A Behavioral Heterogeneous Agent New Keynesian Model

Oliver Pfäuti* Fabian Seyrich†

First version: January 2022
This version: May 2023
Link to most recent version

Abstract

We analyze how cognitive discounting and household heterogeneity affect the transmission of monetary policy. Under cognitive discounting, households' expectations exhibit an underreaction to news about the aggregate economy, which is consistent with empirical evidence on household expectations. Our model simultaneously accounts for recent empirical findings of the transmission of monetary policy: (i) monetary policy affects consumption largely through indirect effects, (ii) households are unequally exposed to aggregate fluctuations and income risk is countercyclical, (ii) forward guidance is less powerful than contemporaneous monetary policy, (iv) and the economy remains stable at the zero lower bound. In contrast to demand shocks, supply shocks are amplified through both, cognitive discounting and household heterogeneity, such that inflation increases more than twice as strong as when abstracting from cognitive discounting and household heterogeneity.

Keywords: Monetary Policy, Heterogeneous Households, Behavioral Macroeconomics, Forward Guidance, Lower Bound, Inflation, Macroeconomic Stabilization

JEL Codes: E21, E52, E62, E71

*University of Mannheim, pfaeuti.oliver@gmail.com

†Berlin School of Economics, DIW Berlin, and Freie Universität Berlin, fabian.seyrich@gmail.com.
We thank Klaus Adam, George-Marios Angeletos, Adrien Auclert (discussant), Neele Balke, Nicholas Barberis, Christian Bayer, Florin Bilbiie, Eduardo Dávila, Davide Debortoli, Stéphane Dupraz (discussant), Xavier Gabaix, Morteza Ghomi (discussant), Zhen Huo, Diego Känzig, Alexander Kriwoluzky, Max Jager, Timo Reinelt, Ricardo Reis, Hannah Seidl, Alp Simsek, Alejandro Van der Ghote (discussant), Maximilian Weiß (discussant) and numerous seminar and conference participants for helpful comments and suggestions. Oliver Pfäuti gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project C02) and the Stiftung Geld & Währung. Fabian Seyrich gratefully acknowledges financial support by the Leibniz Association through the project "Distributional effects of macroeconomic policies in Europe".

1 Introduction

Recent empirical evidence has led to a rethinking of how monetary policy is transmitted to the economy: (i) monetary policy affects household consumption to a large extent through changing people's incomes rather than directly through changes in the real interest rate. These *indirect effects* tend to amplify the effects of conventional monetary policy on consumption as (ii) the incomes of households that exhibit higher marginal propensities to consume are found to be more exposed to aggregate income fluctuations induced by monetary policy; (iii) announcements of future monetary policy changes, in contrast, have relatively weak effects on current economic activity; and (iv) advanced economies have not experienced large instabilities in times in which the nominal interest rate has been stuck at the lower bound.[1]

In this paper, we propose a new framework that accounts for these four facts *simultaneously*: the behavioral Heterogeneous Agent New Keynesian model—or *behavioral HANK model*, for short. The model features a standard New Keynesian core with nominal rigidities, but we allow for household heterogeneity and bounded rationality in the form of cognitive discounting. The presence of both—household heterogeneity and bounded rationality—is key to account for the four facts jointly. In contrast to existing models, our model accounts for the four facts without having to rely on a specific monetary or fiscal policy.

We first illustrate how cognitive discounting interacts with household heterogeneity under a specific calibration of our model for which we obtain a closed-form solution but that still captures the key features of the model. Households that exhibit higher marginal propensities to consume are more exposed to monetary policy which is crucial to account for the fact that monetary policy is amplified through indirect general equilibrium effects. Under cognitive discounting, households' expectations underreact to aggregate news—consistent with what we document for household survey expectations—which dampens the effects of announced future monetary policy changes and ensures that the model remains stable at the effective lower bound. Second, we then show numerically that all our results carry over to the full model. This holds true, even when households over- or underreact to idiosyncratic shocks or when households are heterogenous in their behavioral biases.

Accounting for these four facts simultaneously has important implications for macroeconomic stabilization. In particular, we uncover a new amplification channel of adverse supply shocks: the unequal exposure of households, their behavioral bias and the interaction of the two lead to a substantial increase in the output gap and inflation. Inflation increases more than twice as strong as when abstracting from these model features. As a consequence of this amplification channel,

[1] See, e.g., Ampudia et al. (2018), Slacalek et al. (2020) and Holm et al. (2021) for the empirical relevance of indirect channels in the transmission of monetary policy, Auclert (2019), Patterson (2023) and Slacalek et al. (2020) for evidence on households' income exposure and their marginal propensities to consume, and see, for example, Del Negro et al. (2015), D'Acunto et al. (2022), and Roth et al. (2021) for empirical evidence on the (in-)effectiveness of monetary policy announcements about its future actions, and Debortoli et al. (2020) and Cochrane (2018) on the stability at the lower bound.

there is a strong trade-off for monetary policy between price stability on the one side and fiscal and distributional consequences on the other side after an inflationary supply shock. If monetary policy wants to fully stabilize inflation, it needs to increase interest rates much more aggressively, which pushes up the government debt level and inequality more strongly.

Our model builds on the recent heterogeneous-agent New Keynesian literature (HANK) which combines the typical Bewley-Huggett-Aiyagari incomplete markets setup with nominal rigidities. Ex-ante identical households face uninsurable idiosyncratic productivity risk, incomplete markets and borrowing constraints. In contrast to that literature, households in our model do not necessarily hold rational expectations. In particular, we allow for *cognitive discounting* of aggregate variables: households anchor their expectations about future macroeconomic variables to the steady state and cognitively discount expected future deviations as in Gabaix (2020). As a result, expectations then underreact to aggregate news, as we show to be the case empirically across all income groups and which is also consistent with findings in D'Acunto et al. (2022) or Roth et al. (2021).[2]

We start by showing that for a specific calibration, the model simplifies such that it can be solved in closed form. In particular, the household block can be represented as if there were two representative households. Yet, the model still shares the key features with our full model, namely unequal exposure of households to aggregate shocks, a precautionary savings motive of households and borrowing constraints as well as cognitive discounting of aggregate shocks.[3] The two *as-if* representative households differ in the following respects: the first group is "unconstrained", in the sense that they participate in financial markets and are on their Euler equation. The second group consists of "hand-to-mouth" households who consume all their disposable income. They exhibit high marginal propensities to consume (MPCs) and their income is more exposed to monetary policy in line with the data. As unconstrained households face a risk of becoming hand-to-mouth, they exhibit a precautionary-savings motive.

Given this specific calibration, the model can then be represented in just three equations exactly like the textbook Representative Agent New Keynesian (RANK) model. The key novelty is a new aggregate IS equation. In contrast to the textbook model, our IS equation features a lower sensitivity of current output to changes in expected future output due to households' cognitive discounting and a stronger sensitivity of current output to changes in the real interest rate as households with higher MPCs are more exposed to monetary policy.

As a result of the lower sensitivity of current output to future expected output, announced policies that increase future output, such as announced future interest rate cuts, are less effective

[2] Angeletos and Lian (2023) show how other forms of bounded rationality or lack of common knowledge can be observationally equivalent. For further evidence on underreaction of expectations or general patterns of inattention, see, e.g., Coibion and Gorodnichenko (2015), Coibion et al. (2022) or Angeletos et al. (2021). Kučinskas and Peters (2022) and Born et al. (2022) show that even when agents overreact to micro news, they underreact to macro news.

[3] Models with a similar household structure are often referred to as TANK (Two Agent New Keynesian) models with type switching or as THANK (Tractable HANK) models (Bilbiie (2021)). We therefore refer to this special calibration of our model as *tractable* behavioral HANK model.

in stimulating current output. After such an announced future interest rate cut, unconstrained households want to consume more already today as they want to smooth their consumption intertemporally. Additionally, their precautionary savings motive decreases as they would be better off in case they become hand-to-mouth in the future because hand-to-mouth households benefit more from the future boom. Cognitive discounting weakens *both* of these channels and thus, explains the lower sensitivity of current output to future expected output. The farther away in the future the announced interest rate cut takes place, the smaller its effect on today's output. Hence, the model does not suffer from the *forward guidance puzzle,* which describes the paradoxical finding in many models that announced future interest-rate changes are at least as effective in stimulating current output than contemporaneous interest-rate changes (Del Negro et al. (2015), McKay et al. (2016)). In addition, our model remains determinate under an interest-rate peg and remains stable at the effective lower bound (ELB).

The second deviation from the textbook IS equation—the stronger sensitivity of current output to changes in the real interest rate—arises because households with higher MPCs are more exposed to monetary policy. An expansionary monetary policy shock increases the income of the hand-to-mouth households more than one-for-one. As these households consume all their disposable income, this leads to a stronger response of aggregate consumption than if all households would be exposed equally to monetary policy. Thus, the model features amplification of conventional monetary policy shocks due to indirect general equilibrium effects. A decomposition into direct and indirect effects shows that indeed the major share of the monetary policy transmission works through indirect effects.

We then relax our specific calibration and show that none of our results depends on it. In particular, we build on a calibration that is standard in the HANK literature extended by cognitive discounting and the unequal exposure of households to monetary policy shocks found in the data. Consequently, the model now features a non-degenerate wealth distribution and can only be solved numerically. We show that the model still accounts for facts (i)-(iv) simultaneously.

That our model simultaneously generates amplification of conventional monetary policy through indirect effects and rules out the forward-guidance puzzle is in stark contrast to rational models. Rational HANK models that generate amplification through indirect effects exacerbate the forward-guidance puzzle. Rational models that resolve the forward-guidance puzzle, on the other hand, cannot simultaneously generate amplification of monetary policy through indirect effects (see Werning (2015), Acharya and Dogra (2020), and Bilbiie (2021)).

We extend our model in several ways. First, we consider an extension in which households are heterogeneous with respect to their cognitive discounting. We find in the data that the degree of rationality is slightly positively correlated with the income of households. Introducing this into our model, we find that this extension has only minor quantitative impacts on our results, while the model continues to account for facts (i) - (iv) simultaneously. This even applies to a version in which a subgroup of households is fully rational. Second, we allow for bounded rationality also with respect to households' idiosyncratic risk. Recent empirical findings by Kučinskas and Peters (2022) and Born et al. (2022) show that even though agents' expectations underreact to aggregate shocks, they tend to overreact to idiosyncratic shocks. We show that overreaction with respect to idiosyncratic news has only a small impact on our results: the extended model also accounts for facts (i) - (iv) simultaneously and even quantitatively, the results are barely affected by introducing bounded rationality also with respect to idiosyncratic risk.

We then show that accounting for facts (i) - (iv) simultaneously, matters greatly for the model's policy implications. Many advanced economies have recently experienced a dramatic surge in inflation which is partly attributed to disruptions in production (see di Giovanni et al. (2022)). We analyze these supply disruptions by considering a negative productivity shock.[4] We uncover a novel amplification channel of these supply shocks as both—the underlying heterogeneity and bounded rationality—amplify the inflationary pressure from the supply shock and the two mutually reinforce each other: the positive output gap redistributes towards households with higher MPCs increasing the output gap further and, thus, calls for higher interest rates in each period. As households cognitively discount these higher (future) interest rates, this further increases the output gap amplifying the redistribution to high MPCs households and therefore the increase in the output gap until the economy ends up in an equilibrium with a higher output gap and higher inflation. As a consequence, inflation increases by more than twice as much as in a model without household heterogeneity and bounded rationality.

That both—the unequal exposure of households and cognitive discounting—amplify supply shocks is in stark contrast to demand shocks. In response to persistent demand shocks, the unequal exposure of households amplifies the shock whereas cognitive discounting dampens it. Consequently, our model predicts inflation and the output gap to be less responsive to persistent demand shocks but more responsive to supply shocks compared to the rational model.

The amplification channel also implies a more pronounced trade-off for monetary policy between price stability on the one side and fiscal and distributional consequences on the other side after an inflationary supply shock. If monetary policy wants to fully stabilize inflation, it needs to hike interest rates much more aggressively to counteract the amplification forces. These stronger interest-rate hikes create side effects. In particular, they have strong fiscal implications as they increase the cost of government debt, which leads to a larger increase in government debt. Furthermore, consumption inequality increases strongly. The reason is that wealthy households benefit more from higher interest rates than asset-poor households.

Related literature. The literature treats the facts (i)-(iv) mostly independent from each other. The heterogeneous-household literature has highlighted the transmission of monetary policy through indirect, general equilibrium effects (Kaplan et al. (2018), Auclert (2019), Auclert et al. (2020),

[4]We also consider cost-push shocks as an alternative explanation for high inflationary pressure and find similar implications for monetary and fiscal policy.

Bilbiie (2020), Luetticke (2021)), and proposed potential resolutions of the forward guidance puzzle (McKay et al. (2016, 2017), Hagedorn et al. (2019), Acharya and Dogra (2020), McKay and Wieland (2022)). Werning (2015) and Bilbiie (2021) combine the themes of policy amplification and forward guidance puzzle in HANK and establish a trade-off inherent in models with household heterogeneity: if HANK models amplify contemporaneous monetary (and fiscal) policy through redistribution towards high MPC households, they dampen precautionary savings desires after a forward guidance shock which aggravates the forward guidance puzzle.

Few resolutions of this trade-off—what Bilbiie (2021) calls the *Catch-22*—have been put forward. In contrast to our model, they all rely on a specific design for either monetary or fiscal policy. Bilbiie (2021) shows that if monetary policy follows a Wicksellian price level targeting rule or fiscal policy follows a nominal bond rule, his tractable HANK model can simultaneously account for facts (i)-(iv).[5] Hagedorn et al. (2019) shows how introducing nominal government bonds and coupling it with a particular nominal bond supply rule can resolve the forward guidance puzzle in a quantitative HANK model (following the theoretical arguments in Hagedorn (2016) and Hagedorn (2018)). In contrast, we account for the four facts even in the case in which monetary policy follows a standard Taylor rule and absent any nominal bonds or specific fiscal rules.

Farhi and Werning (2019) also combine household heterogeneity with some form of bounded rationality, but focus entirely on resolving the forward-guidance puzzle. Our model accounts for a number of additional empirical facts, such as the transmission of monetary policy through indirect effects in a setting with unequal exposure of households to monetary policy and countercyclical income risk. We also consider a different form of bounded rationality, cognitive discounting, while Farhi and Werning (2019) focus on level-k thinking. Our setup is consistent with the empirical findings in Roth et al. (2021) who show that households adjust their interest-rate expectations only by about half of what the Fed announces, even when being told the Fed's intended interest-rate path.[6] In contrast to these papers, we consider supply shocks and show that the interaction of household heterogeneity and bounded rationality has qualitatively different implications for supply shocks than for forward guidance shocks.

Few other papers share the combination of nominal rigidities, household heterogeneity and some deviation from full information rational expectations (FIRE). Laibson et al. (2021) introduces *present bias* in a model of household heterogeneity but the model is set in partial equilibrium and they do not consider how the power of forward guidance or the stability at the lower bound are affected by the presence of the two frictions. Auclert et al. (2020) incorporate sticky information into a HANK model to generate hump-shaped responses of macroeconomic variables to aggregate shocks while simultaneously matching intertemporal MPCs. Their paper, however, does

[5] Bilbiie (2021) proposes an additional resolution: a pure risk channel which can, in theory, break the comovement of income risk and inequality. However, it requires a calibration which is at odds with the data.

[6] In an extension, we consider the case in which some households (financial markets, for example) fully incorporate the announced interest-rate paths into their expectations (see Section 4.3 where we discuss heterogeneous degrees of cognitive discounting) and show that our results remain robust in that scenario.

not discuss the implications of the deviation from FIRE and heterogeneity for forward guidance or stability at the lower bound.[7]

Outline. The rest of the paper is structured as follows. We present our behavioral HANK model in Section 2. In Section 3, we consider a special calibration that allows us to solve the model in closed form and, thus, to build intuition for our results. In Section 4, we then move to a more standard calibration and show that all results remain robust in that case. We further discuss the role of heterogeneity in the behavioral bias for our results and non-rationality with respect to the idiosyncratic risk. We then use the quantitative model to study the policy implications of inflationary supply-side shocks in Section 5. Section 6 concludes.

2 Model

This section presents our model that incorporates household heterogeneity, cognitive discounting, and nominal rigidities. Initially, we focus on a scenario where prices are sticky, and wages are entirely flexible. This assumption follows the approach of Bilbiie (2021) and Acharya and Dogra (2020), who have demonstrated the inadequacy of HANK models in accounting for facts (i)-(iv) simultaneously.[8] However, the HANK literature increasingly focuses on the case with flexible prices and rigid wages (Broer et al. (2020), Auclert et al. (2021)). In Sections 3.5 and 4.2, we therefore show that all our findings hold when we introduce rigid wages and flexible prices.

2.1 Households

Time is discrete and denoted by $t = 0, 1, 2, ...$ The economy is populated by a unit mass of households, indexed by $i \in [0, 1]$. Households obtain utility from (non-durable) consumption, $C_{i,t}$, and dis-utility from working $N_{i,t}$. Households discount future utility at rate $\beta_{i,t} \in (0, 1)$. We assume a standard CRRA utility function

$$\mathcal{U}(C_{i,t}, N_{i,t}) \equiv \begin{cases} \frac{C_{i,t}^{1-\gamma}}{1-\gamma} - \frac{N_{i,t}^{1+\varphi}}{1+\varphi}, & \text{if } \gamma \neq 1, \\ \log\left(C_{i,t}\right) - \frac{N_{i,t}^{1+\varphi}}{1+\varphi}, & \text{if } \gamma = 1, \end{cases} \tag{1}$$

where φ denotes the inverse Frisch elasticity and γ the relative risk aversion.

Household i faces the budget constraint

$$C_{i,t} + \frac{B_{i,t+1}}{1 + r_t} = B_{i,t} + W_t z(e_{i,t}) N_{i,t} + D_t d(e_{i,t}) - \tau_t(e_{i,t}) \tag{2}$$

[7] Wiederholt (2015), Angeletos and Lian (2018), Andrade et al. (2019), Gabaix (2020) consider deviations from FIRE and Michaillat and Saez (2021) introduce wealth in the utility function (all in non-HANK setups) and show how to resolve the forward guidance puzzle. See, e.g., Broer et al. (2022) and Ilut and Valchev (2023) for recent contributions to how household heterogeneity and deviations from FIRE interact in settings abstracting from nominal rigidities.

[8] Similarly, McKay et al. (2016), who resolve the forward guidance puzzle in a HANK model, also focus on the case with sticky prices and flexible wages.

and the borrowing constraint $B_{i,t+1} \geq \underline{B}$, where $\underline{B}$ denotes an exogenous borrowing limit, B denotes the household's bond holdings, r_t denotes the net real interest rate, W_t the real wage and $e_{i,t}$ the household's exogeneous idiosyncratic state that follows a Markov chain with time-invariant transition matrix $\mathcal{P}$. The process for $e_{i,t}$ is the same for all households and the mass of households in state e at any point in time equals the probability of being in that state in the stationary equilibrium, $p(e)$. Conditional on their exogenous idiosyncratic state, households have the idiosyncratic productivity $z(e_{i,t})$, they receive a share $d(e_{i,t})$ of total dividends D_t, and pay taxes $\tau_t(e_{i,t})$. We introduce taxes in such a way that they are non-distortionary in the sense that they do not show up in the household's first-order conditions. We also allow households' time discount factor to be a function of e, $\beta(e_{i,t})$.

Given their beliefs, households maximize their expected lifetime utility subject to their budget constraint (2) and the borrowing constraint. This yields the Euler equation

$$C_{i,t}^{-\gamma} \geq \beta(e_{i,t}) R_t \mathbb{E}_t^{BR} \left[C_{i,t+1}^{-\gamma} \right], \tag{3}$$

and the labor-leisure equation

$$N_{i,t}^{\varphi} = z(e_{i,t}) W_t C_{i,t}^{-\gamma}, \tag{4}$$

where $R_t \equiv 1 + r_t$ denotes the gross real interest rate. The Euler equation (3) holds with equality when the borrowing constraint does not bind, while it holds with strict inequality when the borrowing constraint binds. $\mathbb{E}_t^{BR}$ denotes the *boundedly-rational* expectations operator which we discuss next.

Bounded rationality. We assume that households are fully rational with respect to their idiosyncratic risk, but they cognitively discount the effects of aggregate shocks (we relax the assumption that households are rational with respect to their idiosyncratic risk in Section 4.4). To model cognitive discounting, we follow Gabaix (2020) but extend it to an economy with a whole distribution of households rather than focusing on a representative consumer.[9] Let X_t be a random variable (or vector of variables) and let us define $\bar{X}_t$ as some default value the agent may have in mind and let $\widetilde{X}_{t+1} \equiv X_{t+1} - \bar{X}_t$ denote the deviation from this default value.[10] The behavioral agent's expectation about X_{t+1} is then defined as

$$\mathbb{E}_t^{BR} \left[X_{t+1} \right] = \mathbb{E}_t^{BR} \left[\bar{X}_t + \widetilde{X}_{t+1} \right] \equiv \bar{X}_t + \bar{m} \mathbb{E}_t \left[\widetilde{X}_{t+1} \right], \tag{5}$$

[9] While Gabaix (2020) embeds bounded rationality in a NK model the basic idea of behavioral inattention (or sparsity) has been proposed by Gabaix earlier already (see Gabaix (2014, 2017)) and a handbook treatment of behavioral inattention is given in Gabaix (2019). We present a way how to microfound cognitive discounting as a noisy-signal extraction problem in Appendix D.7, but note, that the exact microfoundation or underlying behavioral friction which leads to underreaction is not crucial for the rest of our analysis. Angeletos and Lian (2023) show how other forms of bounded rationality or lack of common knowledge can lead to observationally-equivalent expectations.

[10] Gabaix (2020) focuses on the case in which X_t denotes the state of the economy. He shows (Lemma 1 in Gabaix (2020)) that this form of cognitive discounting also applies to all other variables. Appendix D.6 derives our results following the approach in Gabaix (2020). The results remain exactly the same.

where $\mathbb{E}_t[\cdot]$ is the rational expectations operator and $\bar{m} \in [0,1]$ is the cognitive discounting parameter. A higher $\bar{m}$ denotes a smaller deviation from rational expectations and rational expectations are captured by $\bar{m} = 1$. Our setup therefore nests the rational expectations model as a special case.

When $\bar{m} < 1$, the behavioral agent anchors her expectations to the default value and cognitively discounts expected future deviations from this default value. Given that households are perfectly rational with respect to their idiosyncratic risk and only cognitively discount the implications of aggregate shocks, we assume that the default value $\bar{X}_t$ is given by the variable's stationary equilibrium counterpart. Thus, when there is no aggregate shock and the economy is in the stationary equilibrium, $\widetilde{X}_{t+1} = 0$, households are fully rational.

To see how cognitive discounting matters in our model, note that the only forward-looking equation in the household block is the Euler equation (3). Let $\bar{C}_{i,t} \equiv C(e_{i,t}, B_{i,t}, \bar{Z})$ denote consumption of household i in period t with exogeneous idiosyncratic state $e_{i,t}$ and asset holdings $B_{i,t}$ when all aggregate variables are in steady state, indicated by $\bar{Z}$. Here, Z potentially denotes a whole matrix of aggregate variables, including, for example, news shocks (i.e., forward guidance shocks). In other words, $\bar{C}_{i,t}$ denotes consumption of household i with exogeneous state $e_{i,t}$ and asset holdings $B_{i,t}$ in the stationary equilibrium, and thus, the household's default (or anchor) value of consumption. In case an aggregate shock occurs, $Z_t \neq \bar{Z}$, consumption is denoted by $C_{i,t} = C(e_{i,t}, B_{i,t}, Z_t)$. We can then write the Euler equation with bounded rationality (BR) in terms of the rational expectations operator $\mathbb{E}_t[\cdot]$ as

$$\begin{aligned} C_{i,t}^{-\gamma} &\geq \beta(e_{i,t}) R_t \mathbb{E}_t^{BR}\left[C_{i,t+1}^{-\gamma}\right] \\ &= \beta(e_{i,t}) R_t \mathbb{E}_t^{BR}\left[\bar{C}_{i,t+1}^{-\gamma} + \left(C_{i,t+1}^{-\gamma} - \bar{C}_{i,t+1}^{-\gamma}\right)\right] \\ &= \beta(e_{i,t}) R_t \mathbb{E}_t\left[\bar{C}_{i,t+1}^{-\gamma} + \bar{m}\left(C_{i,t+1}^{-\gamma} - \bar{C}_{i,t+1}^{-\gamma}\right)\right], \end{aligned} \tag{6}$$

where the rational expectations operator $\mathbb{E}_t[\cdot]$ denotes the expectations that a fully rational household would have in the behavioral economy.

