizers in the U.S. Business Cycle," *Econometrica*, January 2016, *84*, 141–194.

_ and _ , "Optimal Automatic Stabilizers," Working Paper 2019.

— , **Emi Nakamura, and Jon Steinsson**, "The Power of Forward Guidance Revisited," *American Economic Review*, October 2016, *106* (10), 3133.

Nisticò, Salvatore, "Optimal Monetary Policy And Financial Stability In A Non-Ricardian Economy," *Journal of the European Economic Association*, October 2016, *14* (5), 1225–1252.

Nuño, Galo and Carlos Thomas, "Optimal Monetary Policy with Heterogeneous Agents," Working Papers 1624, Banco de Espana November 2019.

Ravn, Morten O. and Vincent Sterk, "Job Uncertainty and Deep Recessions," *Journal of Monetary Economics*, 2017, *90* (C), 125–141.

— **and** — , "Macroeconomic Fluctuations with HANK & SAM: An Analytical Approach," *Journal of the European Economic Association*, Forthcoming.

Rotemberg, Julio J., "Monopolistic Price Adjustment and Aggregate Output," *Review of Economic Studies*, 1982, *49* (4), 517–531.

Wang, Neng, "Caballero Meets Bewley: The Permanent-Income Hypothesis in General Equilibrium," *American Economic Review*, June 2003, *93* (3), 927–936.

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

Appendix

A Proof of Proposition 1

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

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

s.t.

$$c_t^s(i)+q_t b_{t+1}^s(i) \quad = \quad w_t\ell_t^s(i)+b_t^s(i)+T_t \tag{52}$$

where $b_s^s(i)=0$ and $w_t=(1-\tau)\widetilde{w}_t$. The optimal labor supply decisions of houshold i is given by:

$$\ell_t^s(i)=\rho\ln w_t-\rho\gamma c_t^s(i)+\xi_t^s(i) \tag{53}$$

and the Euler equation is given by:

$$e^{-\gamma c_t^s(i)}=\beta R_t\mathbb{E}_t e^{-\gamma c_{t+1}^s(i)} \tag{54}$$

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

$$c_t^s(i)=\mathcal{C}_t+\mu_t x_t^s(i) \tag{55}$$

where $x_t^s(i)=b_{t-1}^s(i)+w_t\left(\xi_t^s(i)-\bar{\xi}\right)$ denotes "*virtual cash-on-hand*". Notice that $x_{t+1}^s(i)$ is normally distributed and so given the guess (55), $c_{t+1}^s(i)$ is also normally distributed with mean:

$$\mathbb{E}_t c_{t+1}^s(i)=\mathcal{C}_{t+1}+\mu_{t+1}\frac{R_t}{\vartheta}\left[x_t^s(i)+w_t\left(\rho\ln w_t+\bar{\xi}\right)+T_t-\left(1+\rho\gamma w_t\right)c_t^s(i)\right]$$

and variance:

$$\mathbb{V}_t\left(c_{t+1}^s(i)\right)=\mu_{t+1}^2 w_{t+1}^2\sigma_{t+1}^2$$

Taking logs of (54) and using the two expressions above:

$$\begin{aligned}
c_t^s(i) &= -\frac{1}{\gamma}\ln\beta R_t-\frac{1}{\gamma}\ln\mathbb{E}_t e^{-\gamma c_{t+1}^s(i)} \\
&= -\frac{1}{\gamma}\ln\beta R_t+\mathbb{E}_t c_{t+1}^s(i)-\frac{\gamma}{2}\mathbb{V}_t\left(c_{t+1}^s(i)\right) \\
&= -\frac{1}{\gamma}\ln\beta R_t+\mathcal{C}_{t+1}+\mu_{t+1}\frac{R_t}{\vartheta}\left[x_t^s(i)+w_t\left(\rho\ln w_t+\bar{\xi}\right)+T_t-\left(1+\rho\gamma w_t\right)c_t^s(i)\right]-\frac{\gamma\mu_{t+1}^2 w_{t+1}^2\sigma_{t+1}^2}{2}
\end{aligned}$$

Combining the $c_t^s(i)$ terms and using (55), the above can be rewritten as:

$$\mu_{t+1}\frac{R_t}{\vartheta}\left[(1+\rho\gamma w_t)+\frac{\vartheta}{R_t}\mu_{t+1}^{-1}\right]\{\mathcal{C}_t+\mu_t x_t^s(i)\} = -\frac{1}{\gamma}\ln\beta R_t+\mathcal{C}_{t+1}+\mu_{t+1}\frac{R_t}{\vartheta}\left[x_t^s(i)+w_t\left(\rho\ln w_t+\bar{\xi}\right)+T_t\right] - \frac{\gamma\mu_{t+1}^2 w_{t+1}^2\sigma_{t+1}^2}{2} \tag{56}$$

Matching coefficients:

$$\mathcal{C}_t = -\frac{\vartheta\mu_t}{\mu_{t+1}R_t}\frac{1}{\gamma}\ln\beta R_t+\frac{\vartheta\mu_t}{\mu_{t+1}R_t}\mathcal{C}_{t+1}+\mu_t\left[w_t\left(\rho\ln w_t+\bar{\xi}\right)+T_t\right]-\frac{\vartheta}{R_t}\frac{\mu_t}{\mu_{t+1}}\frac{\gamma\mu_{t+1}^2 w_{t+1}^2\sigma_{t+1}^2}{2} \tag{57}$$

$$\mu_t^{-1} = (1+\rho\gamma w_t)+\frac{\vartheta}{R_t}\mu_{t+1}^{-1} \tag{58}$$

Notice that (58) is the same as (12) in the main text. Next, in equilibrium, aggregate hours worked is given by:

$$\ell_t = \rho\ln w_t-\gamma\rho\mathcal{C}_t+\bar{\xi}$$

and hence aggregate income is:

$$y_t = w_t\ell_t+T_t = w_t\rho\ln w_t-\gamma\rho w_t\mathcal{C}_t+w_t\bar{\xi}+T_t$$

Using this in (57) and the fact that $\mathcal{C}_t = y_t$ yields equation (11) in the main text.

B Planner's Objective function

B.1 Consumption is normally distributed within cohort

Given the consumption function (9) and the normality of shocks, to 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 $\sigma_a^2(s+1,s) = \left(\frac{R_s}{\theta}\right)^2\left[1-(1+\gamma\rho w_s)\mu_s\right]^2 w_s^2\sigma_s^2$. 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

$$\sigma_c^2(t,s) = \mu_t^2\sigma_a^2(t,s)+\mu_t^2 w_t^2\sigma_t^2 \tag{59}$$

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

$$\sigma_a^2(t,s) = \frac{R_{t-1}^2}{\theta^2}\left[1-(1+\gamma\rho w_{t-1})\mu_{t-1}\right]^2\left[\sigma_a^2(t-1,s)+w_{t-1}^2\sigma_{t-1}^2\right] \tag{60}$$

B.2 Objective function of planner

Substituting labor supply (10) 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:

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

where we have used labor supply (10), consumption function (9) and the fact that in equilibrium $\mathcal{C}_t = y_t$. We assume that the planner puts a weight of 1 on every individual alive at date 0 and $(\beta\vartheta)^t$ on individuals' lifetime welfare who will be born at date $t > 0$. Then the social welfare function can be written as:

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

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

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

where Σ_t is defined as:

$$\Sigma_t = (1-\vartheta)\sum_{s=-\infty}^{t} \vartheta^{t-s} \int e^{-\gamma\mu_t x_t^s(i)} di \tag{61}$$

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

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

B.2.1 Derivation of Σ_t recursion

Write (61) as:

$$\begin{aligned}
\Sigma_t &= (1-\vartheta)\int e^{-\gamma\mu_t x_t^t(i)} di + (1-\vartheta)\sum_{s=-\infty}^{t-1} \vartheta^{t-s}\int e^{-\gamma\mu_t x_t^s(i)} di \\
&= (1-\vartheta)\, e^{\frac{\gamma^2\mu_t^2 w_t^2\sigma_t^2}{2}} + (1-\vartheta)\sum_{s=-\infty}^{t-1} \vartheta^{t-s}\int e^{-\gamma\mu_t\left\{a_t^s(i) + w_t\left(\xi_t^s(i)-\bar{\xi}\right)\right\}} di \\
&= (1-\vartheta)\, e^{\frac{\gamma^2\mu_t^2 w_t^2\sigma_t^2}{2}} + (1-\vartheta)\sum_{s=-\infty}^{t-1} \vartheta^{t-s}\int e^{-\gamma\mu_t a_t^s(i)} di \int e^{-\gamma\mu_t w_t\left(\xi_t^s(i)-\bar{\xi}\right)} di \\
&= e^{\frac{\gamma^2\mu_t^2 w_t^2\sigma_t^2}{2}}\left[(1-\vartheta) + (1-\vartheta)\,\vartheta\sum_{s=-\infty}^{t-1} \vartheta^{t-1-s}\int e^{-\gamma\mu_t \frac{R_{t-1}}{\theta}\left[\mu_{t-1}^{-1} - (1+\gamma\rho w_{t-1})\right]\mu_{t-1} x_{t-1}^s(i)} di\right]
\end{aligned}$$

where we have used the fact that $a_t^s(i) = \frac{R_{t-1}}{\theta}\left[1-(1+\gamma\rho w_{t-1})\mu_{t-1}\right]x_{t-1}^s(i)$. Also, at all dates $t > 0$, we know that $\mu_{t-1}^{-1} - (1 + \gamma\rho w_{t-1}) = \frac{\theta}{R_{t-1}}\mu_t^{-1}$. Using this, we have for all $t > 0$:

$$\ln \Sigma_t = \frac{\gamma^2 \mu_t^2 w_t^2 \sigma_t^2}{2} + \ln\left(1 - \vartheta + \vartheta \Sigma_{t-1}\right) \tag{62}$$

which is the same as (22) in the main text. Next, imposing steady state up till date 0, we know that $\mu^{-1} - (1 + \gamma\rho w) = \widetilde{\beta}\left(\mu_0^e\right)^{-1}$ where μ_0^e denotes the date -1 expectation of μ_0. Then we can write Σ_0 as:

$$\Sigma_0 \quad = \quad e^{\frac{\gamma^2 \mu_0^2 w_0^2 \sigma_0^2}{2}} \left[(1-\vartheta) + (1-\vartheta)\,\vartheta \sum_{s=-\infty}^{-1} \vartheta^{-1-s} \int e^{-\gamma\left(\frac{\mu_0}{\mu_0^e}\right)\mu x_{-1}^s(i)} di\right] \tag{63}$$

and the variance of consumption for cohorts born at date $s < 0$ is $\gamma^2\mu^2\sigma_x^2(-1, s) = -s\Lambda$. Therefore:

$$\Sigma_0 = e^{\frac{\gamma^2 \mu_0^2 w_0^2 \sigma_0^2}{2}} \left[\frac{1-\vartheta}{1 - \vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mu_0^e}\right)^2}}\right]$$

We also know that in steady state $(\Sigma_{-1} = \Sigma)$, $[1 - \vartheta + \vartheta\Sigma_{-1}]\left(1 - \vartheta e^{\frac{\Lambda}{2}}\right) = (1 - \vartheta)$. Plugging this in the previous expression yields (23) in the main text:

$$\ln \Sigma_0 = \frac{\gamma^2 \mu_0^2 w_0^2 \sigma_0^2}{2} + \ln\left[1 - \vartheta + \vartheta\Sigma_{-1}\right] + \ln\left(\frac{1 - \vartheta e^{\frac{\Lambda}{2}}}{1 - \vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mu_0^e}\right)^2}}\right)$$

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

Lemma 2. *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.:

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

Rearranging:

$$f(\Lambda) \equiv \frac{\Lambda}{\left(1 - \beta\theta e^{\frac{\Lambda}{2}}\right)^2} = \frac{\sigma^2}{\rho^2}\left(\frac{\gamma\rho w}{1 + \gamma\rho w}\right)^2 \tag{64}$$

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 (64) must satisfy $\Lambda \leq \overline{\Lambda} < \Lambda^* < 1$. By construction, for any $\Lambda < \Lambda^*$, $\widetilde{\beta} = \beta\theta e^{\frac{\Lambda}{2}} < 1$. □

Lemma 3. *For $\sigma < [0, \overline{\sigma})$, 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:

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

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

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

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

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

which is the same as $\overline{\sigma}$ 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 $\sigma \in [0, \overline{\sigma})$, we have $\vartheta e^{\frac{\Lambda}{2}} < 1$. □

Corollary 1. *The following is true:*

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

Proof.

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

□

D First-order condition of the planning problem

The planning problem can be written as:

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

s.t.

$$
\begin{aligned}
\gamma y_t &= \gamma y_{t+1} - \ln \beta\theta + \ln \mu_{t+1} + \ln\left[\mu_t^{-1} - (1+\gamma\rho w_t)\right] - \frac{\gamma^2 \mu_{t+1}^2 w^2 \sigma^2 e^{2\phi(y_{t+1}-y)}}{2} \\
(\Pi_t - 1)\Pi_t &= \frac{\lambda}{\Phi(\lambda-1)}\left[1 - \frac{z_t}{(1-\tau)\lambda w_t}\right] + \zeta_t^{-1}(\Pi_{t+1}-1)\Pi_{t+1} \\
\ln \Sigma_t &= \frac{\gamma^2 \mu_t^2 w^2 \sigma^2 e^{2\phi(y_t-y)}}{2} + \ln\left[1-\vartheta+\vartheta\Sigma_{t-1}\right] + \mathbb{I}(t=0)\ln\left(\frac{1-\vartheta e^{\frac{\Lambda}{2}}}{1-\vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mu_0^e}\right)^2}}\right) \\
y_t &= z_t \frac{\rho \ln w_t + \bar{\xi}}{1+\gamma\rho z_t + \Delta_t}
\end{aligned}
$$

where $\zeta_t = \frac{1}{\vartheta}\left[\mu_t^{-1} - (1+\gamma\rho w_t)\right]\mu_{t+1}\left(\frac{y_{t+1}z_t w_t}{y_t z_{t+1} w_{t+1}}\right)$.