Equation (6) illustrates that when households form expectations about their marginal utility in the next period, their expectations about the marginal utilities associated with each possible individual state are anchored to the marginal utilities associated with these states in stationary equilibrium. Thus, the household's default value of her future marginal utility is a whole distribution of marginal utilities, depending on her individual state $(e_{i,t}, B_{i,t})$.

Underreaction in the data. Given $\bar{m} < 1$, expectations underreact to aggregate news about the future compared to the rational expectations case, that is, they do not fully incorporate aggregate news into their expectations. We now show that households indeed show patterns of underreaction in the data. We follow Coibion and Gorodnichenko (2015) and regress forecast errors on forecast revisions as follows

$$\underbrace{x_{t+4} - \mathbb{E}_t^{e,BR} x_{t+4}}_{\text{Forecast errors}} = c^e + b^{e,CG} \underbrace{\left(\mathbb{E}_t^{e,BR} x_{t+4} - \mathbb{E}_{t-1}^{e,BR} x_{t+3}\right)}_{\text{Forecast revision}} + \epsilon_t^e, \tag{7}$$

and we do so for different income groups, indexed by e. As we show in Appendix B, $b^{e,CG} > 0$ is consistent with underreaction and the corresponding cognitive discounting parameter can be obtained from

$$\bar{m}^e = \left(\frac{1}{1 + b^{e,CG}}\right)^{1/4}. \tag{8}$$

As we focus on bounded rationality with respect to aggregate shocks, we consider expectations about aggregate variables, namely, unemployment changes, the unemployment level, and inflation which we obtain from the Survey of Consumers from the University of Michigan. We split households into three groups based on their income. The bottom and top income groups each contain the 25% households with the lowest and highest income, respectively, and the remaining 50% are assigned to the middle income group. As the expectations in the forecast revisions in equation (7) are about the variable at different points in time (due to data limitations), we instrument forecast revisions by the *main business cycle shock* obtained from Angeletos et al. (2020).

We find that in all cases $\widehat{b}^{e,CG}$ is positive, suggesting that households of all income groups tend to underreact, consistent with our assumption of $\bar{m} < 1$ (Table 2 in the Appendix provides the details). Using equation (8) we obtain estimates of $\bar{m}^e$ equal to 0.57, 0.59 and 0.64 for the bottom 25%, the middle 50% and the top 25%, respectively for the estimates from the IV regressions when focusing on expected unemployment changes. When we consider unemployment levels rather than changes, the estimated $\bar{m}^e$ equal 0.86, 0.87 and 0.88. If we consider inflation expectations instead of unemployment expectations, we obtain estimated cognitive discounting parameters of 0.70, 0.75 and 0.78 for the bottom 25%, the middle 50% and the top 25%, respectively.[11]

There are two take-aways from this empirical exercise: first, households of all income groups underreact in their expectations. Second, the estimated cognitive discounting parameters tend to be between 0.6 and 0.85, consistent with values used in Gabaix (2020).

Consistent with our findings, Kučinskas and Peters (2022) and Born et al. (2022) find that professional forecasters and firms, respectively, underreact to aggregate shocks. However, they also find evidence of *overreaction* (as in Bordalo et al. (2020)) to idiosyncratic shocks. We discuss this case where households underreact to aggregate shocks but overreact to idiosyncratic shocks in Section 4.4.

2.2 Firms

We assume a standard New Keynesian firm side with sticky prices and where firms have rational expectations (the case with flexible prices and sticky wages is discussed in Sections 3.5 and 4.2,

[11]Estimates using OLS rather than IV are similar (see Appendix B).

and the case with boundedly-rational firms in 5 and Appendix D.5). All households consume the same aggregate basket of individual goods, $j \in [0,1]$, $C_t = \left(\int_0^1 C_t(j)^{\frac{\epsilon-1}{\epsilon}} dj\right)^{\frac{\epsilon}{\epsilon-1}}$, where $\epsilon > 1$ is the elasticity of substitution between the individual goods. Each firm faces demand $C_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} C_t$, where $P_t(j)/P_t$ denotes the individual price relative to the aggregate price index, $P_t^{1-\epsilon} = \int_0^1 P_t(j)^{1-\epsilon} dj$, and produces with the linear technology $Y_t(j) = N_t(j)$. Firms can only update their prices infrequently, as in Calvo (1983) and Yun (1996). The real marginal cost is given by W_t. We assume that the government pays a constant subsidy τ^S on revenues to induce marginal cost pricing in the steady state. This subsidy is financed by a lump-sum tax on firms T_t^F. Hence, the profit function is $D_t(j) = (1+\tau^S)[P_t(j)/P_t]Y_t(j) - W_t N_t(j) - T_t^F$. Total profits are then $D_t = Y_t - W_t N_t$ and are zero in steady state.

2.3 Government

The government consists of a fiscal authority and a monetary authority. The fiscal authority faces the budget constraint

$$\frac{B_{t+1}^G}{R_t} + T_t = B_t^G,$$

where B^G denotes the bonds issued by the government and T_t denotes tax income. We abstract from government spending. Taxes follow a simple debt feedback rule

$$T_t - \bar{T} = \vartheta \frac{B_{t+1}^G - \bar{B}^G}{\bar{Y}}, \tag{9}$$

where $\bar{T}$, $\bar{B}^G$ and $\bar{Y}$ denote the respective steady state values. Further, fiscal policy induces the optimal steady state subsidy financed by lump-sum taxation of firms.

In most of the analysis, we assume that monetary policy either sets the nominal interest rate i_t following a standard (linearized) Taylor rule

$$\widehat{i}_t = \phi \pi_t + \epsilon_t^{MP}, \tag{10}$$

or a *real rate rule*

$$r_t = \bar{r} + \epsilon_t^{MP}, \tag{11}$$

with ϵ_t^{MP} being a monetary policy shock, π_t denoting inflation, $\bar{r}$ the steady-state real interest rate, and where variables with a "$\widehat{\ }$" denote log deviations from the variables' respective steady state values. The parameter ϕ captures how strongly monetary policy responds to inflation. For now, monetary policy shocks are the only source of aggregate uncertainty.

Equilibrium definition. Given an initial price level P_{-1}, initial government debt level B_0^G, and an initial distribution of agents $\Psi_0(B_0, e_0)$, a general equilibrium is a path for prices $\{P_t, W_t, \pi_t, r_t, i_t\}$, aggregates $\{Y_t, C_t, N_t, B_{t+1}^G, T_t, D_t\}$, individual allocation rules $\{C_t(B_t, e_t), B_{t+1}(B_t, e_t)\}$ and joint distributions of agents $\Psi_t(B_t, e_t)$ such that households optimize (given their beliefs), all firms optimize, monetary and fiscal policy follow their rules, and the goods and bond markets clear:

$$\sum_e p(e) \int C_t\left(B_t, e_t\right) \Psi_t\left(B_t, e_t\right) = Y_t$$

$$\sum_e p(e) \int B_{t+1}\left(B_t, e_t\right) \Psi_t\left(B_t, e_t\right) = B_{t+1}^G.$$

3 Analytical Results

To understand how household heterogeneity and cognitive discounting interact, we now calibrate the model such that we can solve the model in closed form. We refer to this specific calibration as *tractable* behavioral HANK model as it nests the tractable rational HANK model of Bilbiie (2020, 2021).

3.1 A Calibration towards a Closed-Form Solution

Solving the model in closed form requires specific functional forms for $\beta(e)$, $z(e)$, $d(e)$, and $\tau(e)$, for the stochastic process of e, as well as $B_t^G = \underline{B} = 0$ for all t. Starting with the process of e, we for now assume that there are only two states, $e \in \{U, H\}$, and denote a household's probability to remain in her current state $p(e_{t+1} = U|e_t = U) = s$ and $p(e_{t+1} = H|e_t = H) = h$. Consequently, $\lambda = \frac{1-s}{2-s-h}$ is the time-constant share of households being in state H. We then assume that $\beta(H) < \beta(U)$ such that the Euler equation (6) always holds with equality for households being in state U, while it always holds with inequality for households being in state H. In other words, H households are always $H$and-to-Mouth, while U households are always Unconstrained. In addition, we assume that $z(e) = 1$ and $\tau(e) = 0$ for both states and $d(H) = \frac{\mu^D}{\lambda}$ and $d(U) = \frac{1-\mu^D}{1-\lambda}$. This leaves two sources of income heterogeneity, namely, different labor supply and different profit shares.

The assumption that $B_t^G = \underline{B} = 0$ for all t means that the government does not issue any bonds and households cannot borrow. It follows that households cannot save in equilibrium and therefore, all H households are identical and all U households are identical, independent of how long they have been in state U or in state H. We can thus solve the model as if there were two representative households, a Hand-to-Mouth and an Unconstrained household. Hence, in this section, we will use superscripts H and U to indicate the two representative households.

As profits are zero in steady state due to the subsidy induced by fiscal policy, it follows that households are identical in steady state, $C^H = C^U = C$. In the log-linear dynamics around this steady state, profits vary inversely with the real wage, $\widehat{d}_t = -\widehat{w}_t$. We allow for steady state inequality in Appendix D and show that our results are not driven by this assumption.

3.2 Log-Linearized Dynamics

We now focus on the log-linearized dynamics around the full-insurance, zero-liquidity steady state. The first key equilibrium equation is the consumption of the hand-to-mouth households written as a function of total output

$$\widehat{c}_t^H = \chi \widehat{y}_t, \tag{12}$$

with

$$\chi \equiv 1 + \varphi\left(1 - \frac{\mu^D}{\lambda}\right) \tag{13}$$

measuring the cyclicality of the H household's consumption (see appendix A.1). Patterson (2023) documents that households with higher MPCs tend to be more exposed to aggregate income fluctuations induced by monetary policy or other demand shocks—fact (ii) in the introduction. We can account for fact (ii) by setting $\chi > 1$. Similarly, Auclert (2019) finds that poorer households tend to exhibit higher MPCs. Together with the findings in Coibion et al. (2017) and Hintermaier and Koeniger (2019) that poorer households' income is on average more exposed to monetary policy shocks, this also implies $\chi > 1$. For given φ, this requires $\mu^D < \lambda$.

Why does $\mu^D < \lambda$ imply that the consumption of hand-to-mouth households moves more than one-for-one with aggregate output after a monetary policy shock? Consider an expansionary monetary policy shock, i.e., an unexpected decrease in the interest rate. Unconstrained households want to consume more and save less, leading to an increase in demand. Firms then increase their labor demand, leading to an increase in wages. Due to the assumption of sticky prices and flexible wages, profits in the New Keynesian model decrease ($\widehat{d}_t = -\widehat{w}_t$). In the representative agent model, the representative agent both incurs the increase in wages and the decrease in profits coming from firms. With household heterogeneity, however, this is not necessarily the case. If $d(H) < 1$, which is the case when $\mu^D < \lambda$, the decrease in profits affects the income of H households less than one-for-one while the increase in the real wage affects their income one-for-one. Thus, the total income of H households increases more than one-for-one with aggregate income.

Combining equation (12) with the goods market clearing condition yields

$$\widehat{c}_t^U = \frac{1 - \lambda\chi}{1 - \lambda}\widehat{y}_t, \tag{14}$$

which implies that consumption inequality is given by:[12]

$$\widehat{c}_t^U - \widehat{c}_t^H = \frac{1 - \chi}{1 - \lambda}\widehat{y}_t. \tag{15}$$

Thus, if $\chi > 1$, inequality is countercyclical as it varies negatively with total output, i.e., inequality increases in recessions and decreases in booms. In line with the empirical evidence on the covariance between MPCs and income exposure, the data also points towards $\chi > 1$ when looking

[12] We denote the case in which unconstrained households consume relatively more than hand-to-mouth households as higher inequality, even though they consume the same amount in steady state. As we move away from the tractable model in Sections 4 and 5, households' consumption levels will differ in the stationary equilibrium.

at the cyclicality of inequality, conditional on monetary policy: Coibion et al. (2017), Mumtaz and Theophilopoulou (2017), Ampudia et al. (2018) and Samarina and Nguyen (2019) all provide evidence of countercyclical inequality conditional on monetary policy shocks.

The second key equilibrium equation is the log-linearized bond Euler equation of U households:

$$\widehat{c}_t^U = s\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] + (1-s)\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^H\right] - \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t^{BR}\pi_{t+1}\right). \tag{16}$$

For the case without idiosyncratic risk, i.e., for $s = 1$, equation (16) boils down to a standard Euler equation under bounded rationality. For $s \in [0, 1)$, however, the household takes into account that she might be hit by an idiosyncratic shock and self-insures against becoming hand-to-mouth next period. How strongly this precautionary savings motive affects the household's consumption away from the stationary equilibrium will depend on the household's degree of bounded rationality. We will, following the assumption in Gabaix (2020), often focus on the case in which households are rational with respect to today's real rate, i.e., we replace $\mathbb{E}_t^{BR}\pi_{t+1}$ with $\mathbb{E}_t\pi_{t+1}$ in equation (16). We show in Appendix D that our results go through with boundedly-rational expectations of today's real rate.

Supply side. For simplicity and to get a clear understanding of the mechanisms driving our results, we focus on a static Phillips curve in this section:

$$\pi_t = \kappa\widehat{y}_t, \tag{17}$$

where $\kappa \geq 0$ captures the slope of the Phillips curve. Such a static Phillips curve arises if we assume that firms are either completely myopic or if they face Rotemberg-style price adjustment costs relative to yesterday's market average price index, instead of their own price (see Bilbiie (2021)). In Appendix D.5, we show that a forward-looking Phillips Curve (rational or behavioral) does not qualitatively affect our results.

3.3 The Closed-Form Solution

Our tractable behavioral HANK model can be summarized by three equations: a Phillips curve, representing the aggregate supply side captured by equation (17), a rule for monetary policy (equation (10) or (11)), which together with the third equation—the aggregate IS equation—determines aggregate demand. To obtain the aggregate IS equation, we combine the hand-to-mouth households' consumption (12) with the consumption of unconstrained households (14) and their Euler equation (16) (see appendix A for all the derivations).

Proposition 1. *The aggregate IS equation is given by*

$$\widehat{y}_t = \psi_f\mathbb{E}_t\widehat{y}_{t+1} - \psi_c\frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t\pi_{t+1}\right), \tag{18}$$

where

$$\psi_f \equiv \bar{m}\delta = \bar{m}\left[1 + (\chi - 1)\frac{1-s}{1-\lambda\chi}\right] \quad \textit{and} \quad \psi_c \equiv \frac{1-\lambda}{1-\lambda\chi}.$$

Compared to the rational representative agent model, two new coefficients show up: ψ_c and ψ_f. ψ_c governs the sensitivity of today's output with respect to the contemporaneous real interest rate. ψ_c is shaped by household heterogeneity, in particular by the share of H households λ and their income exposure χ. As the H households' incomes are more exposed to aggregate income ($\chi > 1$), $\psi_c > 1$ which renders current output more sensitive to changes in the contemporaneous real interest rate due to general equilibrium forces, as we show later.

The second new coefficient in the behavioral HANK IS equation (18), ψ_f, captures the sensitivity of today's output with respect to changes in expected future output. ψ_f is shaped by household heterogeneity and the behavioral friction as it depends on the precautionary-savings motive, captured by δ, *and* the degree of bounded rationality of households as well as the interaction of these two. Given that $\chi > 1$, unconstrained households take into account that they will be more exposed to aggregate income fluctuations in case they become hand-to-mouth. Thus, income risk is countercyclical, which manifests itself in $\delta > 1$ (consistent with the empirical evidence, e.g., in Storesletten et al. (2004) or Guvenen et al. (2014)). Countercyclical risk induces compounding in the Euler equation and, thus, competes with the empirically observed underreaction of aggregate expectations ($\bar{m} < 1$) which induces discounting in the Euler equation. We see in the following sections that even for a small degree of bounded rationality—much smaller than the empirics suggest—that discounting through bounded rationality dominates the compounding through countercyclical income risk. Hence, in the behavioral HANK model it holds that $\psi_f < 1$ which makes the economy less sensitive to expectations and news about the future.

Equation (18) nests IS equations of three classes of models in the literature: first, the representative-agent rational expectations (RANK) model which can be obtained by setting $\bar{m} = 1$ and assuming only one state $e = U$ which would imply $\psi_f = \psi_c = 1$ (see Galí (2015), Woodford (2003)).[13] Second, representative agent models deviating from full-information rational expectations when assuming one state and $\bar{m} \in (0, 1)$ which results in $\psi_c = 1$ and $\psi_f < 1$ as, for example, in Gabaix (2019), Angeletos and Lian (2018) and Woodford (2019). And thirdly, TANK and tractable HANK models as e.g., in Bilbiie (2008), Bilbiie (2021), McKay et al. (2017), or Debortoli and Galí (2018) with again two states but $\bar{m} = 1$ which implies $\psi_f = \delta$. Nesting these models enables us to clearly illustrate why our model is able to account for fact (i) - (iv) simultaneously while these other models cannot.

Calibration. Given that this stylized version of our model is a tractable HANK model with two agents, we calibrate it using standard parameters in the literature on tractable HANK models (see, e.g., Bilbiie (2020, 2021)).[14] That said, we show in Appendix D.1 that our results are robust to a wide range of parameters. We set the share of H agents to one third, $\lambda = 0.33$, and μ^D such that $\chi = 1.35$ which implies $\psi_c = 1.2$. We set $\chi > 1$ to capture that high-MPC households' incomes

[13] Only one state implies that χ vanishes from the model and $\lambda = 0$ and $s = 1$.

[14] In the next section, we then show how we can use our quantitative model to directly match micro evidence from Patterson (2023) on the unequal income exposure of households.

are relatively more sensitive to aggregate fluctuations induced by monetary policy, in line with the findings in Patterson (2023), Coibion et al. (2017) and Auclert (2019). We set the probability of a U household to become hand-to-mouth next period to 5.4%, i.e., $s = 0.946$ (this corresponds to $s = 0.8$ in annual terms). We focus on log utility, $\gamma = 1$, set $\beta(U) = 0.99$, and the slope of the Phillips Curve to $\kappa = 0.02$, as in Bilbiie et al. (2022). The cognitive discounting parameter, $\bar{m}$ is set to 0.85, as explained in Section 2. Note, that even when we vary certain parameters, we keep $\lambda < \chi^{-1}$.

3.4 Monetary Policy

We now show how the behavioral HANK model generates amplification of contemporaneous monetary policy through indirect effects while resolving the forward guidance puzzle at the same time. Additionally, we discuss determinacy conditions and show that the model remains stable at the effective lower bound.

General equilibrium amplification and forward guidance. We start by showing how the behavioral HANK model generates amplification of current monetary policy through indirect general equilibrium effects while simultaneously ruling out the forward guidance puzzle. The forward guidance puzzle states that announcements about future changes in the interest rate affect output today as strong (or even stronger) than contemporaneous changes in the interest rate.[15] Such strong effects of future interest rate changes, however, seem puzzling and are not supported by the data (Del Negro et al. (2015), Roth et al. (2021)).

Let us consider two different monetary policy experiments: (i) a contemporaneous monetary policy shock, i.e., a surprise decrease in the real interest rate today, and (ii) a forward guidance shock, i.e., a news shock today about a decrease in the real interest rate k periods in the future. The monetary authority keeps the real interest rate at its steady state value in all other periods. We focus on real rate changes as this is the set up that McKay et al. (2016) focus on and Farhi and Werning (2019) focus on the case with fully-rigid prices, such that nominal rate changes translate one-for-one to real rate changes. However, all our results are robust when focusing on nominal rate changes and are presented in Appendix D.2

Proposition 2. *In the behavioral HANK model, there is amplification of contemporaneous monetary policy relative to RANK if and only if*

$$\psi_c > 1 \Leftrightarrow \chi > 1, \tag{19}$$

and the forward guidance puzzle is ruled out if

$$\psi_f < 1. \tag{20}$$

[15] Detailed analyses of the forward guidance puzzle in RANK are provided by McKay et al. (2016) and Del Negro et al. (2015).

Let us first focus on equation (19) which tells us that the behavioral HANK model generates amplification of contemporaneous monetary policy with respect to RANK whenever $\chi > 1$, that is, when high-MPC households' consumption is relatively more sensitive to aggregate income fluctuations.

After a decrease in the interest rate, wages increase and profits decline. As H agents receive a relatively smaller share of profits but fully benefit from the increase in wages, their income increases more than one-to-one with aggregate income. As they consume their income immediately, the initial effect on total output increases. The unconstrained households, on the other hand, experience a smaller increase in their income due to the fall in their profit income. As a result, $\psi_c > 1$ and the increase in output is amplified through these general equilibrium effects. To see the importance of general equilibrium or indirect effects, the following Lemma disentangles the direct and indirect effects.

Lemma 1. *The consumption function in the behavioral HANK model is given by*

$$\widehat{c}_t = [1 - \beta(1 - \lambda\chi)]\,\widehat{y}_t - \frac{(1-\lambda)\beta}{\gamma}\widehat{r}_t + \beta\bar{m}\delta(1 - \lambda\chi)\mathbb{E}_t\widehat{c}_{t+1}. \tag{21}$$

Let ρ denote the exogenous persistence and define the indirect effects as the change in total consumption due to the change in total income but for fixed real rates. The share of indirect effects, Ξ^{GE}, out of the total effect is then given by

$$\Xi^{GE} = \frac{1 - \beta(1 - \lambda\chi)}{1 - \beta\bar{m}\delta\rho(1 - \lambda\chi)}.$$

Given our calibration and assuming an AR(1) monetary policy shock with a persistence of 0.6, indirect effects account for about 63%, consistent with larger quantitative models as for example in Kaplan et al. (2018) and thus, the model accounts for fact (i).[16] Holm et al. (2021) state that the overall importance of indirect effects they find in the data is comparable to those in Kaplan et al. (2018), with the difference that these effects unfold after some time, whereas direct effects are more important on impact. Because in our stylized model the response to a monetary policy shock peaks on impact, indirect effects are important right away. Slacalek et al. (2020) provide further evidence that indirect effects are strong drivers of aggregate consumption in response to monetary policy shocks. For comparison, the representative agent model generates an indirect share of $\Xi^{GE} = \frac{1-\beta}{1-\beta\bar{m}\rho}$, which, given our calibration, amounts to about 2%.

Turning to forward guidance, equation (20) in Proposition 2 tells us that the forward guidance puzzle is ruled out if $\psi_f < 1$. What determines whether this condition holds or not? First, note that as in the discussion of contemporaneous monetary policy, with $\chi > 1$ the income of H agents moves more than one for one with aggregate income. In this case, unconstrained households who self-insure against becoming hand-to-mouth in the future want less insurance when they expect a decrease in the interest rate because if they become hand-to-mouth they would benefit more

[16] We write β for $\beta(U)$ for notational simplicity and because $\beta(H)$ does not affect any of our results (as long as it is low enough such that the borrowing constraint always binds for H households).

from the increase in aggregate income. Hence, after a forward guidance shock, unconstrained households decrease their precautionary savings which compounds the increase in output today ($\delta > 1$). Yet, as households are boundedly rational, they cognitively discount these effects taking place in the future. Importantly, unconstrained households cognitively discount both the usual consumption-smoothing response due to the future increase in consumption as well as the general equilibrium implications for their precautionary savings, thereby decreasing the effects of the forward guidance shock on today's consumption. Thus, the model not only accounts for facts (i) and (ii) but simultaneously accounts for fact (iii).

This last part clearly illustrates the main interaction of bounded rationality and household heterogeneity that enables the behavioral HANK model to resolve the forward guidance puzzle while simultaneously generating amplification through indirect effects. Households fully understand their idiosyncratic risk of switching their type as well as the implications of switching type in case there are no aggregate shocks, i.e., in the steady state. If the monetary authority makes an unexpected announcement about its future policy, however, behavioral households do not fully incorporate the effects of this policy on their own income risk and thus, their precautionary savings. Already a small underreaction of the behavioral households is enough to resolve the forward guidance puzzle. Given our calibration there is no forward guidance puzzle in the behavioral HANK model as long as $\bar{m} < 0.966$ which is above the upper bounds for empirical estimates (see Section 2).[17] Figure 10 in Appendix D.1 shows that the solution of the forward guidance in our model is very robust with respect to changes in the heterogeneity parameters.

We now compare the behavioral HANK model to its rational counterpart to show how the behavioral HANK model overcomes a shortcoming inherent in the rational HANK model – the *Catch-22* (Bilbiie (2021); see also Werning (2015)). The Catch-22 describes the tension that the rational HANK model can either generate amplification of contemporaneous monetary policy *or* solve the forward guidance puzzle. To see this, note that with $\bar{m} = 1$ the forward guidance puzzle is resolved when $\delta < 1$ which requires $\chi < 1$, as otherwise $\delta > 1$. Assuming $\chi < 1$, however, leads to *dampening* of contemporaneous monetary policy instead of amplification. We graphically illustrate the Catch-22 of the rational model and its resolution in the behavioral HANK model in Appendix C. Note that also rational TANK models (thus, turning off type switching) or the behavioral RANK model would not deliver amplification and resolve the forward guidance puzzle simultaneously. TANK models would face the same issues as the rational RANK model in the sense that they cannot solve the forward guidance puzzle while bounded rationality in a RANK model does not deliver initial amplification through indirect effects.

[17] A related paradox in the rational model is that as the persistence of the shock increases, the effects become unboundedly large and as the persistence approaches unity, an exogenous increase in the nominal interest rate becomes expansionary. The behavioral HANK model, on the other hand, does not suffer from this. We elaborate these points in more detail in Appendix D.4.

3.4.1 Stability at the Effective Lower Bound

In this section, we revisit the determinacy conditions in the behavioral HANK model and discuss the implications for the stability at the effective lower bound constraint on nominal interest rates. We therefore focus on the case where monetary policy follows the Taylor rule (10) (we discuss more general Taylor rules in Appendix A.4). To derive these results, it is sometimes convenient to combine the IS equation (18) with the static Phillips Curve (17) and the Taylor rule (10) so that we can represent the model in a single first-order difference equation:

$$\widehat{y}_t = \frac{\psi_f + \psi_c \frac{\kappa}{\gamma}}{1 + \psi_c \phi \frac{\kappa}{\gamma}} \mathbb{E}_t \widehat{y}_{t+1} - \frac{\psi_c \frac{1}{\gamma}}{1 + \psi_c \phi \frac{\kappa}{\gamma}} \varepsilon_t^{MP}. \tag{22}$$

According to the Taylor principle, monetary policy needs to respond sufficiently strongly to inflation in order to guarantee a determinate equilibrium. In the rational RANK model the Taylor principle is given by $\phi > 1$, where ϕ is the inflation-response coefficient in the Taylor rule (10). We now derive a similar determinacy condition in the behavioral HANK model and show that both household heterogeneity and bounded rationality affect this condition. The following proposition provides the behavioral HANK Taylor principle.[18]

Proposition 3. *The behavioral HANK model has a determinate, locally unique equilibrium if and only if:*

$$\phi > \phi^* = 1 + \frac{\bar{m}\delta - 1}{\frac{\kappa}{\gamma} \frac{1-\lambda}{1-\lambda\chi}}. \tag{23}$$

We obtain Proposition 3 directly from the difference equation (22). For determinacy, we need that the coefficient in front of $\mathbb{E}_t \widehat{y}_{t+1}$ is smaller than 1 (the eigenvalues associated with any exogenous variables are assumed to be $\rho < 1$, and are thus stable). Solving this condition for ϕ yields Propositon 3. Appendix A.4 outlines the details and extends the result to more general Taylor rules.

To understand the condition in Proposition 3, consider first $\bar{m} = 1$ and, thus, focus solely on the role of household heterogeneity. With $\chi > 1$, it follows that $\phi^* > 1$ and, hence, the threshold is higher than the RANK Taylor principle states. This insufficiency of the Taylor principle in rational HANK models has been shown by Bilbiie (2021) and in a similar way by Ravn and Sterk (2021) and Acharya and Dogra (2020). As a future aggregate sunspot increases the income of households in state H disproportionately, unconstrained households cut back on precautionary savings today which further increases output today. This calls for a stronger response of the central bank to not let the sunspot become self-fulfilling.

On the other hand, bounded rationality $\bar{m} < 1$ relaxes the condition as unconstrained households now cognitively discount both the future aggregate sunspot as well as its implications for their idiosyncratic risk. A smaller response of the central bank is needed in order to prevent the sunspot to become self-fulfilling. Given our calibration the cutoff value for $\bar{m}$ to restore the

[18] We focus on local determinacy and bounded equilibria.

RANK Taylor principle in the behavioral HANK model is 0.966. What is more, given our baseline choice of $\bar{m} = 0.85$, we obtain $\phi^* < 0$. Thus, in our tractable behavioral HANK model it is not necessary that monetary policy responds to inflation at all as the economy features a stable unique equilibrium even under an interest rate peg.

Stability at the effective lower bound. Related to the indeterminacy issues under a peg the traditional New Keynesian model struggles to explain how the economy can remain stable when the effective lower bound (ELB) on nominal interest rates is binding for an extended period of time, as observed in many advanced economies over recent decades (see, e.g., Debortoli et al. (2020) and Cochrane (2018)). If the ELB binds for a sufficiently long time, RANK predicts unreasonably large recessions and, in the limit case in which the ELB binds forever, even indeterminacy.[19] Similar to the forward guidance puzzle, this is even more severe in rational HANK models.

We now show that the behavioral HANK model resolves these issues and thus accounts for fact (iv). To this end, let us add a *natural rate shock* (i.e., a demand shock) $\widehat{r}_t^n$ to the IS equation:

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \left(\widehat{i}_t - \mathbb{E}_t \pi_{t+1} - \widehat{r}_t^n \right).$$

We assume that in period t the natural rate decreases to a value $\widetilde{r}^n$ that is sufficiently negative such that the natural rate in levels is below the ELB. The natural rate stays at $\widetilde{r}^n$ for $k \geq 0$ periods and after k periods the economy returns immediately back to steady state. Agents correctly anticipate the length of the binding ELB. Iterating the IS equation forward, it follows that output in period t is given by

$$\widehat{y}_t = -\frac{1}{\gamma} \psi_c \underbrace{\left(\widehat{i}_{ELB} - \widetilde{r}^n \right)}_{>0} \sum_{j=0}^{k} \left(\psi_f + \frac{\kappa}{\gamma} \psi_c \right)^j, \tag{24}$$

where the term $\left(\widehat{i}_{ELB} - \widetilde{r}^n \right) > 0$ captures the shortfall of the policy response due to the binding ELB. Under rational expectations, we have that $\psi_f > 1$ (and $\frac{\kappa}{\gamma}\psi_c > 0$), meaning that output implodes as $k \to \infty$. The same is true in the rational RANK model which is captured by $\psi_f = \psi_c = 1$. In the behavioral HANK model, however, this is not the case. As long as $\psi_f + \frac{\kappa}{\gamma}\psi_c < 1$ the output response in t is bounded even as $k \to \infty$. It follows that $\bar{m} < 0.94$ is enough to rule out unboundedly-severe recessions at the ELB even if the ELB is expected to persist forever. We graphically illustrate in Appendix C that the behavioral HANK model remains stable also for long spells of the ELB in which output in the rational models collapses.

[19] A forever binding ELB basically implies that the Taylor coefficient is equal to zero and, thus, the nominal rate is pegged at the lower bound, thereby violating the Taylor principle. Note, that this statement also extends to models featuring more elaborate monetary policy rules including Taylor rules responding to output or also the Wicksellian price-level targeting rule, as they all collapse to a constant nominal rate in a world of an ever-binding ELB.

3.5 Sticky Wages

So far, we have assumed that prices are sticky and wages are fully flexible. This assumption, however, is not crucial for our aggregate results. The key difference in our context is the underlying mechanism for the unequal exposure of households.

To highlight this, we now consider wages to be sticky and prices to be fully flexible. Given that prices are fully flexible, we also abstract from monopolistic competition of firms, that is, prices are set to marginal costs. From the aggregate production function, $Y_t = N_t$, it follows that with flexible prices the aggregate price index P_t equals the nominal wage, such that that the real wage W_t is constant and equal to 1 (Auclert et al. (2018)). Further, a labor union allocates hours to workers and we assume that all households work the same amount in the steady state. If there is an aggregate shock, however, and hours deviate from their steady state value, $\widehat{n}_t \neq 0$, the labor union allocates these hours as follows:

$$\widehat{n}_t^H = \zeta \widehat{n}_t,$$

with ζ capturing the H households' sensitivity of hours worked to changes in total hours worked. Absent profits, taxes, and transfers, this allocation rule is the only source of income heterogeneity and, thus, ζ is a sufficient statistic of households' income exposure to monetary policy. Fact (ii)—that households with higher MPCs are more strongly exposed to monetary policy shocks—implies $\zeta > 1$. It directly follows that $\widehat{c}_t^H = \zeta \widehat{y}_t$ from the H households' budget constraint and the production function. Market clearing then yields $\widehat{c}_t^U = \frac{1-\lambda\zeta}{1-\lambda}\widehat{y}_t$.

Using these expressions for $\widehat{c}^H$ and $\widehat{c}^U$ in the unconstrained household's Euler equation yields

$$\frac{1-\lambda\zeta}{1-\lambda}\widehat{y}_t = s\bar{m}\frac{1-\lambda\zeta}{1-\lambda}\mathbb{E}_t\widehat{y}_{t+1} + (1-s)\bar{m}\zeta\mathbb{E}_t\widehat{y}_{t+1} - \frac{1}{\gamma}\widehat{r}_t,$$

which is exactly the same IS equation as the one in Proposition 1 when setting $\zeta = \chi$ (see Appendix A for the algebra). Thus, the effects of conventional monetary policy shocks as well as the effects of forward guidance shocks on total output are exactly the same as in our baseline model with sticky prices and flexible wages. But instead of relying on the countercyclicality of profits, the model with sticky wages and flexible prices relies on the labor union's allocation rule of hours worked outside of the steady state to match fact (ii).

4 Quantitative Results

In this section, we relax the specific calibration choices that we use to solve the model in closed-form and show that all our results carry over. To this end, we build on a standard calibration in the HANK literature which implies that the model features a non-degenerate wealth distribution and, thus, needs to be solved numerically. To account for the micro evidence, we add two new ingredients to the standard calibration featuring the essence of our analysis: first, heterogeneous exposure to monetary policy shocks such that high MPC households tend to be more exposed to these shocks, and second, cognitive discounting of households with respect to aggregate shocks.

Calibration. Table 1 summarizes our baseline calibration. We set the discount factor β to match a steady state real rate of 2% (annualized). In contrast to Section 3, we now abstract from differences in time discounting, $\beta(e_{i,t}) = \beta$ for all $e_{i,t}$, such that borrowing constraints only bind for endogenous reasons. To calibrate the idiosyncratic skill process, we set $z(e_{i,t}) = e_{i,t}$ and we follow McKay et al. (2016) in assuming that $e_{i,t}$ follows an AR(1) process with autocorrelation $\rho_e = 0.966$ and variance $\sigma_e^2 = 0.033$ to match the volatility of the distribution of five-year earnings growth rates found in Guvenen et al. (2014). We then discretize this process into a three-states Markov chain using the Rouwenhorst (1995) method. 25% of households are in the lowest and the highest state, respectively, and 50% in the middle state. We set the amount of government debt to match the aggregate MPC of 0.16 out of an income windfall of 500$, as in Kaplan et al. (2018). This results in a government debt-to-annual-GDP level of 69%. We use standard parameters for our supply side. We set the the price markup to 1.2 and the Calvo probability to reset the price to 0.15 as in Christiano et al. (2011).

Table 1: Baseline Calibration Of the Behavioral HANK Model

Parameter	Description	Value
R	Steady State Real Rate (annualized)	2%
γ	Risk aversion	2
φ	Inverse of Frisch elasticity	2
μ	Markup	1.2
θ	Calvo Price Stickiness	0.15
ρ_e	Autocorrelation of idiosyncratic risk	0.966
σ_e^2	Variance of idiosyncratic risk	0.033
$\tau(e)$	Tax shares	[0, 0, 1]
$d(e)$	Dividend shares	$[\frac{0.06}{0.25}, \frac{0.18}{0.5}, \frac{0.76}{0.25}]$
$\frac{B^G}{4Y}$	Government debt	0.69
$\bar{m}$	Cognitive discounting	0.85

To capture fact (ii)—that higher MPC households tend to be on average more exposed to aggregate income changes induced by monetary policy—we target the estimates in Patterson (2023). Patterson (2023) finds that regressing the income elasticity of households with respect to aggregate changes in output on households' MPC yields a regression coefficient of 1.33. We match this estimate by calibrating the dividend shares the households receive. To do so, we assume that the aggregate income fluctuations are due to monetary policy shocks. We obtain a calibration that implies that households with a higher productivity receive a larger share of the dividends than households with a lower productivity. About 75% of the dividends goes to the highest productivity households. These numbers are consistent with the empirical findings in Kuhn et al. (2020).[20] To

[20]As MPCs are highly negatively correlated with productivity, the intuition why this leads to a higher exposure

show the robustness of our results, we also consider a case in which high MPC households are even more exposed to the business cycle in Section 4.5.

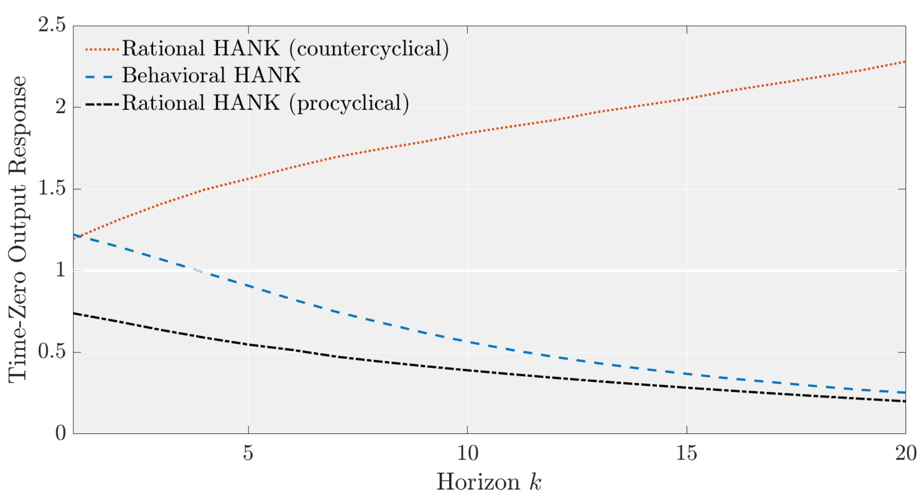
To capture the underreaction of households to aggregate news, we set the cognitive discounting $\bar{m} = 0.85$, which corresponds to the upper bound of our empirical findings presented in Section 2. Since the heterogeneity in cognitive discounting seems to be small in the data (see Section 2), we assume that all households have the same degree of rationality in our baseline calibration. Yet, we also consider the case of heterogeneous degrees of rationality in Section 4.3. In Section 4.4, we discuss how bounded rationality with respect to *idiosyncratic* shocks affect our results.

The rest of the calibration, i.e., $\gamma = 2$, $\varphi = 2$, and the tax shares, is as in McKay et al. (2016).

4.1 Monetary Policy

We now consider two monetary policy experiments. First, a one-time conventional expansionary monetary policy shock and second, a forward guidance shock that is announced today to take place k periods in the future. In particular, we assume that the monetary authority announces in period 0 to decrease the real interest rate by 10 basis points in period k and keeps it at its steady state value in all other periods. We follow Farhi and Werning (2019) and McKay et al. (2016) and assume that the government debt level remains constant, $B_t^G = \bar{B}^G$.

Figure 1: Monetary Policy and Forward Guidance



Note: This figure shows the response of total output in period 0 to anticipated one-time monetary policy shocks occurring at different horizons k, relative to the response in the representative agent model under rational expectations (normalized to 1). The blue-dashed line shows the results for the behavioral HANK model, the orange-dotted line for the rational HANK model with countercyclical inequality and the black-dashed-dotted line for the rational HANK model with procyclical inequality.

Figure 1 shows on the vertical axis the response of output in period 0, dY_0, to an announced real rate change implemented in period k (horizontal axis). The white horizontal line represents the response in the rational RANK model (normalized to 1). The constant response in RANK is

of high MPC households is exactly the same as in Section 3.

a consequence of the assumption that forward guidance is implemented through changes in the real rate.

The blue-dashed line shows the results for the behavioral HANK model. We see that contemporaneous monetary policy has stronger effects than in RANK and the amplification is roughly 20%.[21] The intuition is the same as in the tractable model: as households with higher MPCs tend to be more exposed to aggregate income changes, monetary policy is amplified through indirect general equilibrium effects. Turning again to an AR(1)-process with a persistence of 0.6, we find that indirect effects account for 61% of the total effect in the quantitative behavioral HANK and, thus, for a large part of the transmission consistent with the findings in Kaplan et al. (2018). At the same time, the behavioral HANK model does not suffer from the forward guidance puzzle, as shown by the decline in the blue-dashed line. Interest rate changes announced to take place in the future have relatively weaker effects on contemporaneous output and the effects decrease with the horizon.[22]

In contrast, the orange-dotted and the black-dashed-dotted lines highlight the tension in rational HANK models. When households with high MPCs tend to be more exposed to aggregate income fluctuations—which corresponds to $\chi > 1$ in the tractable model and which we refer to as the *countercyclical* HANK model—contemporaneous monetary policy is as strong as in the behavioral model. But with rational expectations the amplification through indirect effects extends intertemporally and results in an aggravation of the forward guidance puzzle. Indeed, we see from the orange-dotted line that the farther away the announced interest rate change takes place, the stronger the response of output today.

When, in contrast to the data, households with higher MPCs tend to be less exposed to aggregate income fluctuations—$\chi < 1$ in the tractable model and which we refer to as the *procyclical* HANK model—the rational HANK model resolves the forward guidance puzzle (see McKay et al. (2016)). But the procyclical HANK model is unable to generate amplification of contemporaneous monetary policy (see black-dashed-dotted line). Turning to an AR(1)-process, this model implies that indirect effects account only for 12% of the monetary transmission. In addition, this model has quite different policy implications, as we will see in Section 5.

4.2 Sticky Wages

As in the tractable model, we now show that our results hold when we assume that prices are fully flexible but wages are sticky. We again abstract from monopolistic competition of firms, so that prices are set to marginal costs.

Labor hours $N_{i,t}$ are determined by union labor demand. Each worker provides $N_{i,k,t}$ hours

[21] Patterson (2023) also estimates that the unequal exposure of households leads to a 20% amplification compared to an equal exposure benchmark.

[22] We find that for our baseline calibration the behavioral HANK model resolves the forward guidance puzzle as long as $\bar{m} < 0.93$.

of work to a continuum of unions indexed by k. Each union aggregates efficient units of work into a union-specific task. A competitive labor packer then packages these tasks into aggregate employment services according to a CES technology and sells these services to final goods firms at price W_t. We assume that there are quadratic utility costs of adjusting the nominal wage W_{kt}. A union sets a common wage $W_{k,t}$ per efficient unit for each of its members. In doing so, the union trades-off the marginal disutility of working given average hours against the marginal utility of consumption given average consumption as in Wolf (2021). The union then calls upon its members to supply hours according to a specific allocation rule: in stationary equilibrium all households supply the same amount of hours. Outside the stationary equilibrium, we follow Auclert and Rognlie (2020) and assume the allocation rule

$$N_{i,t} = Y_t \frac{(e_{i,t})^{\zeta \log \frac{Y_t}{Y}}}{\mathbb{E}[e^{1+\zeta \log \frac{Y_t}{Y}}]}.$$

If $\zeta = 0$, all households supply the same amount of labor in each period. Assuming $\zeta < 0$, however, implies that the labor supply of less productive households responds more sensitively to changes in aggregate output Y_t and thus, implies countercyclical income risk. We set $\zeta = -1.2$. We further match the MPCs of 0.16 by setting the debt-to-annual-GDP level to 65%. We discretize the $e_{i,t}$ process into 11 states and as in the sticky-price model impose that only the above-median-income households pay taxes.

All in all, our setup leads to a wage Philips curve given by:

$$\pi_t^W = \kappa\Big(v^{'}(N_t) - (\epsilon_n - 1)/\epsilon_n(1 - \tau_t)\frac{W_t}{P_t}u^{'}(C_t)\Big) + \beta\pi_{t+1}^W, \tag{25}$$

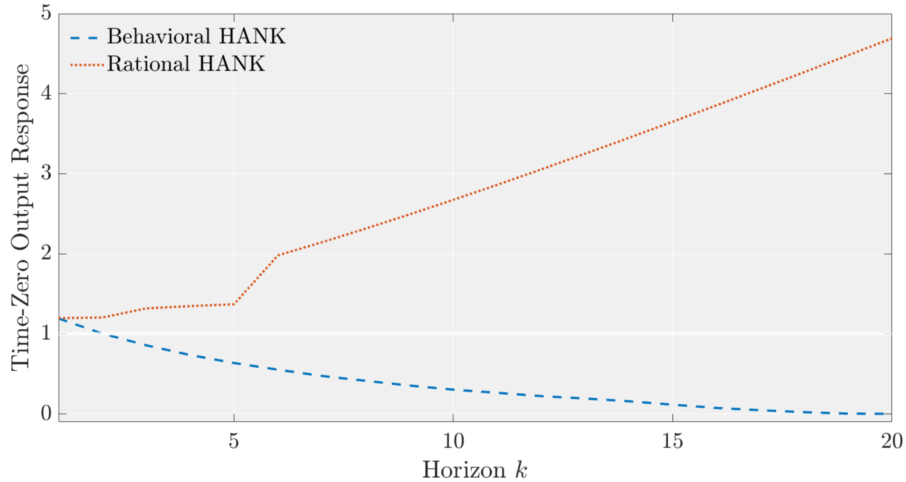
where $\epsilon_n = 11$ is the elasticity of substitution between differentiated labor supply and $\kappa = 0.1$ is the slope of the wage Philips Curve.

Figure 2 shows the effects of conventional monetary policy shocks and of forward guidance shocks on output at time 0 in our sticky-wage behavioral HANK model (blue-dashed line) as well as for the rational HANK model with sticky wages (orange-dotted line). We see that our results are robust. Whereas in both models, contemporaneous monetary policy is to a large share transmitted through indirect effects, the behavioral HANK model rules out the forward guidance puzzle whereas it is aggravated in the rational HANK model compared to the representative agent model.

4.3 Heterogeneous Cognitive Discounting

So far, we have assumed that all households exhibit the same degree of rationality. Yet, as we showed in Section 2, while underreaction is found across all income groups, the data suggests that higher income households deviate somewhat less from rational expectations. To model this, we assume that a household's rationality is a function of her productivity level e: $\bar{m}(e = e_L) = 0.8$, $\bar{m}(e = e_M) = 0.85$ and $\bar{m}(e = e_H) = 0.9$.

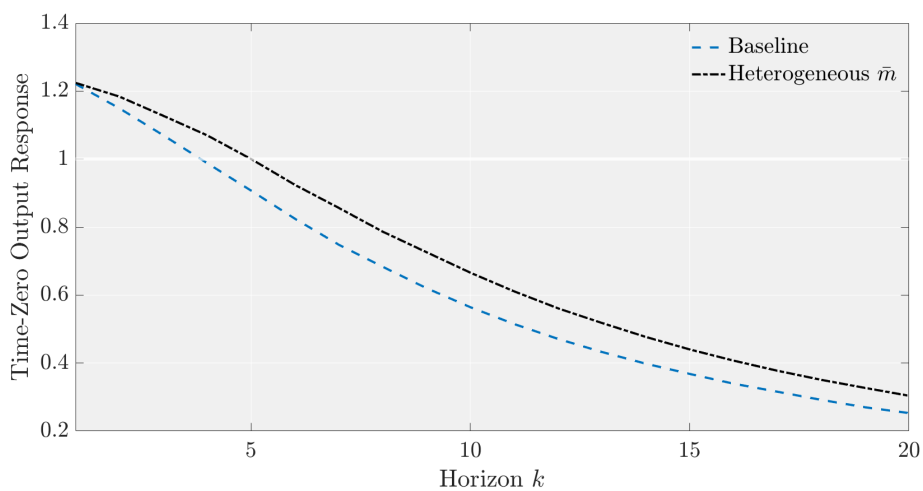
Figure 2: Sticky Wages



Note: This figure shows the response of total output in period 0 to anticipated i.i.d. monetary policy shocks occurring at different horizons k for the case in which prices are flexible and wages are sticky.

This parameterization serves three purposes: first, the lowest-productivity households exhibit the largest deviation from rational expectations and the degree of rationality increases monotonically with productivity. Second, the average degree of bounded rationality remains 0.85 such that we can isolate the effect of heterogeneity in bounded rationality from its overall level. And third, this is a rather conservative parameterization—both in terms of the degree of heterogeneity and in the level of rationality—compared to the results in the data which points towards lower levels of rationality across all households and less dispersion. We discuss an alternative calibration—one in which a subgroup of households is fully rational—in Appendix E.

Figure 3: Heterogeneous $\bar{m}$ and Monetary Policy



Note: This figure shows the response of total output in period 0 to anticipated i.i.d. monetary policy shocks occurring at different horizons k for the baseline calibration with $\bar{m} = 0.85$ for all households (blue-dashed line) and for the model in which households differ in their levels of cognitive discounting (black-dashed-dotted line).