The problem can be written as a Lagrangian:

$$
\begin{aligned}
\mathcal{L} \;=\; & \sum_{t=0}^{\infty}\beta^t\left\{-\frac{1}{\gamma}(1+\gamma\rho w_t)e^{-\gamma y_t}\Sigma_t\right\} \\
& + \sum_{t=0}^{\infty}\beta^t M_{1,t}\left\{\gamma y_{t+1} - \ln\beta\theta + \ln\mu_{t+1} + \ln\left[\mu_t^{-1} - (1+\gamma\rho w_t)\right] - \frac{\gamma^2\mu_{t+1}^2 w^2\sigma^2 e^{2\phi(y_{t+1}-y)}}{2} - \gamma y_t\right\} \\
& + \sum_{t=0}^{\infty}\beta^t M_{2,t}\left\{\frac{\lambda}{\Phi(\lambda-1)}\left[1-\frac{z_t}{(1-\tau)\lambda w_t}\right] + \zeta_t(\Pi_{t+1}-1)\Pi_{t+1} - (\Pi_t-1)\Pi_t\right\} \\
& + M_{3,0}\left\{\frac{\gamma^2\mu_0^2 w_0^2\sigma_0^2}{2} + \ln\left[1-\vartheta+\vartheta\Sigma_{-1}\right] + \ln\left(\frac{1-\vartheta e^{\frac{\Lambda}{2}}}{1-\vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mu_0^e}\right)^2}}\right) - \ln\Sigma_0\right\} \\
& + \sum_{t=1}^{\infty}\beta^t M_{3,t}\left\{\frac{\gamma^2\mu_t^2 w^2\sigma^2 e^{2\phi(y_t-y)}}{2} + \ln\left[1-\vartheta+\vartheta\Sigma_{t-1}\right] - \ln\Sigma_t\right\} \\
& + \sum_{t=0}^{\infty}\beta^t M_{4,t}\left\{y_t - z_t\frac{\rho\ln w_t + \bar{\xi}}{1+\gamma\rho z_t+\Delta_t}\right\}
\end{aligned}
$$

The optimal decisions satisfy:
FOC wrt w_t

$$
\begin{aligned}
& \mathbb{U}_t\frac{\gamma\rho w_t}{1+\gamma\rho w_t} + \beta^{-1}M_{2,t-1}w_t\frac{d\zeta_{t-1}}{dw_t}(\Pi_t-1)\Pi_t - M_{1,t}\frac{\gamma\rho w_t}{\mu_t^{-1}-(1+\gamma\rho w_t)} \\
& + M_{2,t}\left\{\frac{\lambda}{\Phi(\lambda-1)}\frac{z_t}{(1-\tau)\lambda w_t} + w_t\frac{d\zeta_t}{dw_t}(\Pi_{t+1}-1)\Pi_{t+1}\right\} - M_{4,t}\frac{\rho z_t}{1+\gamma\rho z_t+\Delta_t} = 0
\end{aligned}
\tag{65}
$$

FOC y_t

$$
\begin{aligned}
& -\gamma\mathbb{U}_t - \gamma M_{1,t} + \beta^{-1}M_{2,t-1}\frac{d\zeta_{t-1}}{dy_t}(\Pi_t-1)\Pi_t + M_{2,t}\frac{d\zeta_t}{dy_t}(\Pi_{t+1}-1)\Pi_{t+1} \\
& + \beta^{-1}M_{1,t-1}\left\{\gamma - \phi\gamma^2\mu_{t+1}^2 w^2\sigma^2 e^{2\phi(y_t-y)}\right\} + M_{3,t}\phi\gamma^2\mu_t^2 w^2\sigma^2 e^{2\phi(y_t-y)} + M_{4,t} = 0
\end{aligned}
\tag{66}
$$

FOC μ_t

$$-M_{1,t}\frac{\mu_t^{-1}}{\mu_t^{-1}-1-\gamma\rho w_t}+\beta^{-1}M_{1,t-1}\left[1-\gamma^2\sigma^2w^2\mu_t^2e^{2\phi(y_t-y)}\right]+\beta^{-1}M_{2,t-1}\mu_t\frac{d\zeta_{t-1}}{d\mu_t}\left(\Pi_t-1\right)\Pi_t$$
$$+M_{2,t}\mu_t\frac{d\zeta_t}{d\mu_t}\left(\Pi_{t+1}-1\right)\Pi_{t+1}+M_{3,t}\gamma^2\sigma^2w^2\mu_t^2e^{2\phi(y_t-y)}+\mathbb{I}(t=0)M_{3,0}\frac{\vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mu_0^e}\right)^2}}{1-\vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mu_0^e}\right)^2}}\Lambda\left(\frac{\mu_0}{\mu_0^e}\right)^2 = 0 \tag{67}$$

FOC Σ_t

$$\mathbb{U}_t-M_{3,t}+\beta M_{3,t+1}\frac{\vartheta\Sigma_t}{1-\vartheta+\vartheta\Sigma_t}=0 \tag{68}$$

FOC Π_t

$$\beta^{-1}M_{2,t-1}\zeta_{t-1}\left(2\Pi_t-1\right)-M_{2,t}\left(2\Pi_t-1\right)+M_{4,t}z_t\frac{\rho\ln w_t+\overline{\xi}}{\left[1+\gamma\rho z_t+\Delta_t\right]^2}\Psi\left(\Pi_t-1\right)=0 \tag{69}$$

D.1 Steady state of the optimal plan

Next, we impose steady state and set $\Pi=1$. This yields the τ and the multipliers that are consistent with $\Pi=1$ being optimal in the long-run. Using $\Pi=1$ in (69):

$$\left(\beta^{-1}\widetilde{\beta}-1\right)M_2=0$$

which implies that $M_2=0$ in steady state. Next, from (68) we have $m_3=(1-\widetilde{\beta})^{-1}$ where $m_i=M_i/\mathbb{U}$ for $i=\{1,2,3,4\}$. Then equations (65),(66),(67) can be manipulated to yield:

$$m_1 = \frac{\widetilde{\beta}}{1-\widetilde{\beta}}\left[\frac{\Lambda}{1-\beta^{-1}\widetilde{\beta}\left(1-\Lambda\right)}\right] \tag{70}$$

$$\frac{w-1}{1+\gamma\rho w} = \Omega \tag{71}$$

$$m_4 = \gamma\frac{\left(1-\beta^{-1}\widetilde{\beta}\right)\left(1-\Lambda\right)}{1-\beta^{-1}\widetilde{\beta}\left(1-\Lambda\right)}\left(1+\Omega\right) \tag{72}$$

where $\Omega=\frac{\Theta-(1-\Lambda)}{(1-\Lambda)\left(1-\widetilde{\beta}\right)}$ and $\Theta=1-\frac{\Lambda\phi}{\gamma}$. Finally, using (71) and since $\Pi=1$ in (25) implies that $\frac{1}{(1-\tau)\lambda w}=1$, we can write:

$$w=\frac{1+\Omega}{1-\gamma\rho\Omega}\qquad\text{and}\qquad\tau=\frac{\lambda-1}{\lambda}+\frac{1+\gamma\rho}{\lambda}\frac{\Omega}{\Omega+1}$$

which are the same expressions in (27) in the main text.

E Linearized first order conditions

Linearizing the first-order conditions from the planners problem yields the following:
FOC w

$$
\begin{aligned}
-\gamma(1+\Omega)\widehat{y}_t + (1+\Omega)\frac{\widehat{\Sigma}_t}{\Sigma} - \left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)(1+\Omega)\widehat{m}_{1,t} - \left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)^2\left(\frac{\gamma\rho}{1+\gamma\rho}\right)(1+\Omega)^2 m_1\frac{\widehat{w}_t}{w} & \\
-\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}\right)(1+\Omega)m_1\widehat{\mu}_t + \kappa\frac{1+\gamma\rho}{\gamma\rho}\widehat{m}_{2,t} - \frac{\widehat{m}_{4,t}}{\gamma} + \frac{m_4}{\gamma}\frac{\widehat{w}_t}{w} - \frac{m_4}{\gamma}\frac{1}{1+\gamma\rho}\widehat{z}_t &= 0
\end{aligned} \tag{73}
$$

FOC y

$$
\begin{aligned}
-\frac{\gamma\rho(1+\Omega)}{1+\gamma\rho}\frac{\widehat{w}_t}{w} + \gamma\left[1+2\frac{(1-\Theta)^2}{\Lambda}\left(m_3-\frac{m_1}{\beta}\right)\right]\widehat{y}_t - \frac{\widehat{\Sigma}_t}{\Sigma} - \widehat{m}_{1,t} + \frac{\Theta}{\beta}\widehat{m}_{1,t-1} & \\
+2(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)\widehat{\mu}_t + (1-\Theta)\widehat{m}_{3,t} + \frac{\widehat{m}_{4,t}}{\gamma} &= 0
\end{aligned} \tag{74}
$$

FOC π

$$
-\widehat{m}_{2,t} + \frac{1}{\beta R}\widehat{m}_{2,t-1} + \frac{y}{1+\gamma\rho}\Phi m_4\pi_t = 0 \tag{75}
$$

FOC μ (for dates $t \geq 1$)

$$
\begin{aligned}
-\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}\right)\frac{\gamma\rho(1+\Omega)}{1+\gamma\rho}m_1\frac{\widehat{w}_t}{w} + \left[2\Lambda\left(m_3-\frac{m_1}{\beta}\right) - \frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}m_1\right]\widehat{\mu}_t + \Lambda\widehat{m}_{3,t} & \\
+2\gamma(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)\widehat{y}_t - \frac{1}{\widetilde{\beta}}\left(\widehat{m}_{1,t} - \frac{\widetilde{\beta}}{\beta}(1-\Lambda)\widehat{m}_{1,t-1}\right) &= 0
\end{aligned} \tag{76}
$$

For $t=0$

$$
\begin{aligned}
-\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}\right)(1+\Omega)\frac{\gamma\rho}{1+\gamma\rho}m_1\frac{\widehat{w}_0}{w} + \left[2\Lambda\left(m_3-\frac{m_1}{\beta}\right) - \frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}m_1\right]\widehat{\mu}_0 & \\
+2\gamma(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)\widehat{y}_0 + \frac{\vartheta\Lambda\Sigma}{1-\vartheta}(m_3+\widehat{m}_{3,0}) & \\
+\Lambda\widehat{m}_{3,0} + m_3\left(2+\frac{\Lambda(1-\vartheta+\vartheta\Sigma)}{1-\vartheta}\right)\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\widehat{\mu}_0 - \frac{1}{\widetilde{\beta}}\left(\widehat{m}_{1,0} - \frac{\widetilde{\beta}}{\beta}(1-\Lambda)\widehat{m}_{1,-1}\right) &= 0
\end{aligned} \tag{77}
$$

FOC Σ

$$
\frac{\gamma\rho w}{1+\gamma\rho w}\frac{\widehat{w}_t}{w} - \gamma\widehat{y}_t - \widehat{m}_{3,t} + \widetilde{\beta}\widehat{m}_{3,t+1} + \frac{1-\beta^{-1}\widetilde{\beta}^2}{1-\widetilde{\beta}}\frac{\widehat{\Sigma}_t}{\Sigma} = 0 \tag{78}
$$

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

E.1 Deriving the target criterion

Combine the FOC for $\widehat{w}_t$ (73) and for $\widehat{y}_t$ (74):

$$\begin{aligned}
&\gamma\Omega\widehat{y}_t + 2\gamma\frac{(1-\Theta)^2}{\Lambda}\left(m_3 - \frac{m_1}{\beta}\right)\widehat{y}_t + \Omega\frac{\widehat{\Sigma}_t}{\Sigma} - \left[1 + \left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)^2 (1+\Omega)\, m_1\right]\frac{\gamma\rho\,(1+\Omega)}{1+\gamma\rho}\frac{\widehat{w}_t}{w} + \frac{m_4}{\gamma}\frac{\widehat{w}_t}{w} \\
&+ \left[2\,(1-\Theta)\left(m_3 - \frac{m_1}{\beta}\right) - \left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}\right)(1+\Omega)\, m_1\right]\widehat{\mu}_t + \kappa\frac{1+\gamma\rho}{\gamma\rho}\widehat{m}_{2,t} + (1-\Theta)\,\widehat{m}_{3,t} - \frac{m_4}{\gamma}\frac{1}{1+\gamma\rho}\widehat{z}_t \\
&- \frac{1+\Omega\left(1-\widetilde{\beta}\right)}{\widetilde{\beta}}\left[\widehat{m}_{1,t} - \frac{\widetilde{\beta}}{\beta}(1-\Lambda)\,\widehat{m}_{1,t-1}\right] = 0
\end{aligned} \tag{79}$$

Combine with (76):

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

Next, use the GDP definition (21) to substitute out for $\frac{\widehat{w}_t}{w}$:

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

Substitute out for $\widehat{\mu}_t$ using $\widehat{\mu}_t = \Gamma_t - \gamma\left[1 + \left(1-\widetilde{\beta}\right)\Omega\right]\widehat{y}_t$ and using the definitions of m_1, m_3 and m_4, the above can be written as:

$$\begin{aligned}
&\frac{\Lambda}{1-\beta^{-1}\widetilde{\beta}\left(1-\Lambda\right)}\gamma\Omega\left[\left(1-\widetilde{\beta}\right)\Omega + 1 + 2\left(1-\beta^{-1}\widetilde{\beta}\right)\frac{\left(1-\widetilde{\beta}\right)\Omega}{\Lambda^2}\right]\widehat{y}_t - \gamma\Omega\widehat{y}_t \\
&+\frac{\left(1-\beta^{-1}\widetilde{\beta}\right)\left(1-\Lambda\right)}{1-\beta^{-1}\widetilde{\beta}\left(1-\Lambda\right)}\frac{1}{\rho}\left(\Omega+1\right)\left(\widehat{y}_t - \frac{\rho+y}{1+\gamma\rho}\widehat{z}_t\right) - \Omega\frac{\Lambda + 2\left(1-\beta^{-1}\widetilde{\beta}\right)}{1-\beta^{-1}\widetilde{\beta}\left(1-\Lambda\right)}\Gamma_t \\
&+\kappa\frac{1+\gamma\rho}{\gamma\rho}\widehat{m}_{2,t} + \Omega\frac{\widehat{\Sigma}_t}{\Sigma} - \left(1-\widetilde{\beta}\right)\Omega\widehat{m}_{3,t} \quad = \quad 0
\end{aligned} \tag{80}$$