Figure 3 compares the model with heterogeneous degrees of bounded rationality (black-dashed-dotted line) to our baseline quantitative behavioral HANK model (blue-dashed line) for the same monetary policy experiments as above. The effect of a contemporaneous monetary policy shock is practically identical across the two scenarios consistent with the insight that amplification of contemporaneous monetary policy is barely affected by the degree of rationality. At longer horizons, however, monetary policy is more effective in the economy in which households differ in their degrees of rationality.

There are two competing effects: first, high productivity households are now more rational such that they react stronger to announced future changes in the interest rate compared to the baseline which increases the effectiveness of forward guidance. Second, low productivity households are less rational which tends to dampen the effectiveness of forward guidance. Yet, a large share of low productivity households are at their borrowing constraint and, thus, they do not directly react to future changes in the interest rate anyway while most of the high productivity households are unconstrained. Hence, the first effect dominates and forward guidance is more effective compared to the baseline model. Overall, however, the differences across the two calibrations are rather small. As we show in Appendix E, even when the highest productivity households are fully rational the forward guidance puzzle is resolved and the effects of forward guidance vanish quite quickly with the horizon.

4.4 Non-Rational Expectations about Idiosyncratic Shocks

Up to now, we have assumed that households are fully rational with respect to aggregate shocks but that households are perfectly rational with respect to their own idiosyncratic risk. Yet, recent evidence suggests that professional forecasters (Kučinskas and Peters (2022)) and firms (Born et al. (2022)) show patterns of *overreaction* with respect to individual shocks. We now show how simultaneous underreaction to aggregate shocks and overreaction to idiosyncratic shocks affect our results.

To do so, we extend our model and now assume that households overpredict the persistence of their idiosyncratic risk, that is $\tilde{\rho}_e > \rho_e$, where ρ_e denotes the persistence of their actual risk and $\tilde{\rho}_e$ denotes the perceived persistence. In particular, we consider $\tilde{\rho}_e = 0.976$ (instead of $\rho_e = 0.966$). We verify that this implies overreaction to individual news by running the Coibion and Gorodnichenko (2015) regressions (see equation (7)) on model-simulated data with the idiosyncratic productivity being the forecasted variable, $e_t = x_t$. This yields a regression coefficient that is negative, indicating overreaction (Bordalo et al. (2020)).[23]

The orange-dashed line in Figure 4 shows the effects of monetary policy and forward guidance shocks for our extended model where households also misperceive their individual risk. The main

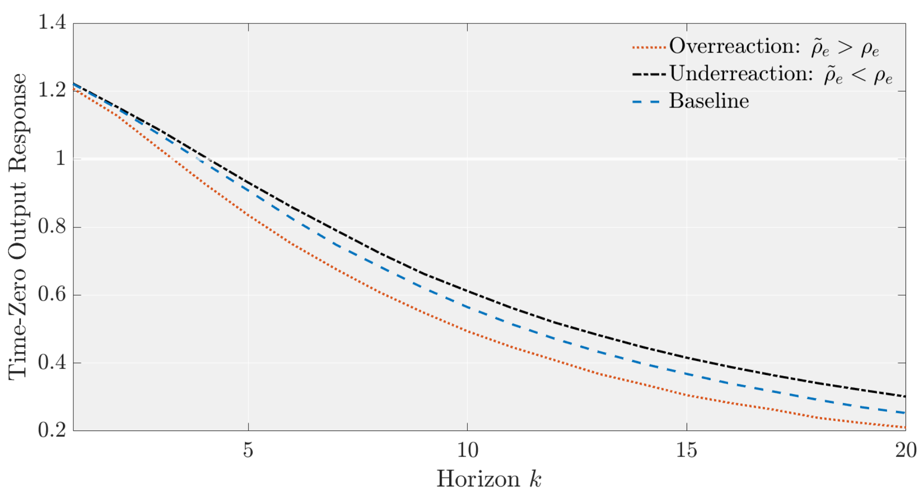
[23] We simulate the model for 1000 households over 500 periods and do this 100 times. For each simulation, we estimate b^{CG} using one-quarter ahead forecast errors. The mean estimate is -0.01 in the case with $\tilde{\rho}_e = 0.976$. Thus, overreaction is quite small, consistent with the empirical findings in Afrouzi et al. (2022) for highly persistent AR(1) processes.

take-away is that our results are very robust to this extension: monetary policy is amplified through indirect, general-equilibrium effects and the effectiveness of forward guidance decreases with the horizons. Quantitatively, the results are pretty similar to our baseline model (blue-dashed line), in which households are fully rational with respect to their idiosyncratic risk. Qualitatively, the effectiveness of forward guidance becomes weaker in the model with overreaction to idiosyncratic shocks. To understand this, recall the aggregate IS equation in our tractable model, equation (18), but replace the actual idiosyncratic risk $1-s$ with the perceived risk $1-\tilde{s}$:

$$\widehat{y}_t = \tilde{\psi}_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma} \widehat{r}_t,$$

with $\tilde{\psi}_f \equiv \bar{m}\left[1+(\chi-1)\frac{1-\tilde{s}}{1-\lambda\chi}\right]$. Overreaction is captured by $1-\tilde{s} < 1-s$. We can thus directly see that $\tilde{\psi}_f < \psi_f$ when $\chi > 1$, which dampens the effects of forward guidance. Intuitively, even though unconstrained households know that they will benefit more from an expansionary forward guidance shock in case they become hand-to-mouth, they underpredict the probability of becoming hand-to-mouth, and hence, forward guidance is further dampened. The same mechanism is at work in our full model, which explains the results in Figure 4.

Figure 4: Non-Rational Expectations about Idiosyncratic Shocks



Note: This figure shows the response of total output in period 0 to anticipated one-time monetary policy shocks occurring at different horizons k, relative to the response in the representative agent model under rational expectations (normalized to 1). The orange-dotted line shows the overreaction case with $\tilde{\rho}_e = 0.976 > \rho_e$, the black-dashed-dotted line the underreaction case with $\tilde{\rho}_e = 0.95 < \rho_e$, and the blue-dashed line our baseline model in which households correctly perceive the persistence of their idiosyncratic risk, $\tilde{\rho}_e = \rho_e = 0.966$.

For completeness, we also consider a case in which households underreact to idiosyncratic shocks. In particular, we set $\tilde{\rho}_e = 0.95$ and, thus, $\tilde{\rho}_e < \rho_e$. The black-dashed-dotted line in Figure 4 shows that also in this case, our main results remain very robust and the quantitative differences to our baseline model are small. Qualitatively, the effectiveness of forward guidance is now a bit less dampened consistent with our intuition above: as households underestimate the persistence of their idiosyncratic productivity, they are more eager to precautionary save and, thus, they react stronger to the relaxation in their precautionary savings risk induced by news about future interest rate decreases. We also run a robustness check with an even more extreme degree of underreaction towards individual news ($\tilde{\rho}_e = 0.85$, not shown) and we find that even in this case, the forward guidance puzzle remains solved.

The take-aways from this section are that our results are robust to allowing for households to deviate from rational expectations about their idiosyncratic risk and that in the probably empirically-relevant case of overreaction, the effects of forward guidance are even further dampened.

4.5 Further Results

Stability at the effective lower bound. To test the stability of the model at the effective lower bound—fact (iv)—we consider a shock to the discount factor that pushes the economy to the ELB for 8 periods, in the behavioral and the rational model. After that the shock jumps back to its steady state value. Consistent with the tractable model, the recession in the rational model is substantially more severe. While output drops on impact by 5% in the behavioral model, it drops by 10% in the rational model (see Appendix E.2 for details).

Unequal exposure: more extreme calibration. In our baseline calibration, we target the finding from Patterson (2023) that a linear regression of households' income elasticity to GDP on their MPC yields a coefficient of 1.33. In Appendix E.1, we show that our results remain robust when we target a more extreme coefficient of 2. In this case, the initial amplification through indirect effects becomes stronger, but the model still resolves the forward guidance puzzle and, thus, is able to account for fact (i) - (iv) simultaneously.

5 Policy Implications of Inflationary Supply Shocks

Having established that the behavioral HANK model is consistent with recent facts about the transmission and effectiveness of monetary policy, we now use the model to revisit the policy implications of inflationary supply shocks. We uncover a novel amplification channel of these shocks that is absent in existing models as it arises due to the interaction of the unequal exposure of households to monetary policy and the behavioral friction, and thus, exactly through the model ingredients that allow the model to simultaneously account for facts (i)-(iv).

Many advanced economies have recently experienced a dramatic surge in inflation and at least part of this is attributed to disruptions in production, such as supply-chain "bottlenecks" (see, e.g., di Giovanni et al. (2022)). We model these disruptions as a negative total factor productivity (TFP) shock. Production of intermediate-goods firm j is now given by $Y_t(j) = A_t N_t(j)$, where A_t is total factor productivity following an AR(1)-process, $A_t = (1 - \rho_A)\bar{A} + \rho_A A_{t-1} + \varepsilon_t^A$, and ε_t^A is a zero-mean i.i.d. shock, $\bar{A}$ the steady-state level of TFP and ρ_A the persistence of A_t which we set to $\rho_A = 0.9$. Each firm can adjust its price with probability 0.15 in a given quarter and we assume that firms have rational expectations to fully focus on the role of bounded rationality on the household side (we discuss the case with behavioral firms later).

Government debt is time-varying and total tax payments, T_t, follow the debt feedback rule, $T_t - \bar{T} = \vartheta \frac{B^G_{t+1} - \bar{B}^G}{\bar{Y}}$, where we set $\vartheta = 0.05$. We start with the case in which monetary policy follows a simple Taylor rule (10) with an inflation coefficient of 1.5. Later on, we discuss the case in which monetary policy follows a strict inflation-targeting rule and implements a zero inflation rate in all periods.

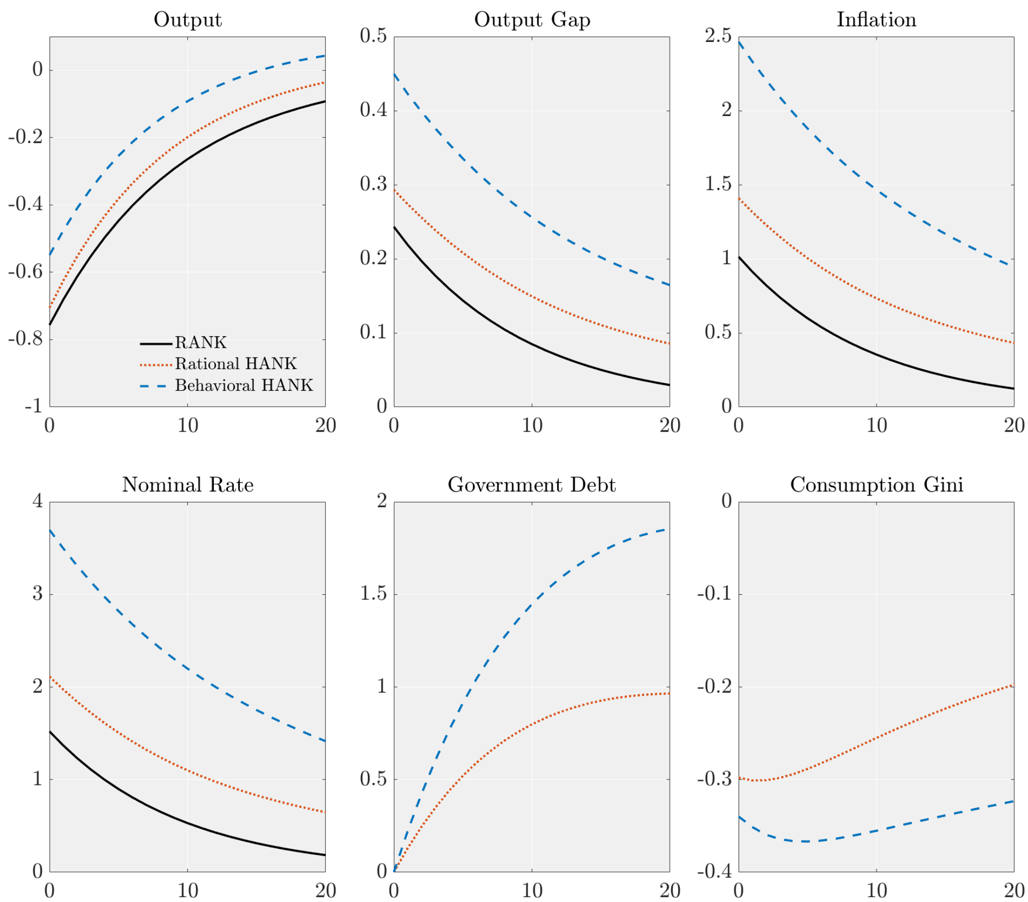
The size of the shock is such that output in the model with fully-flexible prices, complete markets and rational expectations—what we from now on call *potential output*—decreases by 1% in terms of deviations from its steady state. We normalize the leisure parameter in the complete markets, flexible price model such that it has the same steady state output as our behavioral HANK model. The *output gap* is then defined as the difference between actual output and potential output divided by steady state output.

Figure 5 shows the impulse-response functions of output, the output gap, inflation, nominal interest rates, government debt and the consumption Gini index as a measure of inequality after the negative supply shock. The blue-dashed lines show the responses in the behavioral HANK model, the orange-dotted lines in the rational HANK model, and the black-solid lines in RANK. We assume government debt to be constant in RANK.[24]

Qualitatively, the impulse responses are the same across all models: in response to the supply shock, monetary policy increases the nominal interest rate, which pushes down output. Yet output falls by less than potential output, leading to positive output gaps which pushes up inflation. Yet, *quantitatively* there are large differences across the models. In particular, the increase in inflation is roughly 2.5 times as large in the behavioral HANK model compared to RANK and 1.7 times as strong as in the rational HANK model even though the (nominal and real) interest rate increases most strongly in the behavioral HANK model. The reason is a novel amplification channel due to household heterogeneity, cognitive discounting and the interaction of the two. The positive output gap increases wages and decreases profits relative to the outcome without nominal rigidities in the same way as expansionary policy shocks in Sections 3 and 4 do. This redistributes on average towards lower income and higher MPC households which further increases the output gap and inflation. In addition, the higher expected real interest rates in response to the negative supply shock lead to a negative deviation of expected consumption from its stationary equilibrium counterpart. In the behavioral HANK model, households cognitively discount the expected higher interest rates and, hence, their consumption expectations decrease by less. As a result, households decrease today's consumption by less compared to fully rational households. This further increases the output gap which amplifies the redistribution to high MPCs households which again amplifies the increase in the output gap until the economy ends up in an equilibrium with a higher output

[24] As we implement taxes such that they do not show up in the first-order conditions of household, in the RANK version of our model Ricardian equivalence holds and, thus, the path of debt does not matter anyway.

Figure 5: Inflationary supply shock: Taylor rule



Note: This figure shows the impulse responses after a productivity shock for the case where monetary policy follows a Taylor rule. Output and the output gap are shown as percentage deviations from steady state output, the nominal interest rate and inflation as annualized percentage points and the government debt level as percentage point deviations of the debt-per annual-GDP level. The lower-right figure shows the change in the consumption Gini index as a percentage deviation from the stationary equilibrium.

gap and higher inflation.

While inflation and the output gap increase substantially, consumption inequality decreases both in the rational as well as in the behavioral HANK model and it decreases even more in the behavioral model (see lower-right panel in Figure 5). While higher interest rates redistribute to relative consumption-rich households, this effect on consumption inequality is dominated by the increase in the output gap which redistributes to relatively consumption-poor households. Finally, the higher real interest rates increase the cost of government debt which is (partly) financed by issuing more debt. Thus, the government debt level increases, especially in the behavioral HANK model where the increase in real interest rates is larger.

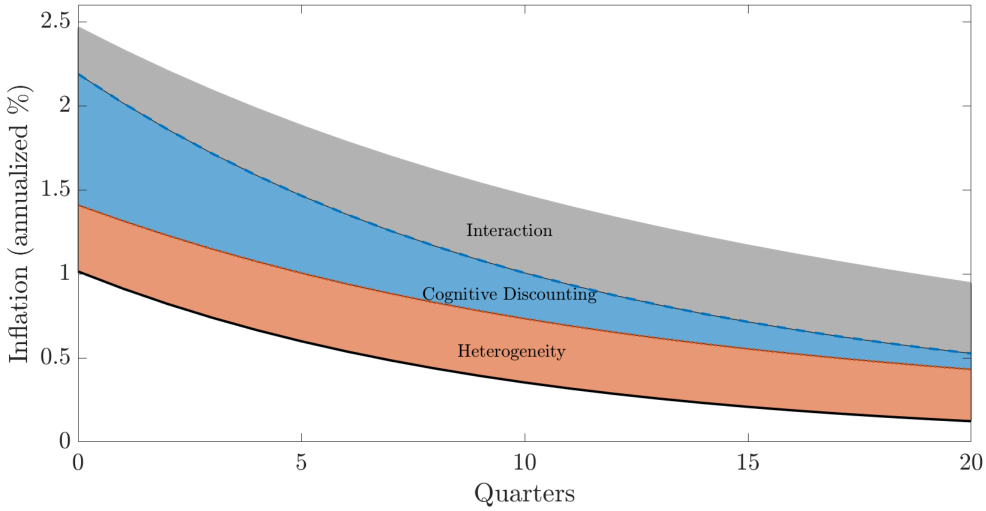
Given the larger sensitivity of inflation to supply shocks due to this novel amplification channel, our model may hence offer a (partial) explanation for why many advanced economies have seen large inflation increases following the Covid-19 pandemic. In particular, our model predicts that when the shock redistributes towards high-MPC households and when the central bank's response to inflation is underpredicted, inflationary supply shocks can lead to a substantial increase in inflation.

Decomposition of the amplification channel. How much of the additional inflation increase in the behavioral HANK model compared to RANK is due to the underlying heterogeneity, how much is due to cognitive discounting and how much is due to the interaction of the two? Figure 6 decomposes the amplification channel into these three components. It shows the additional inflation increase in the behavioral HANK model compared to the inflation increase in the RANK benchmark and its components. The black-solid line shows the inflation response in RANK. The orange-shaded area denotes the additional inflation increase that arises solely due to household heterogeneity. The blue-shaded area the fraction of the overall increase due to cognitive discounting alone. Thus, the gray-shaded area captures the additional inflation increase that is due to the interaction of household heterogeneity and cognitive discounting.

Under our baseline calibration, this complementarity amounts to about 27% on impact of the inflation response in RANK (the inflation response in RANK is 1 percentage point on impact). As the additional increase in the behavioral HANK is about 1.45 percentage points, the complementarity explains about 19% of the *additional increase*. Cognitive discounting alone accounts for 27% while household heterogeneity accounts for 54% of the amplification over RANK. Figure 15 in Appendix F.1 considers an alternative calibration of the discounting parameter where we set it to 0.6 and thus the lower bound of the empirical estimates instead of the upper bound. In this case, inflation increases more than 3.5 times as much as in the RANK model with the interaction between household heterogeneity and cognitive discounting accounting for 1/3 of the initial amplification and the interaction itself accounts for a larger share of the overall additional increase than the underlying heterogeneity.

Supply vs. demand shocks. The fact that the behavioral HANK model amplifies persistent supply shocks more than the rational HANK model is in contrast to persistent demand shocks. While both the underlying heterogeneity and bounded rationality amplify persistent supply shocks, these both model features work in opposite direction in response to persistent demand shock. For example, in response to an expansionary monetary policy shock, the heterogeneous exposure of households amplifies the effects of the shock. The high-MPC households benefit more strongly from the shock which triggers an amplification of the shock, as discussed extensively in Sections 3 and 4. Cognitive discounting, however, would dampen the effect because households would discount the persistent decrease in the interest rate. As a result, persistent demand shocks are less strong in the behavioral HANK model compared to the rational HANK model: an expansionary monetary shock of 1 percentage point with a persistence of 0.6 increases inflation on impact by by 1.24pp. (annualized) in the behavioral HANK while it does by 1.44pp. in the rational HANK model.

Figure 6: Decomposition of the Additional Inflation Increase



Note: This figure shows the decomposition of the additional inflation increase in the behavioral HANK model compared to the rational RANK model. The orange-shaded area represents the additional increase that is solely due to the heterogeneous exposure of households, the blue area the increase due to cognitive discounting and the gray area the additional increase that is due to the interaction of heterogeneity and cognitive discounting.

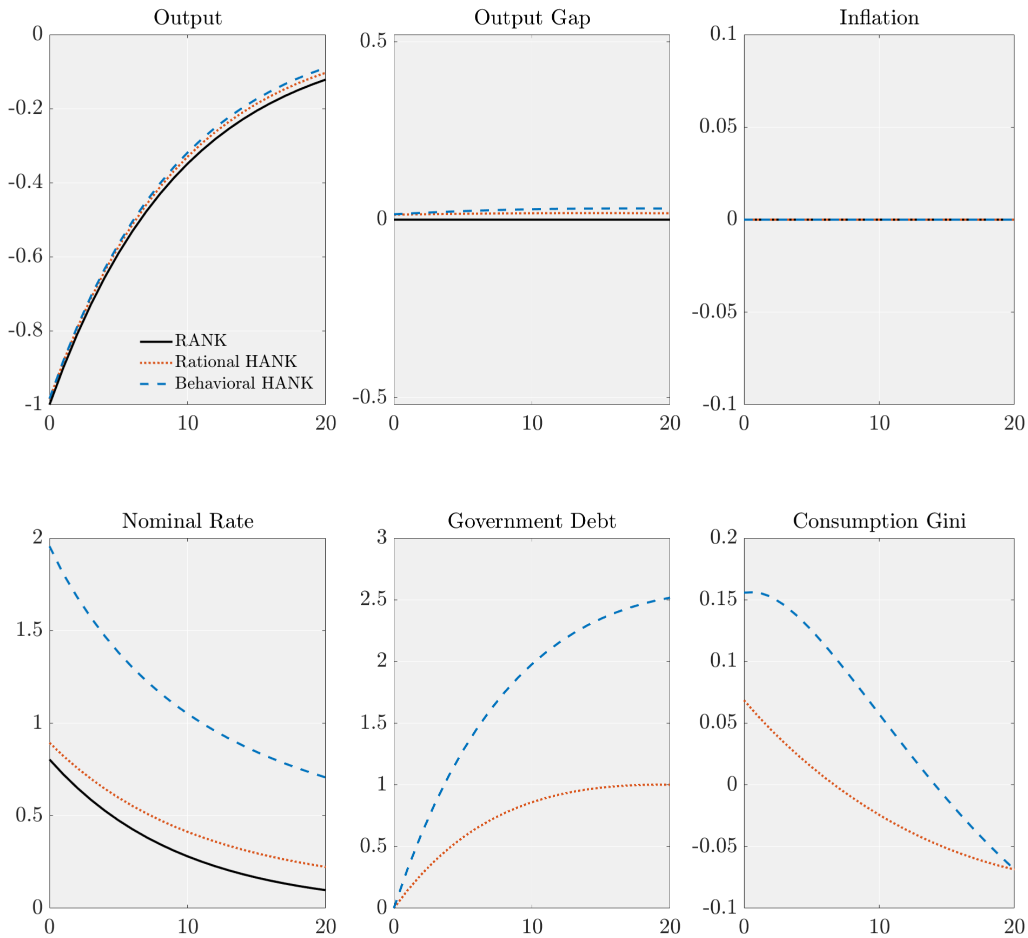
Strict inflation targeting. What if monetary policy reacts more hawkish to inflation? Figure 7 shows the limiting case, in which monetary policy follows a strict inflation targeting rule and, hence, keeps inflation at zero at all times. We see that the output responses are almost indistinguishable across the two models and practically identical to the fall in potential output such that the output gap is essentially zero.

Yet, the reaction of monetary policy differs significantly across the two models. The nominal (and real) interest rate in the behavioral HANK model increases twice as much on impact as in the rational HANK model. The reason is that behavioral households cognitively discount the future higher interest rates that they expect due to the persistence of the shock. Hence, these expected higher future rates are less effective in stabilizing inflation today. Thus, to induce zero inflation in every period, monetary policy needs to increase interest rates by more than in the rational HANK model, in which the expected future interest rate hikes are very powerful. As this line of reasoning applies in each period, the interest rate in the behavioral HANK model remains above the interest rate in the rational model.

Raising interest rates increases the cost of debt for the government which it finances in the short run by issuing additional debt. The bottom-middle panel in Figure 7 shows that government debt in the behavioral model increases by more than twice as much as in the rational model and by more than when monetary policy follows a simple Taylor rule. Thus, the fiscal footprint of monetary policy is larger because monetary policy needs to respond more strongly to counteract the inflationary pressures in the behavioral model.