Guess that:

$$\widehat{m}_{3,t} = \frac{1}{1-\widetilde{\beta}}\frac{\widehat{\Sigma}_t}{\Sigma} + \gamma\Omega\widehat{y}_t + a_z\widehat{z}_t \tag{81}$$

and use this in (78) with $\frac{\widehat{w}_t}{w}$ substituted out using the definition of GDP:

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

using the fact that $\widehat{z}_{t+1} = \varrho_z\widehat{z}_t$. Using the date $t+1$ Σ recursion (45):

$$\Lambda\Gamma_{t+1} = \frac{1-\widetilde{\beta}}{\widetilde{\beta}}\left[\gamma\left(1+\Omega\right)\frac{y}{1+\gamma\rho} + a_z\left(1-\widetilde{\beta}\varrho_z\right)\right]\widehat{z}_t$$

Using the fact that $\Gamma_t = \frac{\gamma y}{1+\gamma\rho}\frac{\left(1-\widetilde{\beta}\right)\left(1+\Omega\right)}{1-\widetilde{\beta}\varrho_z\left(1-\Lambda\right)}\widehat{z}_t$, we have:

$$\Lambda\varrho_z\frac{\gamma y}{1+\gamma\rho}\frac{\left(1-\widetilde{\beta}\right)\left(1+\Omega\right)}{1-\widetilde{\beta}\varrho_z\left(1-\Lambda\right)}\widehat{z}_t = \frac{1-\widetilde{\beta}}{\widetilde{\beta}}\left[\gamma\left(1+\Omega\right)\frac{y}{1+\gamma\rho} + a_z\left(1-\widetilde{\beta}\varrho_z\right)\right]\widehat{z}_t$$

which implies that a_z must satisfy:

$$a_z = -\frac{\gamma\left(1+\Omega\right)}{1-\widetilde{\beta}\rho_z\left(1-\Lambda\right)}\frac{y}{1+\gamma\rho} \tag{82}$$

Using $\widehat{m}_{3,t} = \frac{1}{1-\widetilde{\beta}}\frac{\widehat{\Sigma}_t}{\Sigma} + \gamma\Omega\widehat{y}_t + a_z\widehat{z}_t$ in (80):

$$\alpha\left(\Omega\right)\left(\widehat{y}_t - \delta\left(\Omega\right)\widehat{y}_t^n\right) = \eta_t \qquad \text{for} \qquad t > 0 \tag{83}$$

where we have defined $\eta_t = -\frac{\kappa\left(1+\gamma\rho\right)}{m_4}\widehat{m}_{2,t}$ where

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

and

$$\delta(\Omega) = \alpha(\Omega)^{-1}\left\{1 + \frac{\gamma\left(1-\widetilde{\beta}\right)\Omega}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{\rho y}{\rho + y}\left[\frac{1+\Lambda}{1-\Lambda}\right]\right\}$$

Similarly, for the date 0 target criterion, combine the date 0 version of (79) with (77), (81) and the date 1 Σ recursion (45) to get:

$$\alpha_0(\Omega)\left(\widehat{y}_0 - \delta_0(\Omega)\widehat{y}_0^n - \chi(\Omega)\right) \quad = \quad \eta_0 \tag{84}$$

where

$$\alpha_0(\Omega) = \alpha(\Omega) + \frac{\rho}{m_4}\frac{\gamma^2\left[1+\left(1-\widetilde{\beta}\right)\Omega\right]^2}{1-\widetilde{\beta}}\left[2+\Lambda+2\frac{\Lambda\vartheta\Sigma}{1-\vartheta}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta} \tag{85}$$

$$\delta_0(\Omega) \quad = \quad \alpha_0(\Omega)^{-1}\left[\alpha(\Omega)\delta(\Omega) + \gamma^2\frac{(\Omega+1)}{m_4}\left[1+2\Lambda+2\frac{\Lambda\vartheta\Sigma}{1-\vartheta}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\frac{1+\left(1-\widetilde{\beta}\right)\Omega}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{\rho y}{y+\rho}\right] \tag{86}$$

$$\chi(\Omega) = \alpha_0(\Omega)^{-1}\frac{\gamma\rho}{m_4}\left[\frac{1+(1-\widetilde{\beta})\Omega}{1-\widetilde{\beta}}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}$$

In summary, using (84)-(83) and the target criterion can be written as:

$$\eta_t = \begin{cases} \alpha_0(\Omega)\left(\widehat{y}_t - \delta_0(\Omega)\widehat{y}_t^n - \chi(\Omega)\right) & \text{for } t = 0 \\ \alpha(\Omega)\left(\widehat{y}_t - \delta(\Omega)\widehat{y}_t^n\right) & \text{for } t \geq 1 \end{cases}$$

This is the same as (46) in the main text. Next, multiplying (75) by $\frac{\kappa(1+\gamma\rho)}{m_4}$ yields:

$$\eta_t = \frac{1}{\beta R}\eta_{t-1} - \frac{\lambda}{\lambda - 1}y\pi_t$$

which is the same as equation (47) in the main text. This concludes the derivation of the expressions in Proposition 4.

E.1.1 Derivation of expressions in Proposition 3

Define the scalars $\alpha = \alpha_0(0)$ and $\delta = \delta_0(0)$ and $\chi = \chi(0)$:

$$\begin{aligned}
\alpha \equiv \alpha_0(0) \quad &= \quad 1 + \frac{\rho}{m_4}\frac{\gamma^2}{1-\widetilde{\beta}}\left[2+\Lambda+2\frac{\Lambda\vartheta\Sigma}{1-\vartheta}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta} > 1 \\
\delta \equiv \delta_0(0) \quad &= \quad \alpha^{-1}\left[1 + \frac{\rho}{m_4}\frac{\gamma^2}{1-\widetilde{\beta}}\left[1+2\Lambda+2\frac{\Lambda\vartheta\Sigma}{1-\vartheta}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\frac{1-\widetilde{\beta}}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{1}{1+\rho}\right] \\
\chi = \chi(0) \quad &= \quad \alpha^{-1}\frac{\gamma\rho}{m_4}\frac{\vartheta\Lambda\Sigma}{1-\vartheta} > 0
\end{aligned}$$

Since $\frac{1-\widetilde{\beta}}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{1}{1+\rho} < 1$, it is clear that $\delta < 1$. It is also clear by inspection that $\alpha(0) = \delta(0) = 1$. Thus, the optimal target criterion when $\Omega = 0$ can be written as:

$$\eta_t = \begin{cases} \alpha\left(\widehat{y}_t - \delta \widehat{y}_t^n - \chi\right) & \text{for } t = 0 \\ \widehat{y}_t - \widehat{y}_t^n & \text{for } t \geq 1 \end{cases}$$

which is the same as equation (38). Finally, when $\Omega = 0$, $y = 1$ and so (47) becomes (39).

E.1.2 Optimal target criterion in RANK

Recall that in RANK we have $\Omega = 0$ and $\Lambda = 0$. In this case, it is clear from inspection that the coefficients further simplify to $\alpha = \delta = 1$, $\chi = 0$ and the target criterion can be written as:

$$\eta_t = \widehat{y}_t - \widehat{y}_t^n \qquad \text{for } t \geq 0$$

which is the same as (29) in the main text

E.2 Properties of coefficients $\alpha_0(\Omega), \delta_0(\Omega), \alpha(\Omega), \delta(\Omega)$

Claim: $\alpha(\Omega) > 1$

Proof.