On top of the stronger increase in government debt and interest rates, consumption inequality

Figure 7: Inflationary supply shock: strict inflation-targeting



Note: This figure shows the impulse responses after a TFP shock that decreases potential output by 1% in the inflation-stabilizing monetary policy regime. Output and the output gap are shown as percentage deviations from steady state output, the nominal interest rate and inflation as annualized percentage points and the government debt level as percentage point deviations of the debt-per-annual GDP level. The lower-right figure shows the change in the consumption Gini index as a percentage deviation from the stationary equilibrium.

increases more strongly in the behavioral model compared to the rational model. The reason is that along the wealth distribution, increases in the real interest rate redistribute to wealthier households and, hence, to households who already tend to have a higher consumption level. As the increases in the real interest rate are higher in the behavioral HANK model, these redistribution effects are more pronounced. Because monetary policy fully stabilizes inflation and the output gap, dividends and wages fall by the same relative amount after the productivity shock, such that each household's labor and dividend income falls by the same amount. Hence, the redistribution channels present in Sections 3 and 4 after policy shocks are muted here. Put differently, monetary policy turns off the amplification mechanism that works through the unequal exposure of households when implementing zero inflation.

Overall, our model suggests that accounting for facts (i)-(iv) simultaneously has important implications for policy. In particular, there is a strong trade off for monetary policy following an inflationary supply shock. Simply following a Taylor rule and thus, not responding very aggressively to the inflationary pressures, can lead to a significant increase in inflation through the mutual reinforcement of households' unequal exposure to the overheating of the economy as well as households' cognitive discounting of the monetary authority's future response to inflation. Counteracting these forces and implementing a zero inflation rate, however, requires a much stronger monetary policy response which, in turn, leads to a strong increase in government debt and inequality.

Comparison to the procyclical HANK model. One of the reasons why the behavioral HANK model amplifies supply shocks is that it is less responsive to expected future interest rates. A natural question is then: how do its policy implications compare to those derived in rational HANK models that are calibrated to resolve the forward guidance puzzle? As shown in Section 4, when all households receive an equal share of the dividends, the rational model can resolve the forward guidance puzzle (McKay et al. (2016)). This implies that households with high MPCs benefit less from income increases induced by monetary policy, thereby violating fact (ii).

Figure 16 in the appendix shows that this "procyclical" rational HANK model predicts a much weaker response of inflation to the same supply shock in the case of a standard Taylor rule. The reason is that now the positive output gap redistributes on average to high-income and low MPC households which further dampens aggregate demand. In other words, this model features a dampening channel compared to RANK after supply shocks instead of an amplification channel as in the behavioral HANK model.[25]

The two models also differ in terms of their cross-sectional implications: in the procyclical HANK model, consumption inequality increases strongly whereas it decreases in the behavioral HANK model.

Behavioral firms. In Appendix F.3, we discuss the case in which firms cognitively discount the future in the same way as households. The increase in inflation when monetary policy follows a Taylor rule is somewhat muted whereas the increase in the output gap is amplified compared to the case in which firms are rational. The reason is that firms discount the increase in their future marginal costs and thus increase their prices not as strongly. According to the Taylor rule this then leads to a smaller increase in interest rates so that households consume more, leading to an increase in demand and thus, the output gap.

Cost-push shocks. So far, we have focused on the inflationary pressure coming from negative TFP shocks. We show in Appendix F.4 that if the inflationary pressure comes from a cost-push shock instead, the monetary and fiscal implications are very similar. Inflation and the output gap increases much more in the behavioral HANK model compared to the rational HANK model

[25] Another take-away is that for a given persistent demand shock, the behavioral HANK model and a recalibrated version of the procyclical HANK model could be observationally equivalent in terms of the output and inflation response. Yet, these two models then differ drastically after supply shock.

although interest rates increases more. Accordingly, if the central bank wants to fully stabilize inflation, it needs to raise interest rates much more strongly in the behavioral HANK model than in the rational HANK model to fully stabilize inflation. This pushes up the government debt level, especially in the behavioral HANK model.

6 Conclusion

In this paper, we develop a new framework for business-cycle and policy analysis: the behavioral HANK model. To arrive at our framework, we introduce bounded rationality in the form of cognitive discounting and household heterogeneity into a New Keynesian model. The model can account for recent empirical findings on the transmission mechanisms of monetary policy. In particular, households with higher marginal propensities to consume tend to be more exposed to changes in aggregate income that are induced by monetary policy, leading to an amplification of conventional monetary policy through indirect effects. Simultaneously, the model rules out the forward guidance puzzle and remains stable at the effective lower bound. The model thus overcomes a tension in existing models with household heterogeneity: when accounting for the underlying heterogeneity, these models tend to aggravate the forward guidance puzzle and the instability issues at the lower bound. Both, bounded rationality and household heterogeneity, are crucial to arrive at our results.

Simultaneously accounting for these facts matters greatly for the model's policy implications. In particular, we uncover a new amplification mechanism of inflationary supply shocks through cognitive discounting and the unequal exposure of households. After a negative productivity shock the behavioral HANK model predicts a substantially larger inflation increase. If the monetary authority wants to stabilize inflation after such an inflationary supply shock, it needs to hike the nominal interest rate much more strongly than under rational expectations which leads to a strong increase in government debt and inequality.

Given its consistency with empirical facts about the transmission of monetary policy, the behavioral HANK model provides a natural laboratory for both business-cycle and policy analysis. Our framework can also easily be extended along many dimensions, some of which we have explored in the paper, whereas others are left for future work.

References

Acharya, S. and K. Dogra (2020): "Understanding HANK: Insights from a PRANK," *Econometrica*, 88, 1113–1158.

Adam, K., O. Pfaeuti, and T. Reinelt (2022): "Subjective Housing Price Expectations, Falling Natural Rates and the Optimal Inflation Target," .

Afrouzi, H., S. Y. Kwon, A. Landier, Y. Ma, and D. Thesmar (2022): "Overreaction in expectations: Evidence and theory," *The Quarterly Journal of Economics (forthcoming).*

Ampudia, M., D. Georgarakos, J. Slacalek, O. Tristani, P. Vermeulen, and G. Violante (2018): "Monetary policy and household inequality," .

Andrade, P., G. Gaballo, E. Mengus, and B. Mojon (2019): "Forward guidance and heterogeneous beliefs," *American Economic Journal: Macroeconomics*, 11, 1–29.

Angeletos, G.-M., F. Collard, and H. Dellas (2020): "Business-cycle anatomy," *American Economic Review*, 110, 3030–70.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2021): "Imperfect macroeconomic expectations: Evidence and theory," *NBER Macroeconomics Annual*, 35, 1–86.

Angeletos, G.-M. and C. Lian (2018): "Forward guidance without common knowledge," *American Economic Review*, 108, 2477–2512.

——— (2023): "Dampening general equilibrium: incomplete information and bounded rationality," in *Handbook of Economic Expectations*, Elsevier, 613–645.

Auclert, A. (2019): "Monetary policy and the redistribution channel," *American Economic Review*, 109, 2333–67.

Auclert, A., B. Bardóczy, and M. Rognlie (2021): "MPCs, MPEs, and multipliers: A trilemma for New Keynesian models," *The Review of Economics and Statistics*, 1–41.

Auclert, A. and M. Rognlie (2020): "Inequality and aggregate demand," .

Auclert, A., M. Rognlie, and L. Straub (2018): "The intertemporal keynesian cross," .

——— (2020): "Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model," Tech. rep., National Bureau of Economic Research.

Bhandari, A., J. Borovička, and P. Ho (2019): "Survey data and subjective beliefs in business cycle models," .

Bilbiie, F. O. (2008): "Limited asset markets participation, monetary policy and (inverted) aggregate demand logic," *Journal of economic theory*, 140, 162–196.

——— (2020): "The new Keynesian cross," *Journal of Monetary Economics*, 114, 90–108.

——— (2021): "Monetary policy and heterogeneity: An analytical framework," .

Bilbiie, F. O., D. R. Känzig, and P. Surico (2022): "Capital and income inequality: An aggregate-demand complementarity," *Journal of Monetary Economics*, 126, 154–169.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in macroeconomic expectations," *American Economic Review*, 110, 2748–82.

Born, B., Z. Enders, M. Menkhoff, G. J. Müller, and K. Niemann (2022): "Firm Expectations and News: Micro v Macro," .

Broer, T., N.-J. Harbo Hansen, P. Krusell, and E. Öberg (2020): "The New Keynesian transmission mechanism: A heterogeneous-agent perspective," *The Review of Economic Studies*, 87, 77–101.

Broer, T., A. Kohlhas, K. Mitman, and K. Schlafmann (2022): "Information and Wealth Heterogeneity in the Macroeconomy," .

Calvo, G. A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of Monetary*

Economics, 12, 383–398.

Carlson, J. A. and M. Parkin (1975): "Inflation expectations," *Economica*, 42, 123–138.

Christiano, L., M. Eichenbaum, and S. Rebelo (2011): "When is the government spending multiplier large?" *Journal of Political Economy*, 119, 78–121.

Cochrane, J. H. (2018): "Michelson-Morley, Fisher, and Occam: The radical implications of stable quiet inflation at the zero bound," *NBER Macroeconomics Annual*, 32, 113–226.

Coibion, O. and Y. Gorodnichenko (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 105, 2644–78.

Coibion, O., Y. Gorodnichenko, E. S. Knotek II, and R. Schoenle (2022): "Average inflation targeting and household expectations," *Journal of Political Economy Macroeconomics (forthcoming)*.

Coibion, O., Y. Gorodnichenko, L. Kueng, and J. Silvia (2017): "Innocent Bystanders? Monetary policy and inequality," *Journal of Monetary Economics*, 88, 70–89.

Debortoli, D. and J. Galí (2018): "Monetary policy with heterogeneous agents: Insights from TANK models," .

Debortoli, D., J. Galí, and L. Gambetti (2020): "On the empirical (ir) relevance of the zero lower bound constraint," *NBER Macroeconomics Annual*, 34, 141–170.

Del Negro, M., M. Giannoni, and C. Patterson (2015): "The forward guidance puzzle," *Journal of Political Economy Macroeconomics (forthcoming)*.

di Giovanni, J., Kalemli-Özcan, A. Silva, and M. A. Yildirim (2022): "Global Supply Chain Pressures, International Trade, and Inflation," .

D'Acunto, F., D. Hoang, and M. Weber (2022): "Managing households' expectations with unconventional policies," *The Review of Financial Studies*, 35, 1597–1642.

Farhi, E. and I. Werning (2019): "Monetary policy, bounded rationality, and incomplete markets," *American Economic Review*, 109, 3887–3928.

Fuhrer, J. C. and G. D. Rudebusch (2004): "Estimating the Euler equation for output," *Journal of Monetary Economics*, 51, 1133–1153.

Gabaix, X. (2014): "A sparsity-based model of bounded rationality," *The Quarterly Journal of Economics*, 129, 1661–1710.

——— (2017): "Behavioral macroeconomics via sparse dynamic programming," .

——— (2019): "Behavioral inattention," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 261–343.

——— (2020): "A behavioral New Keynesian model," *American Economic Review*, 110, 2271–2327.

Galí, J. (2015): *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*, Princeton University Press.

Guvenen, F., S. Ozkan, and J. Song (2014): "The nature of countercyclical income risk," *Journal of Political Economy*, 122, 621–660.

Hagedorn, M. (2016): *A demand theory of the price level.*

——— (2018): "Prices and inflation when government bonds are net wealth," .

Hagedorn, M., J. Luo, I. Manovskii, and K. Mitman (2019): "Forward guidance," *Journal of Monetary Economics*, 102, 1–23.

Hintermaier, T. and W. Koeniger (2019): "Differences in euro-area household finances and their relevance for monetary-policy transmission," .

Holm, M. B., P. Paul, and A. Tischbirek (2021): "The transmission of monetary policy under the microscope," *Journal of Political Economy*, 129, 2861–2904.

Ilut, C. and R. Valchev (2023): "Economic agents as imperfect problem solvers," *The Quarterly Journal of Economics*, 138, 313–362.

Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

Kučinskas, S. and F. S. Peters (2022): "Measuring under-and overreaction in expectation formation," *The Review of Economics and Statistics*, 1–45.

Kuhn, M., M. Schularick, and U. I. Steins (2020): "Income and wealth inequality in America, 1949–2016," *Journal of Political Economy*, 128, 3469–3519.

Laibson, D., P. Maxted, and B. Moll (2021): "Present bias amplifies the household balance-sheet channels of macroeconomic policy," .

Luetticke, R. (2021): "Transmission of monetary policy with heterogeneity in household portfolios," *American Economic Journal: Macroeconomics*, 13, 1–25.

Mankiw, N. G. (2000): "The savers-spenders theory of fiscal policy," *American economic review*, 90, 120–125.

McKay, A., E. Nakamura, and J. Steinsson (2016): "The power of forward guidance revisited," *The American Economic Review*, 106, 3133–3158.

——— (2017): "The discounted euler equation: A note," *Economica*, 84, 820–831.

McKay, A. and J. F. Wieland (2022): "Forward guidance and durable goods demand," *American Economic Review: Insights*, 4, 106–22.

Michaillat, P. and E. Saez (2021): "Resolving New Keynesian anomalies with wealth in the utility function," *Review of Economics and Statistics*, 103, 197–215.

Mumtaz, H. and A. Theophilopoulou (2017): "The impact of monetary policy on inequality in the UK. An empirical analysis," *European Economic Review*, 98, 410–423.

Patterson, C. (2023): "The matching multiplier and the amplification of recessions," *American Economic Review*, 113, 982–1012.

Ravn, M. O. and V. Sterk (2021): "Macroeconomic fluctuations with HANK & SAM: An analytical approach," *Journal of the European Economic Association*, 19, 1162–1202.

Roth, C., M. Wiederholt, and J. Wohlfart (2021): "The Effects of Forward Guidance: Theory with Measured Expectations," .

Rouwenhorst, K. G. (1995): "Asset pricing implications of equilibrium business cycle models," 294–330.

Samarina, A. and A. D. Nguyen (2019): “Does monetary policy affect income inequality in the euro area?” .

Slacalek, J., O. Tristani, and G. L. Violante (2020): “Household balance sheet channels of monetary policy: A back of the envelope calculation for the euro area,” *Journal of Economic Dynamics and Control*, 115, 103879.

Storesletten, K., C. I. Telmer, and A. Yaron (2004): “Cyclical dynamics in idiosyncratic labor market risk,” *Journal of political Economy*, 112, 695–717.

Werning, I. (2015): “Incomplete markets and aggregate demand,” .

Wiederholt, M. (2015): “Empirical properties of inflation expectations and the zero lower bound,” *manuscript, Goethe University.*

Wolf, C. (2021): “Interest Rate Cuts vs. Stimulus Payments: An Equivalence Result,” .

Woodford, M. (2003): “Interest and prices,” .

——— (2019): “Monetary policy analysis when planning horizons are finite,” *NBER Macroeconomics Annual*, 33, 1–50.

Yun, T. (1996): “Nominal price rigidity, money supply endogeneity, and business cycles,” *Journal of monetary Economics*, 37, 345–370.

Appendix

A Analytical Results: Proofs and Details

A.1 Derivation of χ

In Section 3, we stated that

$$\widehat{c}_t^H = \chi \widehat{y}_t, \tag{26}$$

where $\chi \equiv 1 + \varphi \left(1 - \frac{\mu^D}{\lambda}\right)$ is *the* crucial statistic coming from the limited heterogeneity setup. We now show how we arrive at equation (26) from the H-household's budget constraint, optimality conditions and market clearing.

The labor-leisure condition of the H households is given by $(N_t^H)^\varphi = W_t (C_t^H)^{-\gamma}$, and similarly for the U households. As we focus on the steady state with no inequality, we have that in steady state $C = C^H = C^U$ and $N = N^U = N^H$ and market clearing and the production function imply $Y = C = N$, which we normalize to 1.

Log-linearizing the labor-leisure conditions yields $\varphi \widehat{n}_t^H = \widehat{w}_t - \gamma \widehat{c}_t^H$ and $\varphi \widehat{n}_t^U = \widehat{w}_t - \gamma \widehat{c}_t^U$. Since both households work for the same wage, we obtain

$$\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H = \varphi \widehat{n}_t^U + \gamma \widehat{c}_t^U \tag{27}$$

Log-linearizing the market clearing conditions yields $\widehat{n}_t = \lambda \widehat{n}_t^H + (1-\lambda)\widehat{n}_t^U$ and $\widehat{c}_t = \lambda \widehat{c}_t^H + (1-\lambda)\widehat{c}_t^U$, which can be re-arranged as (using $\widehat{y}_t = \widehat{c}_t = \widehat{n}_t$)

$$\widehat{n}_t^U = \frac{1}{1-\lambda}\left(\widehat{y}_t - \lambda \widehat{n}_t^H\right)$$
$$\widehat{c}_t^U = \frac{1}{1-\lambda}\left(\widehat{y}_t - \lambda \widehat{c}_t^H\right).$$

Replacing $\widehat{n}_t^U$ and $\widehat{c}_t^U$ in equation (27) then gives

$$\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H = (\varphi + \gamma)\widehat{y}_t. \tag{28}$$

The budget constraint of H households (accounting for the fact that bond holdings are zero in equilibrium) is given by $C_t^H = W_t N_t^H + \frac{\mu^D}{\lambda} D_t$. In log-linearized terms, we get

$$\widehat{c}_t^H = \widehat{w}_t + \widehat{n}_t^H + \frac{\mu^D}{\lambda}\widehat{d}_t, \tag{29}$$

and using that $\widehat{w}_t = -\widehat{d}_t = \varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H$, we get

$$\widehat{c}_t^H = \left(\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H\right)\left(1 - \frac{\mu^D}{\lambda}\right) + \widehat{n}_t^H. \tag{30}$$

Using (28) to solve for $\widehat{n}_t^H$ and plugging it into (30) yields

$$\widehat{c}_t^H = \widehat{c}_t^H \gamma \left(1 - \frac{\mu^D}{\lambda}\right) + \chi\left(\frac{\varphi+\gamma}{\varphi}\widehat{y}_t - \frac{\gamma}{\varphi}\widehat{c}_t^H\right).$$

Grouping terms, we obtain

$$\widehat{c}_t^H = \chi \widehat{y}_t,$$

with $\chi \equiv 1 + \varphi\left(1 - \frac{\mu^D}{\lambda}\right)$, as stated above.

A.2 Proof of Proposition 1.

When linearizing the model around the steady state, our bounded rationality assumptions imply

$$\mathbb{E}_t^{BR}\left[\widehat{x}_{t+1}\right] = \bar{m}\mathbb{E}_t\left[\widehat{x}_{t+1}\right]. \tag{31}$$

Combining equations (12) and (14) with (31), we have

$$\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^H\right] = \bar{m}\mathbb{E}_t\left[\widehat{c}_{t+1}^H\right] = \bar{m}\chi\mathbb{E}_t\left[\widehat{y}_{t+1}\right]$$
$$\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] = \bar{m}\mathbb{E}_t\left[\widehat{c}_{t+1}^U\right] = \bar{m}\frac{1-\lambda\chi}{1-\lambda}\mathbb{E}_t\left[\widehat{y}_{t+1}\right].$$

Plugging these two equations as well as equation (14) into the Euler equation of unconstrained households (16) yields

$$\frac{1-\lambda\chi}{1-\lambda}\widehat{y}_t = s\bar{m}\frac{1-\lambda\chi}{1-\lambda}\mathbb{E}_t\left[\widehat{y}_{t+1}\right] + (1-s)\bar{m}\chi\mathbb{E}_t\left[\widehat{y}_{t+1}\right] - \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t\pi_{t+1}\right).$$

(Note, that this is exactly the same expression as in section 3.5 but with χ instead of ζ.) Combining the $\mathbb{E}_t\left[\widehat{y}_{t+1}\right]$ terms and dividing by $\frac{1-\lambda\chi}{1-\lambda}$ yields the following coefficient in front of $\mathbb{E}_t\left[\widehat{y}_{t+1}\right]$:

$$\psi_f \equiv \bar{m}\left[s + (1-s)\chi\frac{1-\lambda}{1-\lambda\chi}\right] = \bar{m}\left[1 - 1 + s + (1-s)\chi\frac{1-\lambda}{1-\lambda\chi}\right]$$

$$= \bar{m}\left[1 - \frac{1-\lambda\chi}{1-\lambda\chi} + s + (1-s)\chi\frac{1-\lambda}{1-\lambda\chi}\right] = \bar{m}\left[1 - \frac{1-\lambda\chi}{1-\lambda\chi} + \frac{(1-\lambda\chi)s}{1-\lambda\chi} + (1-s)\chi\frac{1-\lambda}{1-\lambda\chi}\right]$$
$$= \bar{m}\left[1 + (\chi - 1)\frac{1-s}{1-\lambda\chi}\right].$$

Defining $\psi_c \equiv \frac{1-\lambda}{1-\lambda\chi}$ yields the behavioral HANK IS equation in Proposition 1:

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t \pi_{t+1}\right).$$

A.3 Proof of Proposition 2.

We prove here the more general case where forward guidance is implemented as changes in the nominal rather than the real rate and where the supply side is captured by the Phillips Curve (17). The case with real rate changes is a special case of the nominal rate case and can be captured by setting $\kappa = 0$.

The first part of Proposition 2 follows from the fact that amplification is obtained when

$$\psi_c = \frac{1-\lambda}{1-\lambda\chi} > 1,$$

which requires $\chi > 1$, given that we assume throughout $\chi\lambda < 1$.

For the second part, recall how we define the forward guidance experiment (following Bilbiie (2021)). We assume a Taylor coefficient of 0, i.e., $\phi = 0$, such that the nominal interest rate is given by $\widehat{i}_t = \varepsilon_t^{MP}$. Replacing inflation using the Phillips curve (17), i.e., $\pi_t = \kappa\widehat{y}_t$, we can re-write the behavioral HANK IS equation from Proposition 1 as

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\left(\varepsilon_t^{MP} - \kappa \mathbb{E}_t \widehat{y}_{t+1}\right)$$
$$= \left(\psi_f + \psi_c \frac{1}{\gamma}\kappa\right)\mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\varepsilon_t^{MP}$$

The forward guidance puzzle is ruled out if and only if

$$\left(\psi_f + \psi_c \frac{1}{\gamma}\kappa\right) < 1,$$

which is the same as:

$$\bar{m}\delta + \frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\kappa < 1.$$

A.4 Proof of Proposition 3.

Replacing $\widehat{i}_t$ by $\phi\pi_t = \phi\kappa\widehat{y}_t$ and $\mathbb{E}_t\pi_{t+1} = \kappa\mathbb{E}_t\widehat{y}_{t+1}$ (which follows from the Taylor rule and the static Phillips Curve) in the IS equation (18), we get

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\left(\phi\kappa\widehat{y}_t - \kappa \mathbb{E}_t \widehat{y}_{t+1}\right),$$

which can be re-written as

$$\widehat{y}_t\left(1 + \psi_c \frac{1}{\gamma}\phi\kappa\right) = \mathbb{E}_t \widehat{y}_{t+1}\left(\psi_f + \psi_c \frac{1}{\gamma}\kappa\right).$$

Dividing by $\left(1 + \psi_c \frac{1}{\gamma}\phi\kappa\right)$ and plugging in for ψ_f and ψ_c yields

$$\widehat{y}_t = \frac{\bar{m}\delta + \frac{(1-\lambda)\kappa}{\gamma(1-\lambda\chi)}}{1 + \kappa\phi\frac{1}{\gamma}\frac{(1-\lambda)}{1-\lambda\chi}} \mathbb{E}_t \widehat{y}_{t+1}.$$

To obtain determinacy, the term in front of $\mathbb{E}_t \widehat{y}_{t+1}$ has to be smaller than 1. Solving this for ϕ yields

$$\phi > \phi^* = 1 + \frac{\bar{m}\delta - 1}{\frac{\kappa}{\gamma}\frac{1-\lambda}{1-\lambda\chi}}, \tag{32}$$

which is the condition in Proposition 3. This illustrates how bounded rationality raises the likelihood that the Taylor principle ($\phi^* = 1$) is sufficient for determinacy, as the Taylor principle can only hold if $\bar{m}\delta \leq 1$. In the rational model, this boils down to $\delta \leq 1$. However, the Taylor principle can be sufficient under bounded rationality, i.e., $\bar{m} < 1$, even when $\delta > 1$, thus, even when allowing for amplification. Note that we could also express condition (32) as

$$\phi > \phi^* = 1 + \frac{\psi_f - 1}{\frac{\kappa}{\gamma}\psi_c}.$$

Generalizations of Proposition 3. Proposition 3 can easily be extended to allow for Taylor rules of the form

$$\widehat{i}_t = \phi_\pi \pi_t + \phi_y \widehat{y}_t$$

and in which the behavioral agents do not have rational expectations about the real interest rate but rather perceive the real interest rate to be equal to

$$\widehat{r}_t^{BR} \equiv \widehat{i}_t - \bar{m}^r \mathbb{E}_t \pi_{t+1},$$

where $\bar{m}^r$ can be equal to $\bar{m}$ or can potentially differ from it (if it equals 1, we are back to the case in which the behavioral agent is rational with respect to real interest rates).