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

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

$$\alpha(\Omega) > 1 + \frac{\rho\gamma\Omega}{1+\Omega}\frac{1+\Lambda}{(1-\Lambda)^2} > 1$$

□

Claim: If risk is countercyclical $\Theta > 1$, then $\delta(\Omega) < 1$

Proof. Countercyclical risk or $\Theta > 1$ implies that $\Omega > \frac{\Lambda}{(1-\widetilde{\beta})(1-\Lambda)}$ □

Proof. Notice that $\delta(\Omega)$ can be written as:

$$\delta(\Omega) = \frac{1 + \Omega + (\Omega + \Omega^2)\frac{\gamma\left(1-\widetilde{\beta}\right)}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{\rho y}{\rho + y}\left[\frac{1+\Lambda}{1-\Lambda}\right]}{1 + (1-\rho\gamma)\Omega + \rho\gamma\Omega^2\left(\frac{2}{\Lambda(1-\Lambda)} - 1\right)\left(1-\widetilde{\beta}\right)}$$

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

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

This expression can be simplified to yield:

$$1+\frac{\left(1-\widetilde{\beta}\right)}{1-\widetilde{\beta}\rho_z\left(1-\Lambda\right)}\frac{y}{\rho+y}\left(\frac{1+\Lambda}{1-\Lambda}\right) < \Omega\left(1-\widetilde{\beta}\right)\left[\left(\frac{2}{\Lambda\left(1-\Lambda\right)}-1\right)-\frac{1}{1-\widetilde{\beta}\rho_z\left(1-\Lambda\right)}\frac{y}{\rho+y}\left(\frac{1+\Lambda}{1-\Lambda}\right)\right] \tag{87}$$

First, we show that the term in the square brackets on the RHS of (87) is positive, i.e.

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

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:

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

which is true since $\widetilde{\beta} < 1$ and $\frac{2}{2-(1-\Lambda)\Lambda} > 1$ since we know that $0 < \Lambda < 1$ from Appendix XXX. Thus, the term in the square brackets on the RHS of (87) is positive.

Next, to show that (87) holds with countercyclical risk, it suffices to show that it holds for the lowest Ω consistent with non-procyclical risk, i.e. $\Omega = \frac{\Lambda}{(1-\widetilde{\beta})(1-\Lambda)}$. Plug in $\Omega = \frac{\Lambda}{(1-\widetilde{\beta})(1-\Lambda)}$ into (87), i.e:

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

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: If risk is countercyclical $\Theta > 1$, then $\alpha_0(\Omega) > 1$

Proof. Since $\alpha(\Omega) > 1$, it follows from (85) that $\alpha_0(\Omega) > 1$. □

F The model with nominal debt

In this extension, we change the assumption that financial intermediaries trade real claims with the households. Instead, we assume that now these intermediaries trade nominal claims at a nominal price $\frac{\vartheta}{1+i_t}$ which pay a dollar tomorrow. Then the household's budget constraint can be written as:

$$P_t c_t^s\left(i\right)+\frac{\vartheta}{1+i_t}A_{t+1}^s\left(i\right) = P_t w_t \ell_t^s\left(i\right)+A_t^s\left(i\right)+P_t T_t$$

where $A_{t+1}^s(i)$ is the quantity of nominal claims purchased by the household at date t. Dividing through by P_t, we get:

$$c_t^s(i) + \frac{\vartheta}{1+i_t}\Pi_{t+1}a_{t+1}^s(i) = w_t\ell_t^s(i) + a_t^s(i) + T_t$$

where $a_t^s(i) = \frac{A_t^s(i)}{P_t}$ denotes the real value of wealth held by the household at the beginning of date t. Given these definitions, the rest of the model is the same except for the μ_t recursion which can now be written as:

$$\mu_t^{-1} = 1 + \gamma\rho w_t + \frac{\vartheta\Pi_{t+1}}{1+i_t}\mu_{t+1}^{-1}$$

the Σ_t recursion which we now derive. We start with the budget constraint of a households and plug in the expression for labor supply (10) and the consumption function (9):

$$\begin{aligned} a_t^s(i) &= \frac{1+i_{t-1}}{\vartheta\Pi_t}\left[w_{t-1}\ell_{t-1}^s(i) + a_{t-1}^s(i) + T_{t-1} - c_{t-1}^s(i)\right] \\ &= \frac{R_{t-1}}{\theta}\left[1 - (1+\gamma\rho w_{t-1})\mu_{t-1}\right]x_{t-1}^s(i) \end{aligned}$$

In steady state, (12) implies that $\frac{R}{\theta}\left[1-(1+\gamma\rho w)\mu\right] = 1$, so we can write:

$$a_t^s(i) = \frac{R}{\theta}\left[1-(1+\gamma\rho w)\mu\right]x_{t-1}^s(i) = x_{t-1}^s(i)$$

If we were in steady state up till date 0, the distribution of assets which would obtain at date 0 if there was no inflation between $t=-1$ and $t=0$:

$$\tilde{a}_0^s(i) = w\sum_{k=s}^{-1}\left(\xi_k^s(i) - \bar{\xi}\right)$$

and

$$\sigma_{\tilde{a}}^2(s,0) = \sigma^2 w^2(-s)$$

But if there is inflation, then actual $a_0^s(i)$ is:

$$a_0^s(i) = w\sum_{k=s}^{-1}\left(\xi_k^s(i) - \bar{\xi}\right)\Pi_0^{-1}$$

This implies that for cohort born at date $s<0$, $a_0^s(i)$ is normally distributed with mean 0 and variance $-sw^2\sigma^2\Pi_0^{-2}$. Next, we use this information to derive the expression for Σ_0. Using the same Pareto weights as in the baseline model, we have from (63):

$$\begin{aligned} \Sigma_0 &= (1-\vartheta)\sum_{s=-\infty}^{0}\vartheta^{-s}\int e^{-\gamma\mu_0 x_0^s(i)}di \\ &= (1-\vartheta)\sum_{s=-\infty}^{0}\vartheta^{-s}\int e^{-\gamma\mu_0 a_0^s(i)}di\int e^{-\gamma\mu_0 w_0\left(\xi_0^s-\bar{\xi}\right)}di \\ &= (1-\vartheta)\,e^{\frac{1}{2}\gamma^2\mu_0^2 w_0^2\sigma_0^2}\sum_{s=-\infty}^{0}\vartheta^{-s}e^{-\frac{\Lambda}{2}\left(\frac{\mu_0}{\mu\Pi_0}\right)^2 s} = (1-\vartheta)\frac{e^{\frac{1}{2}\gamma^2\mu_0^2 w_0^2\sigma_0^2}}{1-\vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mu\Pi_0}\right)^2}} \end{aligned}$$

where on the last line we have used the fact that for cohorts born at dates $s < 0$, $a_0^s(i) \sim N\left(0, -sw^2\sigma^2\Pi_0^{-2}\right)$. This expression can be manipulated to yield:

$$\ln \Sigma_0 = \frac{1}{2}\gamma^2\mu_0^2 w_0^2\sigma_0^2 + \ln\left[1-\vartheta+\vartheta\Sigma\right] + \underbrace{\ln\left(\frac{1-\vartheta e^{\frac{\Lambda}{2}}}{1-\vartheta e^{\frac{\Lambda}{2}\left(\frac{\mu_0}{\mathbb{E}_{-1}\mu_0}\frac{\mathbb{E}_{-1}\Pi_0}{\Pi_0}\right)^2}}\right)}_{\text{effect of date 0 surprise}}$$