Combining the static Phillips Curve with the generalized Taylor rule and the behavioral HANK IS equation, it follows that

$$\widehat{y}_t = \frac{\psi_f + \frac{\kappa}{\gamma}\psi_c \bar{m}^r}{1 + \frac{\psi_c}{\gamma}\left(\kappa\phi_\pi + \phi_y\right)} \mathbb{E}_t \widehat{y}_{t+1}. \tag{33}$$

From equation (33), it follows that we need

$$\phi_\pi > \bar{m}^r - \phi_y + \frac{\psi_f - 1}{\psi_c \frac{\kappa}{\gamma}} = \bar{m}^r - \phi_y + \frac{\bar{m}\delta - 1}{\frac{1-\lambda}{1-\chi\lambda}\frac{\kappa}{\gamma}} \tag{34}$$

for the model to feature a determinate, locally unique equilibrium. Condition (34) shows that both, $\bar{m}^r < 1$ and $\phi_y > 0$, weaken the condition in Proposition 3. Put differently, bounded rationality with respect to the real rate or a Taylor rule that responds to changes in output, both relax the condition on ϕ_π to yield determinacy.

A.5 Derivation of Lemma 1

Let us first state a few auxiliary results that will prove helpful later. First, in log-linearized terms, the stochastic discount factor is given by

$$\frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t,t+1}^U = \widehat{c}_t^U - s\bar{m}\mathbb{E}_t\widehat{c}_{t+1}^U - (1-s)\bar{m}\mathbb{E}_t\widehat{c}_{t+1}^H$$

and for i periods ahead:

$$\frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t,t+i}^U = \widehat{c}_t^U - s\bar{m}^i\mathbb{E}_t\widehat{c}_{t+i}^U - (1-s)\bar{m}^i\mathbb{E}_t\widehat{c}_{t+i}^H.$$

Furthermore, we have:

$$\begin{aligned}\frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t+1,t+2}^U &= \mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U - s\widehat{c}_{t+2}^U - (1-s)\widehat{c}_{t+2}^H\right] \\ &= \bar{m}\mathbb{E}_t\widehat{c}_{t+1}^U - s\bar{m}^2\mathbb{E}_t\widehat{c}_{t+2}^U - (1-s)\bar{m}^2\mathbb{E}_t\widehat{c}_{t+2}^H\end{aligned}$$

and the stochastic discount factor has the property

$$\mathbb{E}_t^{BR}\left[\widehat{q}_{t,t+i}^U\right] = \mathbb{E}_t^{BR}\left[\widehat{q}_{t,t+1}^U + \widehat{q}_{t+1,t+2}^U + ... + \widehat{q}_{t+i-1,t+i}^U\right].$$

Using these results, $\mathbb{E}_t^{BR}\left[\widehat{q}_{t,t+i}^U\right]$ can be written as

$$\begin{aligned}\frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t,t+i}^U = \widehat{c}_t^U + (1-s)\bar{m}\mathbb{E}_t\left[\widehat{c}_{t+1}^U - \widehat{c}_{t+1}^H\right] + (1-s)\bar{m}^2\mathbb{E}_t\left[\widehat{c}_{t+2}^U - \widehat{c}_{t+2}^H\right] + ...+ \\ + (1-s)\bar{m}^i\mathbb{E}_t\left[\widehat{c}_{t+i}^U - \widehat{c}_{t+1}^H\right] - \bar{m}^i\mathbb{E}_t\widehat{c}_{t+i}^U,\end{aligned}$$

or put differently

$$\frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t,t+i}^U + \bar{m}^i\mathbb{E}_t\widehat{c}_{t+i}^U = \widehat{c}_t^U + (1-s)\mathbb{E}_t\sum_{k=1}^{i}\bar{m}^k\left(\widehat{c}_{t+k}^U - \widehat{c}_{t+k}^H\right). \tag{35}$$

The (linearized) budget constraint can be written as

$$\mathbb{E}_t^{BR}\sum_{i=0}^{\infty}\beta^i\left(\frac{1}{\gamma}\widehat{q}_{t,t+i}^U + \widehat{c}_{t+i}^U\right) = \mathbb{E}_t^{BR}\sum_{i=0}^{\infty}\beta^i\left(\frac{1}{\gamma}\widehat{q}_{t,t+i}^U + \widehat{y}_{t+i}^U\right)$$

$$\Leftrightarrow \mathbb{E}_t^{BR}\sum_{i=0}^{\infty}\beta^i\left(\frac{1}{\gamma}\widehat{q}_{t,t+i}^U\right) + \mathbb{E}_t\sum_{i=0}^{\infty}(\beta\bar{m})^i\,\widehat{c}_{t+i}^U = \mathbb{E}_t^{BR}\sum_{i=0}^{\infty}\beta^i\left(\frac{1}{\gamma}\widehat{q}_{t,t+i}^U\right) + \mathbb{E}_t\sum_{i=0}^{\infty}(\beta\bar{m})^i\,\widehat{y}_{t+i}^U.$$

Now, focus on the left-hand side and notice that the sum $\mathbb{E}_t\sum_{i=0}^{\infty}(\beta\bar{m})^i\,\widehat{c}_{t+i}^U$ cancels with the $\bar{m}^i\mathbb{E}_t\widehat{c}_{t+i}^U$ terms in equation (35) when summing them up. The left-hand side of the budget constraint can thus be written as

$$\begin{aligned}&\mathbb{E}_t\sum_{i=0}^{\infty}\beta^i\left(\widehat{c}_t^U + (1-s)\sum_{k=1}^{i}\bar{m}^k\left(\widehat{c}_{t+k}^U - \widehat{c}_{t+k}^H\right)\right) \\ &= \frac{1}{1-\beta}\widehat{c}_t^U + (1-s)\mathbb{E}_t\sum_{i=0}^{\infty}\beta^i\sum_{k=1}^{i}\bar{m}^k\left(\widehat{c}_{t+k}^U - \widehat{c}_{t+k}^H\right) \\ &= \frac{1}{1-\beta}\widehat{c}_t^U + \frac{1-s}{1-\beta}\mathbb{E}_t\sum_{i=1}^{\infty}(\beta\bar{m})^i\left(\widehat{c}_{t+i}^U - \widehat{c}_{t+i}^H\right).\end{aligned}$$

Note, from the Euler equation of the unconstrained households, we obtain the real interest rate

$$-\frac{1}{\gamma}\widehat{r}_t = \widehat{c}_t^U - s\mathbb{E}_t^{BR}\widehat{c}_{t+1}^U - (1-s)\mathbb{E}_t^{BR}\widehat{c}_{t+1}^H = \frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t,t+1}^U,$$

and similarly,

$$-\frac{1}{\gamma}\bar{m}^i\mathbb{E}_t\widehat{r}_{t+i} = \frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t+i,t+i+1}^U,$$

where $\widehat{r}_t$ is the (linearized) real interest rate.

Combining these results, we see that

$$\mathbb{E}_t^{BR}\sum_{i=0}^{\infty}\beta^i\frac{1}{\gamma}\widehat{q}_{t,t+i}^U = -\frac{1}{1-\beta}\frac{1}{\gamma}\beta\mathbb{E}_t\sum_{i=0}^{\infty}(\beta\bar{m})^i\,\widehat{r}_{t+i}.$$

Plugging this into the right-hand side of the budget constraint and multiplying both sides by $1-\beta$ yields

$$\widehat{c}_t^U = -\frac{1}{\gamma}\beta\widehat{r}_t + (1-\beta)\widehat{y}_t^U - (1-s)\mathbb{E}_t\sum_{i=1}^{\infty}(\beta\bar{m})^i\left(\widehat{c}_{t+i}^U - \widehat{c}_{t+i}^H\right)$$
$$-\frac{1}{\gamma}\beta\mathbb{E}_t\sum_{i=1}^{\infty}(\beta\bar{m})^i\,\widehat{r}_{t+i} + (1-\beta)\mathbb{E}_t\sum_{i=1}^{\infty}(\beta\bar{m})^i\,\widehat{y}_{t+i}^U,$$

or written recursively

$$\widehat{c}_t^U = -\frac{1}{\gamma}\beta\widehat{r}_t + (1-\beta)\widehat{y}_t^U + \beta\bar{m}s\mathbb{E}_t\widehat{c}_{t+1}^U + \beta\bar{m}(1-s)\mathbb{E}_t\widehat{c}_{t+1}^H.$$

Now, aggregating, i.e., multiplying the expression for $\widehat{c}_t^U$ by $(1-\lambda)$, adding $\lambda\widehat{c}_t^H$ and using $\widehat{c}_t^H = \chi\widehat{y}_t$ as well as $\widehat{y}_t^U = \frac{1-\lambda\chi}{1-\lambda}\widehat{y}_t$, yields the consumption function

$$\widehat{c}_t = [1-\beta(1-\lambda\chi)]\,\widehat{y}_t - \frac{(1-\lambda)\beta}{\gamma}\widehat{r}_t + \beta\bar{m}\delta(1-\lambda\chi)\mathbb{E}_t\widehat{c}_{t+1}, \tag{36}$$

as stated in the main text.

To obtain the share of indirect effects, note that the model does not feature any endogenous state variables and hence, endogenous variables inherit the persistence of the exogenous variables, ρ. Thus, $\mathbb{E}_t\widehat{c}_{t+1} = \rho\widehat{c}_t$. Plugging this into the consumption function (36), we get

$$\widehat{c}_t = \frac{1-\beta(1-\lambda\chi)}{1-\beta\bar{m}\delta\rho(1-\lambda\chi)}\widehat{y}_t - \frac{(1-\lambda)\beta}{\gamma(1-\beta\bar{m}\delta\rho(1-\lambda\chi))}\widehat{r}_t.$$

The term in front of $\widehat{y}_t$ is the share of indirect general equilibrium effects.

Online Appendix: Extensions and Robustness

B Calibrating $\bar{m}$

In most of our analysis, we set the cognitive discounting parameter $\bar{m}$ to 0.85, as in Gabaix (2020). One way at arriving at this value is by matching estimated IS equations. Fuhrer and Rudebusch (2004), for example, estimate an IS equation and find that the coefficient in front of $\mathbb{E}_t \widehat{y}_{t+1}$ (what we call ψ_f) is approximately 0.65, which together with $\delta > 1$, would imply a $\bar{m}$ much lower than 0.85 and especially our determinacy results would be even stronger under such a calibration.

Another way to calibrate $\bar{m}$ (as pointed out in Gabaix (2020)) is to interpret the estimates in Coibion and Gorodnichenko (2015) through the "cognitive-discounting lens". They regress forecast errors on forecast revisions

$$x_{t+h} - F_t x_{t+h} = c + b^{CG} \left(F_t x_{t+h} - F_{t-1} x_{t+h} \right) + u_t,$$

where $F_t x_{t+h}$ denotes the forecast at time t of variable x, h periods ahead. Focusing on inflation, they find that $b^{CG} > 0$ in consensus forecasts, pointing to *underreaction* (similar results are, for example, found in Angeletos et al. (2021) and Adam et al. (2022) for other variables).

In the linearized model, the law of motion of x is $x_{t+1} = \Gamma \left(x_t + \varepsilon_{t+1} \right)$ whereas the behavioral agents perceive it to be $x_{t+1} = \bar{m}\Gamma \left(x_t + \varepsilon_{t+1} \right)$. It follows that $F_t x_{t+h} = \left(\bar{m}\Gamma \right)^h x_t$ and thus, forecast revisions are equal to

$$\begin{aligned} F_t x_{t+h} - F_{t-1} x_{t+h} &= \left(\bar{m}\Gamma \right)^h x_t - \left(\bar{m}\Gamma \right)^{h+1} x_{t-1} \\ &= \left(\bar{m}\Gamma \right)^h \Gamma (1 - \bar{m}) x_{t-1} + \left(\bar{m}\Gamma \right)^h \varepsilon_t. \end{aligned}$$

The forecast error is given by

$$x_{t+h} - F_t x_{t+h} = \Gamma^h (1 - \bar{m}^h) \Gamma x_{t-1} + \Gamma^h (1 - \bar{m}^h) \varepsilon_t + \sum_{j=0}^{h-1} \Gamma^j \varepsilon_{t+h-j},$$

where $\sum_{j=0}^{h-1} \Gamma^j \varepsilon_{t+h-j}$ is the rational expectations forecast error. Gabaix (2020) shows that b^{CG} is bounded below $b^{CG} \geq \frac{1-\bar{m}^h}{\bar{m}^h}$, showing that $\bar{m} < 1$ yields $b^{CG} > 0$, as found empirically. When replacing the weak inequality with an equality, we get

$$\bar{m}^h = \frac{1}{1 + b^{CG}}.$$

Most recently, Angeletos et al. (2021) estimate b^{CG} (focusing on a horizon $h = 3$) to lie between $b^{CG} \in [0.74, 0.81]$ for unemployment forecasts and $b^{CG} \in [0.3, 1.53]$ for inflation, depending on the considered period (see their Table 1). These estimates imply $\bar{m} \in [0.82, 0.83]$ for unemployment and $\bar{m} \in [0.73, 0.92]$ for inflation, and are thus close to our preferred value of 0.85. Note, however, that these estimates pertain to professional forecasters and should therefore be seen as upper bounds on $\bar{m}$. As outlined in Section 2, we estimate these regressions for households to obtain more direct evidence on $\bar{m}$ for households (of different income groups). The following subsection discusses the data, the empirical strategy and the findings we obtain in more detail.

B.1 Estimating $\bar{m}$ for different Household Groups

To test for heterogeneity in the degree of cognitive discounting, we follow Coibion and Gorodnichenko (2015) and regress forecast errors on forecast revisions as follows

$$x_{t+4} - \mathbb{E}_t^{e,BR} x_{t+4} = c^e + b^{e,CG} \left(\mathbb{E}_t^{e,BR} x_{t+4} - \mathbb{E}_{t-1}^{e,BR} x_{t+4} \right) + \epsilon_t^e, \tag{37}$$

to estimate $b^{e,CG}$ for different groups of households, indexed by e. As shown above, $b^{e,CG} > 0$ is consistent with underreaction and the corresponding cognitive discounting parameter is approximately given by (we calibrate the model at quarterly frequency whereas the data is about 1-year-ahead expectations, thus, the adjustment in the exponent)

$$\bar{m}^e = \left(\frac{1}{1 + b^{e,CG}} \right)^{1/4}. \tag{38}$$

Ideally, we would use actual data and expectations data about future marginal utilities of consumption where changes in these variables only being driven by aggregate shocks. However, that data is not available. Instead, we focus on expectations about future unemployment (and inflation) where it seems reasonable to assume that they are only driven by aggregate shocks and that they matter for household's (actual and expected) marginal consumption utility. The Survey of Consumers from the University of Michigan provides 1-year ahead unemployment expectations and we use the unemployment rate from the FRED database as our measure of actual unemployment. We split the households into three groups based on their income. The bottom and top income groups each contain the 25% households with the lowest and highest income, respectively, and the remaining 50% are assigned to the middle income group.

The Michigan Survey asks households whether they expect unemployment to increase, decrease or to remain about the same over the next twelve months. We follow Carlson and Parkin (1975), Mankiw (2000) and Bhandari et al. (2019) to translate these categorical unemployment expectation into numerical expectations.

Focus on group $e \in \{L, M, H\}$ and let $q_t^{e,D}$, $q_t^{e,S}$ and $q_t^{e,U}$ denote the shares within income group e reported at time t that think unemployment will go down, stay roughly the same, or go up over the next year, respectively. We assume that these shares are drawn from a cross-sectional distribution of responses that are normally distributed according to $\mathcal{N}\left(\mu_t^e, (\sigma_t^e)^2\right)$ and a threshold a such that when a household expects unemployment to remain within the range $[-a, a]$ over the next year, she responds that unemployment will remaine "about the same". We thus have

$$q_t^{e,D} = \Phi\left(\frac{-a - \mu_t^e}{\sigma_t^e}\right) \qquad q_t^{e,U} = 1 - \Phi\left(\frac{a - \mu_t^e}{\sigma_t^e}\right),$$

which after some rearranging yields

$$\sigma_t^e = \frac{2a}{\Phi^{-1}\left(1 - q_t^{e,U}\right) - \Phi^{-1}\left(q_t^{e,D}\right)}$$

$$\mu_t^e = a - \sigma_t^e \Phi^{-1}\left(1 - q_t^{e,U}\right).$$

This leaves us with one degree of freedom, namely a. We make two assumptions. First, a is independent of the income group. The second assumption is that we set $a = 0.5$ which means that if a household expects the change in unemployment to be less than half a percentage point (in absolute terms), she reports that she expects unemployment to be about the same as it is at the time of the survey (our results are quite robust with respect to our choice of a).

As the question in the survey is about the expected change in unemployment, we add the actual unemployment rate at the time of the survey to μ_t^e to construct a time-series of unemployment expectations, as in Bhandari et al. (2019). That said, we will also report the case of expected unemployment *changes*.

Given the so-constructed expectations, we can compute forecast revisions as

$$\mu_t^e - \mu_{t-1}^e$$

and four-quarter-ahead forecast errors using the actual unemployment rate u_t obtained from FRED as

$$u_{t+4} - \mu_t^e. \tag{39}$$

For the case of expected unemployment changes, we replace u_{t+4} with $(u_{t+4} - u_t)$ in equation (39).

Following Coibion and Gorodnichenko (2015), we then regress forecast errors on forecast revisions

$$u_{t+4} - \mu_t^e = c^e + b^{e,CG}\left(\mu_t^e - \mu_{t-1}^e\right) + \epsilon_t^e, \tag{40}$$

to estimate $b^{e,CG}$ for each income group e. Note, however, that the expectations in the forecast revisions are about unemployment at different points in time. To account for this, we instrument forecast revisions by the *main business cycle shock* obtained from Angeletos et al. (2020) (Coibion and Gorodnichenko (2015) use a similar IV strategy when considering expectations from the Michigan Survey).

Table 2: Regression Results of Equation (37)

	IV Regression			OLS		
	Bottom 25%	Middle 50%	Top 25%	Bottom 25%	Middle 50%	Top 25%
$\widehat{b}^{e,CG}$	0.85	0.75	0.63	1.22	1.10	0.90
s.e.	(0.471)	(0.453)	(0.401)	(0.264)	(0.282)	(0.247)
F-stat.	24.76	18.74	17.86	-	-	-
N	152	152	152	157	157	157

Note: This table provides the estimated $\widehat{b}^{e,CG}$ from regression (37) for different income groups. The first three columns show the results when the right-hand side in equation (37) is instrumented using the *main business cycle shock* from Angeletos et al. (2020) and the last three columns using OLS. Standard errors are robust with respect to heteroskedasticity and are reported in parentheses. The row "F-stat." reports the first-stage F-statistic for the IV regressions.

Table 2 shows the results. The first three columns report the estimated $b^{e,CG}$ from the IV regressions and the last three columns the same coefficients estimated via OLS. Standard errors are robust with respect to heteroskedasticity and are reported in parentheses. The row "F-stat." reports the first-stage F-statistic for the IV regressions. We see that in all cases $\widehat{b}^{e,CG}$ is positive, suggesting that households of all income groups tend to underreact, consistent with our assumption of $\bar{m} < 1$.

Using equation (38) we obtain $\bar{m}^e$ equal to 0.86, 0.87 and 0.88 for the bottom 25%, the middle 50% and the top 25%, respectively for the estimates from the IV regressions and 0.82, 0.83 and 0.85 for the OLS estimates. When estimating $\bar{m}^e$ using expected unemployment *changes* instead of the level, the estimated $\bar{m}^e$ equal 0.57, 0.59 and 0.64 for the IV regressions and 0.77, 0.80 and 0.86 for the OLS regressions.

There are two main take-aways from this empirical exercise: first, it further confirms that $\bar{m} = 0.85$ is a reasonable (but rather conservative) deviation from rational expectations. Second, the data suggests that there is heterogeneity in the degree of rationality conditional on households income. In particular, households with higher income tend to exhibit higher degrees of rationality.[26]

If we consider inflation expectations instead of unemployment expectations, we obtain estimated cognitive discounting parameters of 0.70, 0.75 and 0.78 for the bottom 25%, the middle 50% and the top 25%, respectively. Thus, somewhat lower than for unemployment and the differences across income groups are larger. In particular, higher-income households tend to be more rational (they discount less) than lower-income households. The differences, however, are overall rather small.

C Figures to Section 3

C.1 Resolving the Catch-22

We graphically illustrate the Catch-22 (Bilbiie (2021)) of the rational model and the resolution of it in the behavioral HANK model in Figure 8. Figure 8 shows the case of *nominal* rate changes. The figure shows on the vertical axis the response of contemporaneous output relative to the initial response in the RANK model with rational expectations for anticipated i.i.d. monetary policy shocks occurring at different times k on the horizontal axis.[27]

The orange-dotted line represents the baseline calibration of the rational HANK model. We see that this model is able to generate contemporaneous amplification of monetary policy shocks, that is, an output response that is relatively stronger than in RANK. Put differently the GE effects

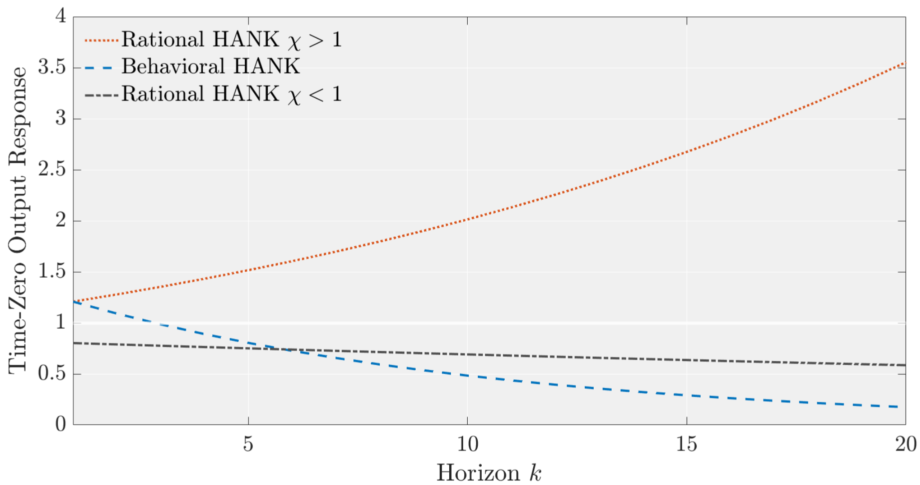
[26] This is consistent with other empirical findings on heterogeneous deviations from FIRE. Broer et al. (2022), for example, document that wealthier households tend to have more accurate beliefs, as measured by forecast errors.

[27] Under fully-rigid prices (i.e., $\kappa = 0$) the RANK model would deliver a constant response for all k. The same is true for two-agent NK models (TANK), i.e., tractable HANK models without type switching. Whether the constant response would lie above or below its RANK counterpart depends on $\chi \lesseqgtr 1$ in the same way the initial response depends on $\chi \lesseqgtr 1$.

amplify the effects of monetary policy shocks. Yet, at the same time, it exacerbates the forward guidance puzzle as shocks occurring in the future have even stronger effects on today's output than contemporaneous shocks.

The black-dashed-dotted line shows how the forward guidance puzzle can be resolved by allowing for $\chi < 1$. Yet, this comes at the cost that the model is unable to generate amplification of contemporaneous monetary policy shocks. Recent empirical findings, however, document that GE effects indeed amplify monetary policy changes (Patterson (2023), Auclert (2019)).

Figure 8: Resolving the Catch-22



Note: This figure shows the response of total output in period 0 to anticipated i.i.d. monetary policy shocks occurring at different horizons k (horizontal axis), relative to the initial response in the RANK model under rational expectations (equal to 1).

The blue-dashed line shows that the behavioral HANK model, on the other hand, generates both: amplification of contemporaneous monetary policy and a resolution of the forward guidance puzzle, both consistent with the empirical facts.