Of course, for dates $t > 0$, the Σ recursion stays the same as in the baseline model since the planner is not able to create any inflation surprise at dates $t \geq 1$. In linearized terms, the Σ recursion can be written as:

$$\frac{\widehat{\Sigma}_t}{\Sigma} = \begin{cases} -\gamma\left(1-\widetilde{\beta}\right)\Omega\widehat{y}_t + \Lambda\Gamma_t + \frac{\vartheta}{\beta R}\frac{\widehat{\Sigma}_{t-1}}{\Sigma} & \text{for } t > 0 \\ -\gamma\left(1-\widetilde{\beta}\right)\Omega\widehat{y}_0 + \Lambda\Gamma_0 + \frac{\vartheta\Lambda\Sigma}{1-\vartheta}\left(\widehat{\mu}_0-\pi_0\right) & \text{for } t = 0 \end{cases}$$

F.1 Optimal monetary policy

The monetary policy problem is very similar to the baseline case. The only difference is the expression for Σ_0 since now the planner can affect the level of consumption inequality by creating an inflation surprise at date 0. In linearized form, the only first order conditions that differ from those in Section E is the date 0 first-order condition with respect to inflation and with respect to μ. In linearized form, the first order conditions for π can now be written as:
For $t = 0$

$$-\widehat{m}_{2,0} + \left(\frac{y}{1+\gamma\rho}\Psi m_4 + m_3\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\right)\pi_0 - \frac{\Lambda\vartheta\Sigma}{1-\vartheta}\widehat{m}_{3,0} - m_3\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\left(2+\Lambda+\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\right)\left(\widehat{\mu}_0-\pi_0\right) - m_3\frac{\vartheta\Lambda\Sigma}{1-\vartheta} = 0 \tag{88}$$

and for $t \geq 1$

$$-\widehat{m}_{2,t} + \frac{1}{\beta R}\widehat{m}_{2,t-1} + \frac{y}{1+\gamma\rho}\Psi m_4\pi_t = 0 \tag{89}$$

and the FOCs for μ can be written as:
For $t = 0$

$$\begin{aligned} -\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}\right)(1+\Omega)\frac{\gamma\rho}{1+\gamma\rho}m_1\frac{\widehat{w}_0}{w} + \left[2\Lambda\left(m_3-\frac{m_1}{\beta}\right) - \frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}m_1\right]\widehat{\mu}_0 \\ +2\gamma\left(1-\Theta\right)\left(m_3-\frac{m_1}{\beta}\right)\widehat{y}_0 + \frac{\vartheta\Lambda\Sigma}{1-\vartheta}\left(m_3+\widehat{m}_{3,0}\right) \\ +\Lambda\widehat{m}_{3,0} + m_3\left(2+\frac{\Lambda\left(1-\vartheta+\vartheta\Sigma\right)}{1-\vartheta}\right)\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\left(\widehat{\mu}_0-\pi_0\right) - \frac{1}{\widetilde{\beta}}\left(\widehat{m}_{1,0} - \frac{\widetilde{\beta}}{\beta}\left(1-\Lambda\right)\widehat{m}_{1,-1}\right) &= 0 \end{aligned} \tag{90}$$

and for $t \geq 1$

$$-\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}\right)(1+\Omega)\frac{\gamma\rho}{1+\gamma\rho}m_1\frac{\widehat{w}_t}{w}+\left[2\Lambda\left(m_3-\frac{m_1}{\beta}\right)-\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}m_1\right]\widehat{\mu}_t \tag{91}$$
$$+2\gamma(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)\widehat{y}_t+\Lambda\widehat{m}_{3,0}-\frac{1}{\widetilde{\beta}}\left(\widehat{m}_{1,t}-\frac{\widetilde{\beta}}{\beta}(1-\Lambda)\widehat{m}_{1,t-1}\right) = 0$$

F.2 Deriving the target criterion

It is clear that the target criterion for dates $t \geq 1$ is unchanged. The only difference is in the date 0 target criterion, which we derive next. In this case with nominal debt, the date 0 version of the combined FOC for w and y (79) can be written as:

$$\gamma\Omega\widehat{y}_0+2\gamma\frac{(1-\Theta)^2}{\Lambda}\left(m_3-\frac{m_1}{\beta}\right)\widehat{y}_0+\Omega\frac{\widehat{\Sigma}_0}{\Sigma}-\left[1+\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}}\right)^2(1+\Omega)m_1\right]\frac{\gamma\rho(1+\Omega)}{1+\gamma\rho}\frac{\widehat{w}_0}{w}+\frac{m_4}{\gamma}\frac{\widehat{w}_0}{w}$$
$$+\left[2(1-\Theta)\left(m_3-\frac{m_1}{\beta}\right)-\left(\frac{1-\widetilde{\beta}}{\widetilde{\beta}^2}\right)(1+\Omega)m_1\right]\widehat{\mu}_0+\kappa\frac{1+\gamma\rho}{\gamma\rho}\widehat{m}_{2,0}+(1-\Theta)\widehat{m}_{3,0}-\frac{m_4}{\gamma}\frac{1}{1+\gamma\rho}\widehat{z}_0$$
$$-\frac{1+\Omega\left(1-\widetilde{\beta}\right)}{\widetilde{\beta}}\left[\widehat{m}_{1,0}-\frac{\widetilde{\beta}}{\beta}(1-\Lambda)\widehat{m}_{1,-1}\right] = 0$$

Multiply (90) by $-\left[1+\Omega\left(1-\widetilde{\beta}\right)\right]$ and add to the equation above to get:

$$\left\{\frac{\Lambda}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}\gamma\Omega\left[\left(1-\widetilde{\beta}\right)\Omega+1+2\left(1-\beta^{-1}\widetilde{\beta}\right)\frac{\left(1-\widetilde{\beta}\right)\Omega}{\Lambda^2}\right]-\gamma\Omega+\frac{m_4}{\rho\gamma}\right\}\widehat{y}_0$$
$$-\frac{m_4}{\rho\gamma}\frac{\rho+y}{1+\gamma\rho}\widehat{z}_0-\Omega\frac{\Lambda+2\left(1-\beta^{-1}\widetilde{\beta}\right)}{1-\beta^{-1}\widetilde{\beta}(1-\Lambda)}\Gamma_0+\kappa\frac{1+\gamma\rho}{\gamma\rho}\widehat{m}_{2,0}+\Omega\frac{\widehat{\Sigma}_0}{\Sigma}-\left(1-\widetilde{\beta}\right)\Omega\widehat{m}_{3,0}$$
$$-\left[1+\left(1-\widetilde{\beta}\right)\Omega\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\widehat{m}_{3,0}-\left[\frac{1+\left(1-\widetilde{\beta}\right)\Omega}{1-\widetilde{\beta}}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}-\left[\frac{1+\left(1-\widetilde{\beta}\right)\Omega}{1-\widetilde{\beta}}\right]\left[2+\frac{\Lambda(1-\vartheta+\vartheta\Sigma)}{1-\vartheta}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\Gamma_0$$
$$+\gamma\frac{\left[1+\left(1-\widetilde{\beta}\right)\Omega\right]^2}{1-\widetilde{\beta}}\left[2+\frac{\Lambda(1-\vartheta+\vartheta\Sigma)}{1-\vartheta}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\widehat{y}_0+\left[\frac{1+\Omega\left(1-\widetilde{\beta}\right)}{1-\widetilde{\beta}}\right]\left[2+\frac{\Lambda(1-\vartheta+\vartheta\Sigma)}{1-\vartheta}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\pi_0 = 0$$