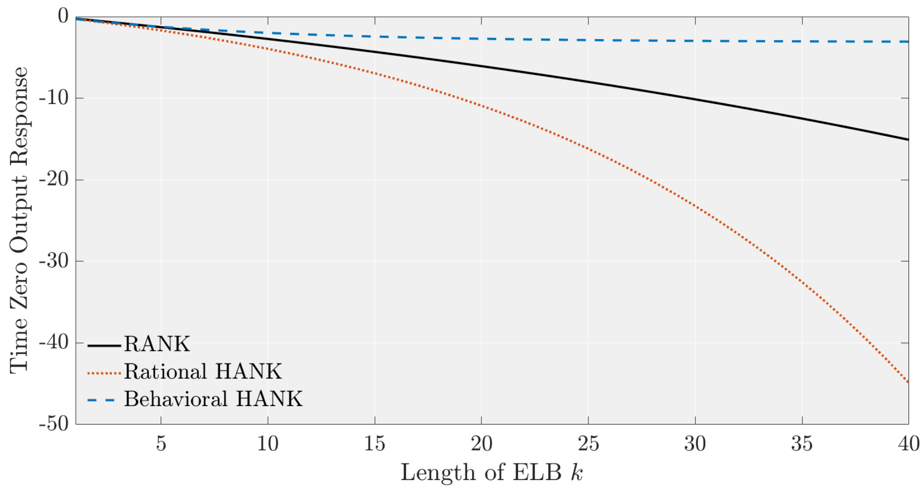
C.2 Stability at the Effective Lower Bound

We illustrate the stability of the behavioral HANK model at the lower bound graphically in Figure 9. Recall from Section 2, the forward-iterated IS equation with a natural rate shock:

$$\widehat{y}_t = -\frac{1}{\gamma}\psi_c \underbrace{\left(\widehat{i}_{ELB} - \widetilde{r}^n\right)}_{>0} \sum_{j=0}^{k} \left(\psi_f + \frac{\kappa}{\gamma}\psi_c\right)^j .$$

Figure 9 shows the output response in RANK, the rational HANK and the behavioral HANK to different lengths of a binding ELB (depicted on the horizontal axis). The shortcoming of monetary policy due to the ELB, i.e., the gap $\left(\widehat{i}_{ELB} - \widetilde{r}^n\right) > 0$, is set to a relatively small value of 0.25% (1% annually), and we set $\bar{m} = 0.85$. Figure 9 shows the implosion of output in the rational RANK (back-solid line) and even more so in the rational HANK model (orange-dotted line): an ELB that is expected to bind for 40 quarters would decrease today's output in the rational RANK

Figure 9: The Effective Lower Bound Problem



Note: This figure shows the contemporaneous output response for different lengths of a binding ELB k (horizontal axis) and compares the responses across different models.

by 15% and in the rational HANK model by 45%. On the other hand—and consistent with recent experiences in advanced economies—output in the behavioral HANK model remains quite stable and drops by a mere 3%, as illustrated by the blue-dashed line.

D Extensions and Robustness of the Analytical Model

D.1 Robustness of Calibration

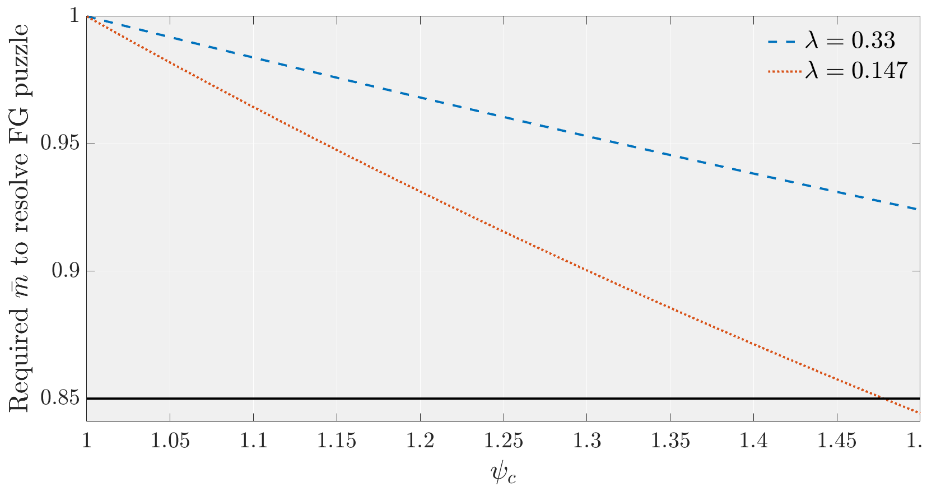
In our baseline calibration, we obtain an amplification of conventional monetary policy shocks of 20% compared to the case in which all households are equally exposed to monetary policy (i.e., $\psi_c = 1.2$) for a given share of hand-to-mouth, λ. In particular, we set $\lambda = 0.33$. This results in $\chi = 1.35$. In the quantitative model, we also obtain an amplification of about 20% but this is implied by targeting the micro evidence from Patterson (2023).

To show the robustness of our results, we show in Figure 10 for different ψ_c (on the horizontal axis) the highest $\bar{m}$ (on the vertical axis) that still resolves the forward-guidance puzzle. The blue-dashed line shows this for $\lambda = 0.33$ and the orange-dotted line for $\lambda = 0.147$ which is the share of borrowing-constrained households in Farhi and Werning (2019).

We see that a $\bar{m}$ of 0.85 (as indicated by the black-solid line) rules out the forward-guidance puzzle in almost all cases. Only at the relatively low λ of 0.147 and a high $\psi_c > 1.48$, we would require a $\bar{m}$ of about 0.84 instead of 0.85 to rule out the forward-guidance puzzle (note, that a ψ_c of 1.47 at $\lambda = 0.147$ implies $\chi = 2.86$). Given that the empirical estimates point towards values of $\bar{m} \in [0.6, 0.85]$, we conclude the resolution of the forward guidance puzzle in the behavioral HANK model with countercyclical income risk is quite robust.

Also note that the values for γ and κ that we use are directly taken from Bilbiie (2021, 2020)

Figure 10: Robustness of Forward Guidance Puzzle Solution



Note: This figures show for different ψ_c (horizontal axis) the required $\bar{m}$ to resolve the forward-guidance puzzle on the vertical axis. The blue-dashed line shows this for our benchmark calibration of $\lambda = 0.33$ and the orange-dotted line for $\lambda = 0.147$.

and are quite standard in the literature. Gabaix (2020), however, sets $\kappa = 0.11$ and $\gamma = 5$. Even though these coefficients differ quite substantially from our baseline calibration, note that our results would barely be affected by this. To see this, note that *amplification* is only determined by λ and χ, both independent of κ and γ. The determinacy condition on the other hand depends on both, κ and γ, but what ultimately matters is the fraction $\frac{\kappa}{\gamma}$ (see Proposition 3). As κ and γ are both approximately five times larger in Gabaix (2020) compared to Bilbiie (2021) and our baseline calibration, the fraction is approximately the same and thus, the determinacy region under an interest-rate peg remains unchanged.

D.2 Nominal Interest Rate Changes

In Section 3, we focused on the case where monetary policy directly controls the *real* rather than the nominal interest rate. We now show that our results are unchanged when instead focusing on nominal rate changes. As in the main text, we consider two different monetary policy experiments: (i) a contemporaneous monetary policy shock, i.e., a surprise decrease in the nominal interest rate today, and (ii) a forward guidance shock, i.e., a news shock today about a decrease in the nominal interest rate k periods in the future. In both cases, we focus on *i.i.d.* shocks and the Taylor response coefficient is zero, $\phi = 0$.[28]

Proposition 4. *In the behavioral HANK model, there is amplification of contemporaneous monetary policy relative to RANK if and only if*

$$\psi_c > 1 \Leftrightarrow \chi > 1, \tag{41}$$

[28]If we instead impose $\phi > 0$, contemporaneous amplification in the following proposition is not affected but the condition to rule out the forward guidance puzzle is further relaxed.

and the forward guidance puzzle is ruled out if

$$\psi_f + \frac{\kappa}{\gamma}\psi_c < 1. \tag{42}$$

We thus see, that the amplification result is unchanged (see Proposition 2) whereas the condition to rule out the forward-guidance puzzle is somewhat stricter as $\frac{\kappa}{\gamma}\psi_c > 0$. This is the case because there is now an inflation feedback effect. An expected decrease in the nominal interest rate in the future increases inflation expectations and thus, lowers the real rate further. Thus, the effects on today's output become stronger.

However, again a relatively small underreaction of the behavioral households is enough to resolve the forward guidance puzzle. Given our calibration there is no forward guidance puzzle in the behavioral HANK model as long as $\bar{m} < 0.94$ which is above the upper bounds for empirical estimates (see Section 2).

D.3 Allowing for Steady State Inequality

In the tractable model, we have assumed that there is no steady state inequality, i.e., $C^H = C^U$. In the following, we relax this assumption and denote steady state inequality by $\Omega \equiv \frac{C^U}{C^H}$. Recall the Euler equation of unconstrained households

$$\left(C_t^U\right)^{-\gamma} = \beta R_t \mathbb{E}_t^{BR}\left[s\left(C_t^U\right)^{-\gamma} + (1-s)\left(C_t^H\right)^{-\gamma}\right],$$

from which we can derive the steady state real rate

$$R = \frac{1}{\beta(s + (1-s)\Omega^\gamma)}.$$

Log-linearizing the Euler equation yields

$$\widehat{c}_t^U = \beta R \bar{m}\left[s\mathbb{E}_t\widehat{c}_{t+1}^U + (1-s)\Omega^\gamma \mathbb{E}_t \widehat{c}_{t+1}^H\right] - \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t\pi_{t+1}\right).$$

Combining this with the consumption functions and the steady state real rate yields the IS equation

$$\widehat{y}_t = \bar{m}\tilde{\delta}\mathbb{E}_t\widehat{y}_{t+1} - \frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\left(\widehat{i}_t - \mathbb{E}_t\pi_{t+1}\right), \tag{43}$$

with

$$\tilde{\delta} \equiv 1 + (\chi - 1)\frac{(1-s)\Omega^\gamma}{s + (1-s)\Omega^\gamma}\frac{1}{1-\lambda\chi}.$$

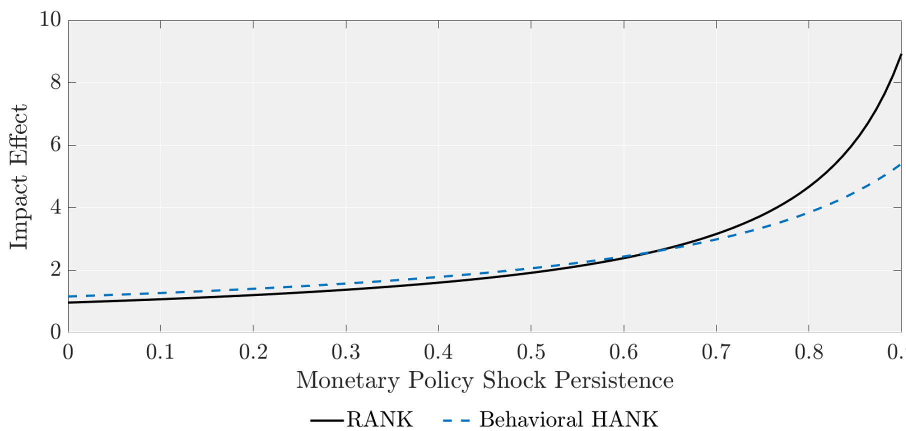
From a qualitative perspective, the whole analysis in Section 3 could be carried out with $\tilde{\delta}$ instead of δ. Quantitatively the differences are small as well. For example, if we set $\Omega = 1.5$, we get $\tilde{\delta} = 1.05$ instead of $\delta = 1.034$. Thus, we need $\bar{m} < 0.93$ instead of $\bar{m} < 0.94$ for determinacy under a peg.

D.4 Persistent Monetary Policy Shocks

In the main text in Section 3, we illustrated the resolution of the Catch-22 by considering i.i.d. monetary policy shocks (following Bilbiie (2021)). The behavioral HANK model delivers initial amplification of these monetary shocks but the effects decrease with the horizon of the shock, i.e., the behavioral HANK model resolves the forward guidance puzzle. Another way to see this is by considering persistent shocks.

Figure 11 illustrates this. The figure shows the response of output in period t to a shock in period t for different degrees of persistence (x-axis). The black-solid line shows the output response in RANK and the blue-dashed line in the behavioral HANK. The forward guidance puzzle in RANK manifests itself in the sense that highly persistent shocks have stronger effects in RANK than in the behavioral HANK. Persistent shocks are basically a form of forward guidance and thus, with high enough persistence in the shocks, the RANK model predicts stronger effects than the behavioral HANK model.

Figure 11: Initial Output Response for Varying Degrees of the Persistence



Note: This figure shows the initial output response to monetary policy shocks with different degrees of persistence.

As the persistence of the monetary policy shock approaches unity, the rational model leads to the paradoxical finding that an exogenous increase in the nominal interest rate leads to an expansion in output. To see this, note that we can write output as

$$\widehat{y}_t = -\frac{\frac{\psi_c}{\gamma}}{1 + \frac{\psi_c}{\gamma}\phi\kappa - \left(\psi_f + \psi_c \frac{\kappa}{\gamma}\right)\rho}\varepsilon_t^{MP}. \tag{44}$$

Given our baseline calibration and a Taylor coefficient of $\phi = 1$, the rational model would produce these paradoxical findings for $\rho > 0.967$. The behavioral HANK model, on the other hand, does not suffer from this as the denominator is always positive, even when $\phi = 0$ and $\rho = 1$.

D.5 Forward-Looking NKPC and Real Interest Rates

In the tractable model, we made the assumption that agents are rational with respect to real interest rates (as in Gabaix (2020)) and assumed a static Phillips Curve for simplicity. We now show that the results are barely affected when considering a forward-looking New Keynesian Phillips Curve (NKPC) and that agents are also boundedly rational with respect to real rates.

Gabaix (2020) derives the NKPC under bounded rationality and shows that it takes the form:

$$\pi_t = \beta M^f \mathbb{E}_t \pi_{t+1} + \kappa \widehat{y}_t,$$

with

$$M^f \equiv \bar{m}\left(\theta + \frac{1-\beta\theta}{1-\beta\theta\bar{m}}(1-\theta)\right),$$

where $1-\theta$ captures the Calvo probability of price adjustment.

Taking everything together (including the bounded rationality with respect to real interest rates), the model can be summarized by the following three equations:

$$\begin{aligned}
\widehat{y}_t &= \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\left(\widehat{i}_t - \bar{m}\mathbb{E}_t \pi_{t+1}\right) \\
\pi_t &= \beta M^f \mathbb{E}_t \pi_{t+1} + \kappa \widehat{y}_t \\
\widehat{i}_t &= \phi \pi_t.
\end{aligned}$$

Plugging the Taylor rule into the IS equation, we can write everything in matrix form:

$$\begin{pmatrix} \mathbb{E}_t \pi_{t+1} \\ \mathbb{E}_t \widehat{y}_{t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{\beta M^f} & -\frac{\kappa}{\beta M^f} \\ \frac{\psi_c}{\gamma \psi_f}\left(\phi - \frac{\bar{m}}{\beta M^f}\right) & \frac{1}{\psi_f}\left(1 + \frac{\psi_c \bar{m} \kappa}{\gamma \beta M^f}\right) \end{pmatrix}}_{\equiv A} \begin{pmatrix} \pi_t \\ \widehat{y}_t \end{pmatrix}. \tag{45}$$

For determinacy, we need

$$det(A) > 1; \quad det(A) - tr(A) > -1; \quad det(A) + tr(A) > -1.$$

The last condition is always satisfied. The first two conditions are satisfied if and only if

$$\phi > \max\left\{\frac{\beta\delta M^f \bar{m} - 1}{\frac{\kappa}{\gamma}\frac{1-\lambda}{1-\lambda\chi}}, \bar{m} + \frac{(\delta\bar{m}-1)(1-\beta M^f)}{\frac{\kappa}{\gamma}\frac{1-\lambda}{1-\lambda\chi}}\right\}.$$

In the case of a static Phillips curve but bounded rationality with respect to the real rate, the second condition is the crucial one. To capture the static Phillips curve, we can simply set $M^f = 0$. We can see that bounded rationality with respect to the real rate relaxes the determinacy condition whereas a forward-looking NKPC tightens it. But even in the case of a forward-looking NKPC (rational or behavioral), cognitive discounting relaxes the determinacy condition and thus, all our results from the static Phillips curve are qualitatively unchanged. Under our baseline calibration and $\theta = 0.875$ and $\beta = 0.99$ as in Gabaix (2020), the model still features determinacy under a peg, even when real interest rate expectations are rational (and therefore, also when they are behavioral).

D.6 Cognitive Discounting of the State Vector

In Section 2, we assume that cognitive discounting applies to all variables, which differs slightly from the assumption in Gabaix (2020) who assumes that cognitive discounting applies to the *state* of the economy (exogenous shocks as well as announced monetary and fiscal policies). He then proves (Lemma 1 in Gabaix (2020)) how cognitive discounting applies as a result (instead of as an assumption) to all future variables, including future consumption choices. For completeness, we show in this section how our results are unaffected when following the approach in Gabaix (2020).

Let X_t denote the (de-meaned) state vector which evolves as

$$X_{t+1} = G^X\left(X_t, \varepsilon_{t+1}\right), \tag{46}$$

where G^X denotes the transition function of X in equilibrium and ε are zero-mean innovations. Linearizing equation (46) yields

$$X_{t+1} = \Gamma X_t + \varepsilon_{t+1}, \tag{47}$$

where ε_{t+1} might have been renormalized. The assumption in Gabaix (2020) is that the behavioral agent perceives the state vector to follow

$$X_{t+1} = \bar{m} G^X(X_t, \varepsilon_{t+1}), \tag{48}$$

or in linearized terms

$$X_{t+1} = \bar{m}\left(\Gamma X_t + \varepsilon_{t+1}\right). \tag{49}$$

The expectation of the boundedly-rational agent of X_{t+1} is thus $\mathbb{E}_t^{BR}\left[X_{t+1}\right] = \bar{m}\mathbb{E}_t\left[X_{t+1}\right] = \bar{m}\Gamma X_t$. Iterating forward, it follows that $\mathbb{E}_t^{BR}\left[X_{t+k}\right] = \bar{m}^k \mathbb{E}_t\left[X_{t+k}\right] = \bar{m}^k \Gamma^k X_t$.

Now, consider any variable $z(X_t)$ with $z(0) = 0$ (e.g., demeaned consumption of unconstrained households $C^U(X_t)$). Linearizing $z(X)$, we obtain $z(X) = b_X^z X$ for some b_X^z and thus

$$\begin{aligned}
\mathbb{E}_t^{BR}\left[z(X_{t+k})\right] &= \mathbb{E}_t^{BR}\left[b_X^z X_{t+k}\right] = b_X^z \mathbb{E}_t^{BR}\left[X_{t+k}\right] \\
&= b_X^z \bar{m}^k \mathbb{E}_t\left[X_{t+k}\right] = \bar{m}^k \mathbb{E}_t\left[b_X^z X_{t+k}\right] \\
&= \bar{m}^k \mathbb{E}_t\left[z(X_{t+k})\right].
\end{aligned}$$

For example, expected consumption of unconstrained households tomorrow (in linearized terms) is given by

$$\mathbb{E}_t^{BR}\left[\widehat{c}^U(X_{t+1})\right] = \bar{m}\mathbb{E}_t\left[\widehat{c}^U(X_{t+1})\right], \tag{50}$$

which we denote in the main text as

$$\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] = \bar{m}\mathbb{E}_t\left[\widehat{c}_{t+1}^U\right]. \tag{51}$$

Now, take the linearized Euler equation (16) of unconstrained households:

$$\widehat{c}_t^U = s\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] + (1-s)\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^H\right] - \frac{1}{\gamma}\widehat{r}_t, \tag{52}$$

where $\widehat{r}_t \equiv \widehat{i}_t - \mathbb{E}_t \pi_{t+1}$.

Using the notation in Gabaix (2020), we can write the Euler equation as

$$\widehat{c}^U(X_t) = s\mathbb{E}_t^{BR}\left[\widehat{c}^U(X_{t+1})\right] + (1-s)\mathbb{E}_t^{BR}\left[\widehat{c}^H(X_{t+1})\right] - \frac{1}{\gamma}\widehat{r}(X_t). \tag{53}$$

Now, applying the results above, we obtain

$$\widehat{c}^U(X_t) = s\bar{m}\mathbb{E}_t\left[\widehat{c}^U(X_{t+1})\right] + (1-s)\bar{m}\mathbb{E}_t\left[\widehat{c}^H(X_{t+1})\right] - \frac{1}{\gamma}\widehat{r}(X_t), \tag{54}$$

which after writing $\widehat{c}^U(X_t)$, $\widehat{c}^U(X_{t+1})$ and $\widehat{c}^H(X_{t+1})$ in terms of total output yields exactly the IS equation in Proposition 1.

D.7 Microfounding $\bar{m}$

Gabaix (2020) shows how to microfound $\bar{m}$ from a noisy signal extraction problem in the case of a representative agent. Following these lines, we show how such a signal-extraction problem offers a potential microfoundation in the heterogeneous agent case, too.

The (linearized) law of motion of the state variable, X_t, is given by $X_{t+1} = \Gamma X_t + \varepsilon_{t+1}$ (a similar reasoning extends to the non-linearized case), where X has been demeaned. Now assume that each households j performs a mental simulation of the future, but receives only noisy signals about that simulation, i.e., the household receives signals S^j_{t+1} of X_{t+1}, and these signals are given by

$$S^j_{t+1} = \begin{cases} X_{t+1} & \text{with probability } p \\ X'_{t+1} & \text{with probability } 1-p \end{cases}$$

where X'_{t+1} is an i.i.d. draw from the unconditional distribution of X_{t+1}, which has an unconditional mean of zero. In words, with probability p the agent j receives perfectly precise information in one particular mental simulation of the future, and with probability $1-p$ agent j receives a signal realization that is completely uninformative. A fully-informed rational agent would have $p = 1$.

The household runs a continuum of these simulations in his head. The conditional mean of X_{t+1}, given the signal S^j_{t+1}, is given by

$$X^e_{t+1} \equiv \mathbb{E}\left[X_{t+1}|S_{t+1} = s^j_{t+1}\right] = p \cdot s^j_{t+1}.$$

To see this, note that the joint distribution of (X_{t+1}, S^j_{t+1}) is

$$f(x_{t+1}, s^j_{t+1}) = pg(s^j_{t+1})\delta_{s^j_{t+1}}(x_{t+1}) + (1-p)g(s^j_{t+1})g(x_{t+1}),$$

where $g(X_{t+1})$ denotes the distribution of X_{t+1} and δ is the Dirac function. Given that the unconditional mean of X_{t+1} is 0, i.e., $\int x_{t+1}g(x_{t+1})dx_{t+1} = 0$, it follows that

$$\begin{aligned} \mathbb{E}_t\left[X_{t+1}|S^j_{t+1} = s^j_{t+1}\right] &= \frac{\int x_{t+1} f(x_{t+1}, s^j_{t+1})dx_{t+1}}{\int f(x_{t+1}, s^j_{t+1})dx_{t+1}} \\ &= \frac{pg(s^j_{t+1})s^j_{t+1} + (1-p)g(s^j_{t+1})\int x_{t+1}g(x_{t+1})dx_{t+1}}{g(s^j_{t+1})} \\ &= ps^j_{t+1}. \end{aligned}$$

Furthermore, we have

$$\mathbb{E}\left[S_{t+1}|X_{t+1}\right] = pX_{t+1} + (1-p)\mathbb{E}\left[X'_{t+1}\right] = pX_{t+1}.$$

So, it follows that the *average* expectation of X_{t+1} over all these simulations is given by

$$\mathbb{E}\left[X^e_{t+1}(S_{t+1})|X_{t+1}\right] = \mathbb{E}\left[p \cdot S_{t+1}|X_{t+1}\right]$$

$$= p \cdot \mathbb{E}\left[S_{t+1}|X_{t+1}\right]$$
$$= p^2 X_{t+1}.$$

Defining $\bar{m} \equiv p^2$ and since $X_{t+1} = \Gamma X_t + \varepsilon_{t+1}$, we have that the agent perceives the law of motion of X to equal

$$X_{t+1} = \bar{m}\left(\Gamma X_t + \varepsilon_{t+1}\right), \tag{55}$$

as imposed in equation (49). The boundedly-rational expectation of X_{t+1} is then given by

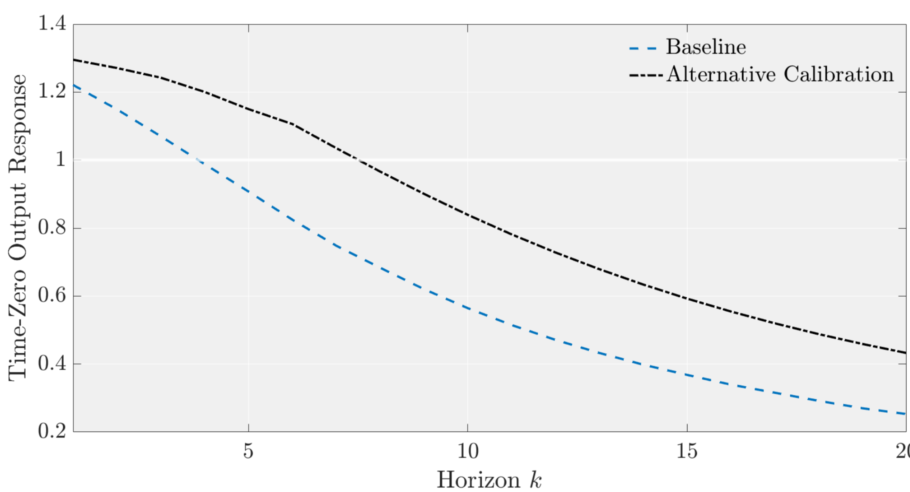
$$\mathbb{E}_t^{BR}\left[X_{t+1}\right] = \bar{m}\mathbb{E}_t\left[X_{t+1}\right].$$

E Details and Extensions to Section 4

E.1 Robustness of Calibration

Following Patterson (2023), we calibrate the unequal income exposure of households such that a linear regression of the income elasticity w.r.t. GDP on MPCs yields a coefficient of 1.33. To show that our results are robust to more extreme calibrations, Figure 12 shows the case where we target a coefficient of 2.0. As one would expect, contemporaneous monetary policy shocks are further amplified and due to the induced countercyclical income risk forward guidance becomes somewhat more effective. Overall, however, we conclude that the forward-guidance puzzle is still clearly ruled out and our results thus robust.

Figure 12: Robustness of Calibration

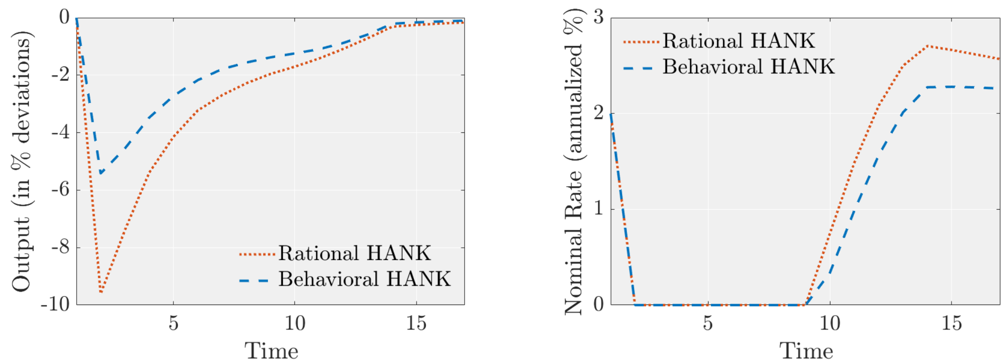


Note: This figure shows the response of total output in period 0 to anticipated i.i.d. monetary policy shocks occurring at different horizons k for a more unequal income exposure of households.

E.2 Stability at the ELB and Fiscal Multipliers

Figure 13 shows the output and nominal interest rate response after a shock to the discount factor in the quantitative behavioral HANK model and in its rational counterpart. In particular, the discount factor jumps on impact by 0.65% for 12 quarters before it returns to steady state.

Figure 13: ELB recession in the quantitative behavioral HANK model



Note: This figure shows the impulse responses of total output and of the nominal interest rate after a discount factor shock that brings the economy to the ELB for 8 quarters.

We see that while the interest-rate path is quite similar across the two models, the output drop in the rational model is about twice as deep as in the behavioral HANK model. The intuition is as in the tractable model (Section 3). The binding ELB acts like a contractionary monetary policy shock because the nominal interest rate cannot keep up with the drop in the natural rate due to the ELB. Under rational expectations, households fully account for this and thus, cut back their consumption quite strongly on impact. Thus, the ELB leads to a large recession. Under cognitive discounting, on the other hand, households discount these future shocks and hence, decrease their consumption by less, leading to a milder recession.

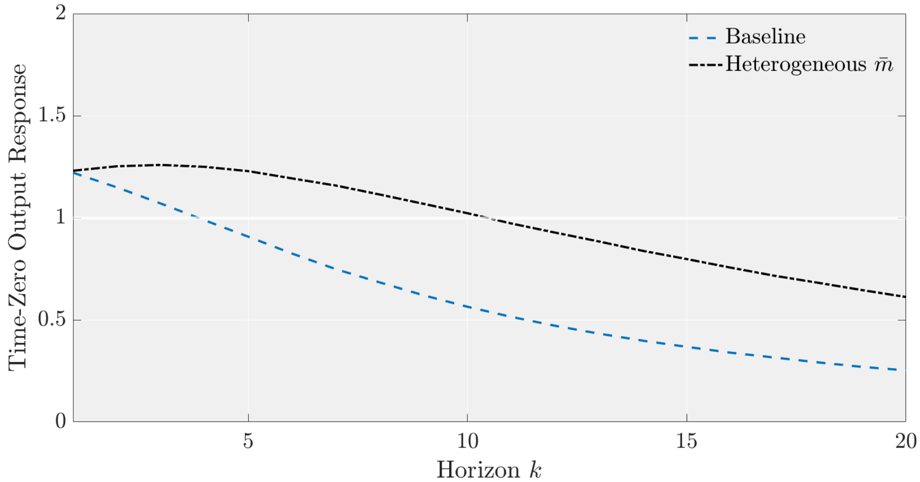
E.3 Heterogeneous $\bar{m}$: Alternative Calibration

The estimated differences in households' underreaction across different income groups are rather small. Nevertheless, one might argue that some agents (financial markets, for example) closely track what the Fed is doing and that they are usually well informed about its actions. To mirror this, we assume that the highest-productivity households are fully rational, i.e., their $\bar{m}$ is equal to 1. To keep the average $\bar{m}$ at 0.85, we then assume that the lowest-productivity households have a $\bar{m}$ of 0.7 and the middle-productivity households of 0.85.

The black-dashed-dotted line in Figure 14 shows the time zero output response (vertical axis) to an announced monetary policy shock taking place at different horizons (horizontal axis).

We see that forward guidance is more powerful than in the baseline calibration as the agents that tend to be more forward looking because they are not at their borrowing constraint are also more rational. Overall, however, our results remain robust. Thus, even when a subpopulation of all households is fully rational, the behavioral HANK model can simultaneously generate amplification of conventional monetary policy through indirect effects and rule out the forward guidance puzzle.

Figure 14: Heterogeneous $\bar{m}$ and Monetary Policy

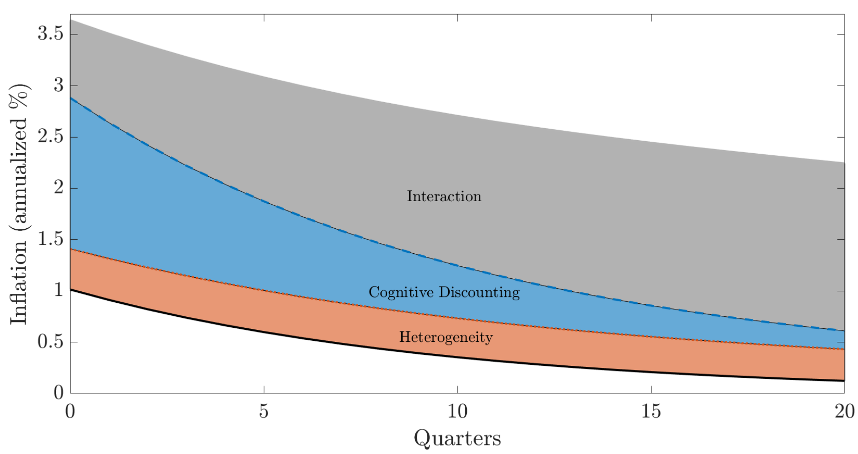


Note: This figure shows the response of total output in period 0 to anticipated i.i.d. monetary policy shocks occurring at different horizons k for the baseline calibration with $\bar{m} = 0.85$ for all households (blue-dashed line), and for the model in which high productivity households have $\bar{m} = 1$, medium-level productivity households have $\bar{m} = 0.85$ and low-productivity households have $\bar{m} = 0.7$ (black-dashed-dotted line).

F Additional Results and Figures to Section 5

F.1 Decomposition of Amplification Channel: More Cognitive Discounting

Figure 15: Decomposition of the Additional Inflation Increase: Lower $\bar{m}$



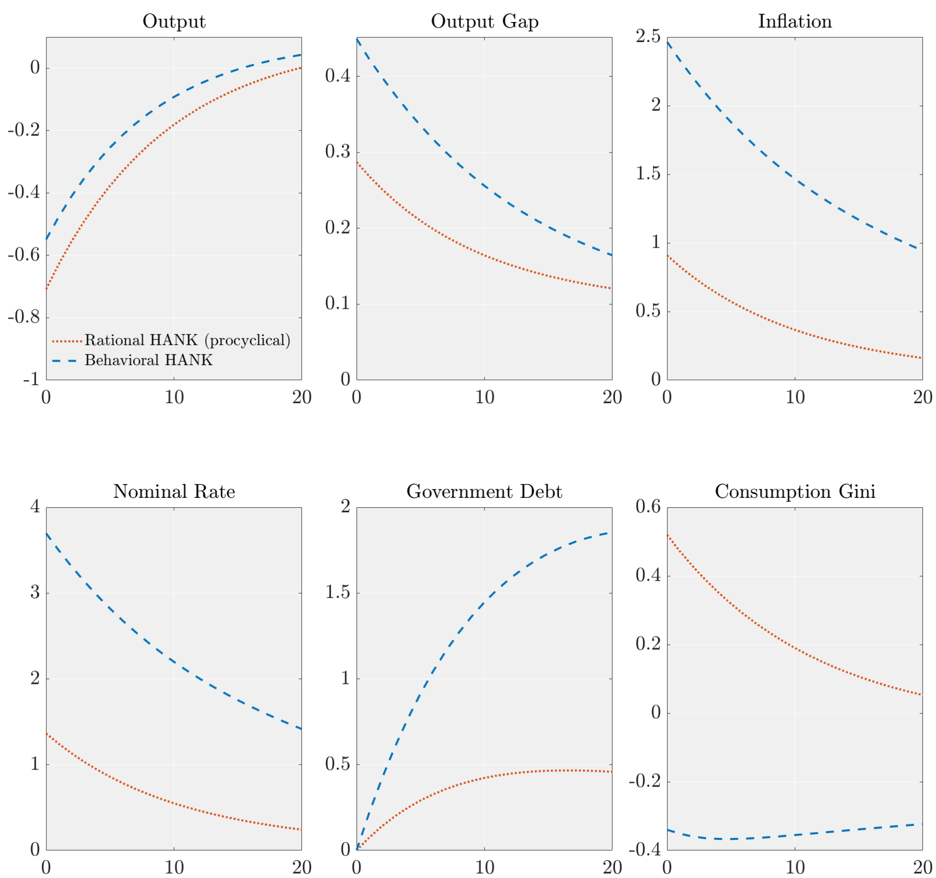
Note: This figure shows the decomposition of the additional inflation increase in the behavioral HANK model for $\bar{m} = 0.6$ compared to the rational RANK model. The orange-shaded area represents the additional increase that is solely due to the heterogeneous exposure of households, the blue area the increase due to cognitive discounting and the gray area the additional increase that is due to the interaction of heterogeneity and cognitive discounting.

As shown in Section 5, household heterogeneity and cognitive discounting interact in such a way that productivity shocks get amplified through both of these ingredients as well as their interaction. Given our baseline calibration, the interaction accounts for about 19% of the additional increase compared to RANK. We now consider an alternative calibration where we set the cognitive discounting parameter $\bar{m}$ to 0.6 instead of 0.85. Thus, somewhat closer to the lower bound of empirical estimates (see Section 2). Figure 15 shows the decomposition of the additional amplification of negative productivity shocks under this alternative calibration.

Two things stand out. First, the overall inflation increase is more than twice as large compared to RANK. Given our discussion in Section 5, this is no surprise. The stronger cognitive discounting induces a larger increase in inflation after the negative productivity shock. Second, the interaction becomes even more important. In fact, the interaction alone accounts for more than the underlying heterogeneity itself. It amounts to more than 75% of the impact inflation response in RANK (1 percentage point) or about 29% of the *additional increase*.

F.2 Procyclical HANK

Figure 16: Procyclical inequality

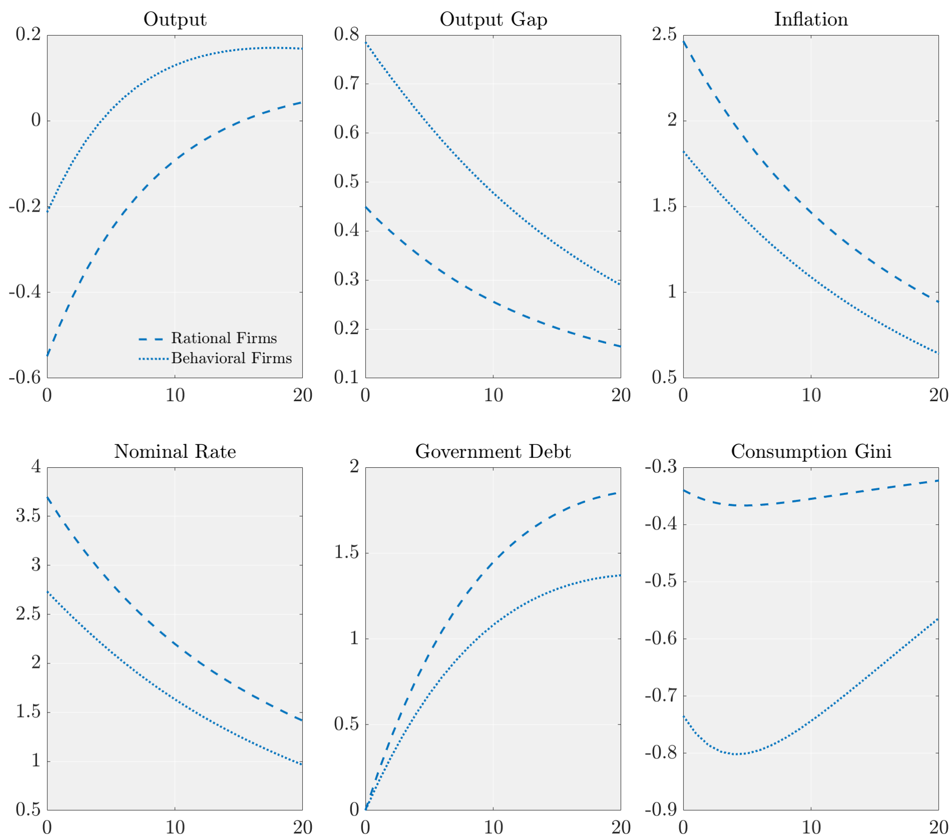


Note: This figure shows the impulse responses after a TFP shock that decreases potential output by 1% when monetary policy follows a Taylor rule for the behavioral HANK model (blue-dashed lines) and for the rational HANK model with procyclical inequality (orange-dotted lines). Output and the output gap are shown as percentage deviations from steady state output, the nominal interest rate and inflation as annualized percentage points and the government debt level as percentage point deviations of the debt-per-annual GDP level. The lower-right figure shows the change in the consumption Gini index as a percentage deviation from the stationary equilibrium.

F.3 Behavioral Firms

Figure 17 shows the impulse-response functions after a negative productivity shock when monetary policy follows a Taylor rule and in which firms are behavioral (with a cognitive discounting factor of 0.85). We see that the increase in inflation when monetary policy follows a Taylor rule is somewhat muted whereas the increase in the output gap is strongly amplified compared to the case in which firms are rational. The reason is that firms discount the increase in their future marginal costs and thus increase their prices not as strongly. According to the Taylor rule this then leads to a smaller increase in the nominal interest rate (both channels inducing a lower real rate) so that households consume more, leading to an increase in demand and thus, the output gap.

Figure 17: Inflationary supply shock: behavioral firms



Note: This figure shows the impulse responses after a productivity shock for the case that monetary policy follows a Taylor rule and firms cognitively discount the future with a cognitive discounting parameter of 0.85. Output and the output gap are shown as percentage deviations from steady state output, the nominal interest rate and inflation as annualized percentage points and the government debt level as percentage point deviations of the debt-per annual-GDP level. The lower-right figure shows the change in the consumption Gini index as a percentage deviation from the stationary equilibrium.

F.4 Cost-Push Shocks

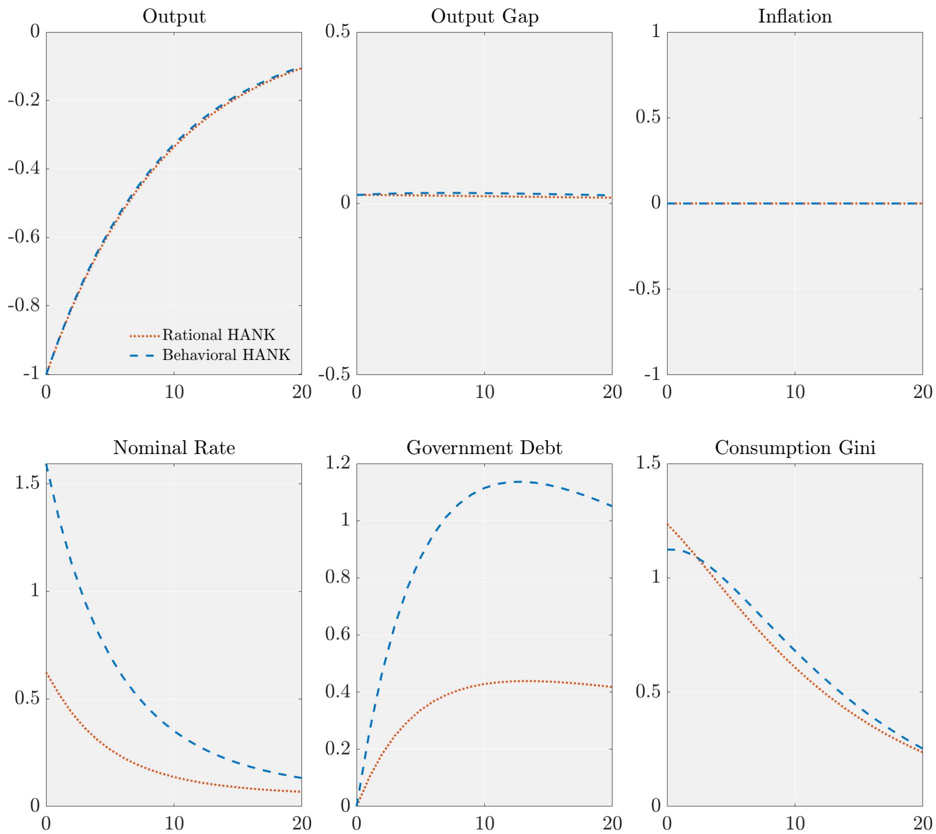
We now show that the fiscal and monetary implications are very similar for an inflationary cost-push shock. To introduce cost-push shocks, we assume that the desired mark-up of firms, μ_t follows an AR(1)-process, $\mu_t = (1 - \rho_\mu)\bar{\mu} + \rho_\mu \mu_{t-1} + \varepsilon_t^\mu$, where ε_t^μ is an i.i.d. shock, $\bar{\mu}$ the steady-state level of the desired markup and ρ_μ the persistence of the shock process which we set to $\rho_\mu = 0.9$. The rest of the model is as in Section 5. Note, that we model the shock such that it also applies to the model under flexible prices, thus moves potential output as well.

Figure 18 shows the impulse-response functions of output, the output gap, inflation, nominal interest rates, government debt and the consumption Gini index as a measure of consumption inequality following an inflationary cost-push shock. The blue-dashed lines show the responses in the behavioral HANK model with homogeneous and heterogeneous $\bar{m}$, respectively, and the orange-dotted lines in the rational HANK model. In both cases, monetary policy fully stabilizes inflation by assumption. Output drops, with the responses being practically identical across the two models. Again, the output gap is practically closed in both models. The required response of the nominal interest rate, however, differs substantially across the behavioral and the rational model, as was the case after a negative productivity shock, discussed in Section 5. In the behavioral HANK model the monetary authority increases the nominal rate much more strongly and more persistently. The reason for this strong response is that households cognitively discount future (expected) interest rate hikes making them less effective for stabilizing inflation today. Thus, in order to achieve the same stabilization outcome in every period, the interest rate needs to increase by more.

Increasing the interest rate more strongly increases the cost of debt for the government which it finances in the short run by issuing more debt. The middle panel on the bottom line in Figure 18 shows that government debt in the behavioral model increases more than three times as much as in the rational model. Furthermore, consumption inequality increases in both models. There are two channels: first and most important, the cost-push shock increases dividends and decreases wages which redistributes from low to high productivity households thereby pushing up consumption inequality. Second, the increase in the real interest rate redistributes towards high wealth households but it is the high productivity households who eventually pay the tax burden. This slightly decreases the consumption of high productivity households and increases the consumption of middle productivity households who hold some assets but do not face tax increases. Thus, the second channel slightly dampens the increase in inequality and, as real interest rates increase by more, this channel is stronger in the behavioral HANK model.

Figure 19 shows the impulse-response functions of output, the output gap, inflation, nominal interest rates, government debt (as a share of annual GDP) and consumption inequality for the same cost-push shock but for the case in which monetary policy follows a simple Taylor rule with a response coefficient of 1.5.

Figure 18: Inflationary cost-push shock: strict inflation targeting

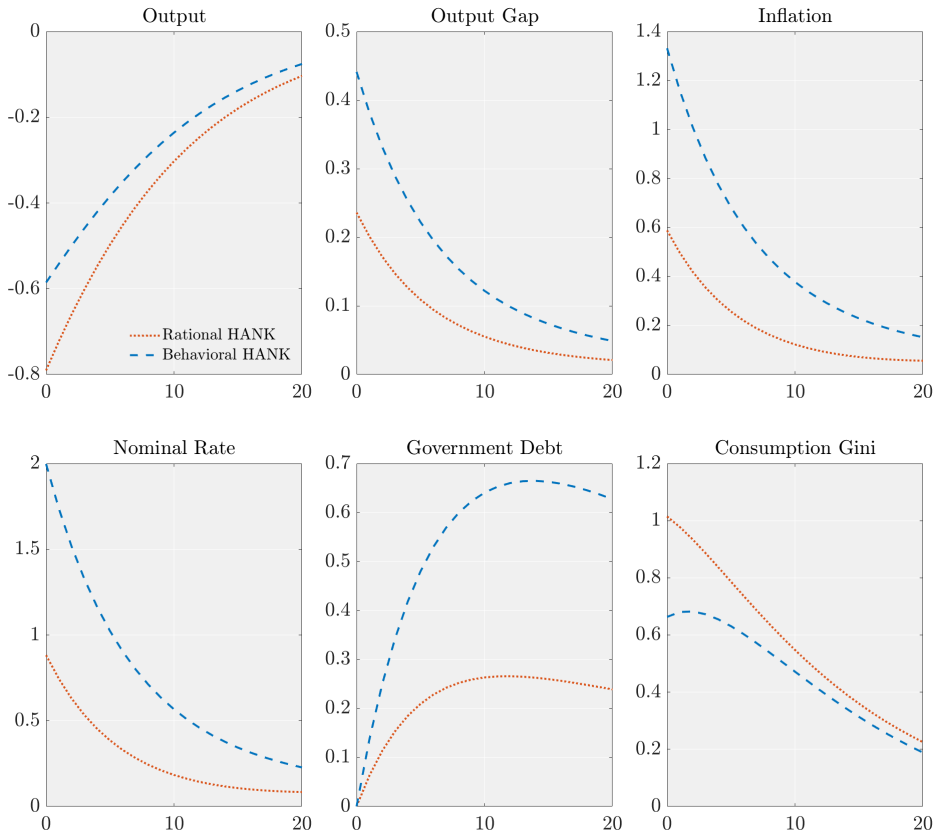


Note: This figure shows the impulse responses after a cost-push shock that decreases potential output by 1% in the inflation-stabilizing monetary policy regime. Output and the output gap are shown as percentage deviations from steady state output, the nominal interest rate and inflation as annualized percentage points and the government debt level as percentage point deviations of the debt-per-annual GDP level. The lower-right figure shows the change in the consumption Gini index as a percentage deviation from the stationary equilibrium.

As in the case where monetary policy fully stabilizes inflation, inflation and the nominal interest rate increase substantially more strongly in the behavioral HANK model than in its rational version. Also government debt increases more substantially.

Consumption inequality increases less strongly than with fully stabilizing inflation. The overheating economy—reflected in the positive output gap and increase in inflation—increases wages and decreases profits (relative to the inflation stabilizing regime) in the same way as expansionary policy shocks in Sections 3 and 4 do, thereby redistributing towards lower income households which dampens the increase in consumption inequality.

Figure 19: Inflationary cost-push shock: Taylor rule



Note: This figure shows the impulse responses after a cost-push shock that decreases potential output by 1% in the Taylor rule monetary policy regime. Output and the output gap are shown as percentage deviations from steady state output, the nominal interest rate and inflation as annualized percentage points and the government debt level as percentage point deviations of the debt-per-annual GDP level. The lower-right figure shows the change in the consumption Gini index as a percentage deviation from the stationary equilibrium.