Guess

$$\widehat{m}_{3,0}=\frac{1}{1-\widetilde{\beta}}\frac{\widehat{\Sigma}_0}{\Sigma}+\gamma\Omega\widehat{y}_0-\frac{\gamma(1+\Omega)}{1-\widetilde{\beta}\rho_z(1-\Lambda)}\frac{y}{1+\gamma\rho}\widehat{z}_0$$

Plug this into the expression above:

$$\left\{\frac{\Lambda}{1-\beta^{-1}\widetilde{\beta}\left(1-\Lambda\right)}\gamma\Omega\left[\left(1-\widetilde{\beta}\right)\Omega+1+2\left(1-\beta^{-1}\widetilde{\beta}\right)\frac{\left(1-\widetilde{\beta}\right)\Omega}{\Lambda^2}\right]+\frac{m_4}{\rho\gamma}-\gamma\Omega\left[1+\left(1-\widetilde{\beta}\right)\Omega\right]\right\}\widehat{y}_0$$
$$-m_4\left[\frac{\left(1-\widetilde{\beta}\right)\Omega}{1-\widetilde{\beta}\rho_z\left(1-\Lambda\right)}\frac{y}{y+\rho}\left(\frac{1+\Lambda}{1-\Lambda}\right)+\frac{1}{\rho\gamma}\right]\widehat{y}_0^n+\kappa\frac{1+\gamma\rho}{\gamma\rho}\widehat{m}_{2,0}$$
$$+\gamma\left[1+\left(1-\widetilde{\beta}\right)\Omega\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\left\{2\Omega+\left[\frac{1+\left(1-\widetilde{\beta}\right)\Omega}{1-\widetilde{\beta}}\right]\left[2\left(1+\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\right)+\Lambda\right]\right\}\widehat{y}_0$$
$$-\left[\frac{1+\left(1-\widetilde{\beta}\right)\Omega}{1-\widetilde{\beta}\rho_z\left(1-\Lambda\right)}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\gamma\left(1+\Omega\right)\frac{y}{y+\rho}\left[1+2\left(\Lambda+\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\right)\right]\widehat{y}_0^n$$
$$+\left[\frac{1+\Omega\left(1-\widetilde{\beta}\right)}{1-\widetilde{\beta}}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\left[2\left(1+\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\right)+\Lambda\right]\pi_0-\left[\frac{1+\left(1-\widetilde{\beta}\right)\Omega}{1-\widetilde{\beta}}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\quad=\quad 0$$

This expression can be rearranged to yield:

$$\alpha_0\left(\Omega\right)\left[\widehat{y}_0-\delta_0\left(\Omega\right)\widehat{y}_0^n+\iota\left(\Omega\right)\pi_0-\chi\left(\Omega\right)\right]=\eta_0$$

where

$$\alpha_0\left(\Omega\right)\iota\left(\Omega\right)=\frac{\gamma\rho}{m_4}\left[\frac{1+\Omega\left(1-\widetilde{\beta}\right)}{1-\widetilde{\beta}}\right]\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\left[2\left(1+\frac{\vartheta\Lambda\Sigma}{1-\vartheta}\right)+\Lambda\right]$$

and $\alpha_0\left(\Omega\right),\delta_0\left(\Omega\right)$ and $\chi\left(\Omega\right)$ are as defined in Appendix E.1.

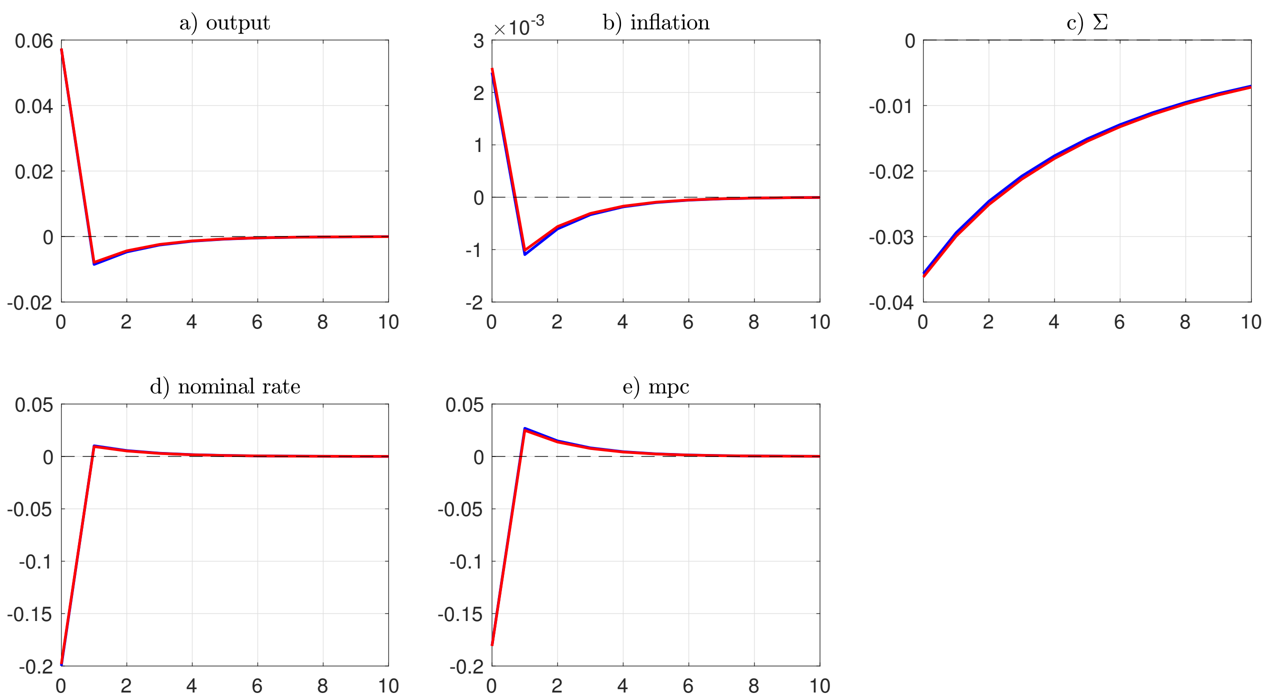

Figure 8. Time inconsistency with nominal debt: Optimal dynamics absent shocks in baseline calibration ($\kappa = 0.01$). Blue line denotes HANK with countercylcical risk and real debt, red line denotes HANK with countercyclical risk and nominal debt. All variables are plotted as percentage deviations from their steady state values.

.