A Behavioral Heterogeneous Agent New Keynesian Model

Oliver Pfäuti* Fabian Seyrich†

April 17, 2022
Link to most recent version

Abstract

We develop a New Keynesian model with household heterogeneity and bounded rationality in the form of cognitive discounting. The resulting behavioral heterogeneous agent New Keynesian model is consistent with recent empirical facts about the effectiveness and the transmission mechanisms of monetary and fiscal policy: monetary policy is amplified through indirect general equilibrium effects, fiscal multipliers on consumption are positive and the model delivers empirically-realistic intertemporal marginal propensities to consume. Simultaneously, and consistent with the data, the model resolves the forward guidance puzzle and remains stable at the effective lower bound as the model features equilibrium determinacy even under an interest-rate peg. The model is analytically tractable and nests a wide range of existing models as special cases, none of which can produce all the listed features within one model. We further show how the main insights from the tractable model extend to a quantitative version of the model, how the model-implied household expectations can be aligned with recent findings from survey data, and how to derive an equivalence result between heterogeneous-household models with bounded rationality and those featuring incomplete information and learning.

Keywords: Behavioral Macroeconomics, Heterogeneous Households, Monetary Policy, Forward Guidance, Fiscal Policy, New Keynesian Puzzles, Determinacy, Lower Bound

JEL Codes: E21, E52, E62, E71

*Department of Economics, University of Mannheim, oliver.pfaeuti@gess.uni-mannheim.de.

†Berlin School of Economics, DIW Berlin, and Freie Universität Berlin, fabian.seyrich@gmail.com.
We thank Klaus Adam, George-Marios Angeletos, Neele Balke, Nicholas Barberis, Florin Bilbiie, Eduardo Dávila, Zhen Huo, Alexander Kriwoluzky, Max Jager, Timo Reinelt, Hannah Seidl, Alp Simsek, Maximilian Weiß (discussant) and seminar and conference participants at Yale University, the SNDE Annual Symposium, the 2^{nd} edition of QuickTalks, the RCEA Conference 2022, the 15^{th} RGS Doctoral Conference, University of Mannheim, DIW Berlin and HU Berlin 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". First version online: January 2022.

1 Introduction

Recent empirical evidence on the transmission mechanisms and effectiveness of monetary and fiscal policies challenges the basic model of monetary policy, the New Keynesian model, along many dimensions: Monetary policy is transmitted to household consumption to a large extent through indirect, general equilibrium effects. Government spending increases private consumption substantially. Households' marginal propensities to consume (MPCs) out of transitory income changes are high on average in the year of the received income windfall and remain high even in the year after. Announcements of future monetary policy changes, on the other hand, have relatively weak effects on current economic activity. Despite these weak effects of forward guidance, advanced economies have not faced large instabilities during long spells at the binding effective lower bound.[1]

In this paper, we propose a new framework which can account for all these empirical facts *simultaneously*. To this end, we develop a New Keynesian model featuring household heterogeneity and bounded rationality in the form of cognitive discounting. The resulting behavioral heterogeneous agent New Keynesian model—or *behavioral HANK*—is analytically tractable which allows for a clear understanding of the role of the two frictions and how they interact. We show that it is indeed the *interaction* of bounded rationality and household heterogeneity that allows our model to be reconciled with the empirical evidence. Moreover the model nests a broad spectrum of existing models—including the standard New Keynesian model, rational HANK models, and representative agent models which depart from the full-information rational expectations hypothesis (FIRE). None of these other models, however, can account for all the listed empirical facts simultaneously.

To arrive at our framework, we extend the textbook New Keynesian model by household heterogeneity and bounded rationality in ways that preserve the tractability of the model. There are two groups of households. One group of households are "unconstrained", in the sense that they participate in financial markets so that they are on their Euler equation. The other group of households are off their Euler equation and consume all their disposable income. We thus refer to these households as "hand-to-mouth". Households face an idiosyncratic risk of switching from one type to the other. This set-up generates heterogeneity in income, MPCs, and precautionary-savings motives. We introduce bounded rationality in the

[1] See, e.g., Ampudia et al. (2018), Slacalek et al. (2020) Samarina and Nguyen (2019) and Holm et al. (2021) for the empirical relevance of indirect channels in the transmission of monetary policy, Galí et al. (2007), Perotti (2007) or Dupor et al. (2021) for empirical evidence on the positive consumption response to fiscal spending, Auclert et al. (2018), Fagereng et al. (2021), Jappelli and Pistaferri (2020), Auclert (2019) and Patterson (2019) document empirical patterns of MPCs and see, for example, Del Negro et al. (2015), D'Acunto et al. (2020), Miescu (2022) and Roth et al. (2021) for empirical evidence on the (in-)effectiveness of forward guidance and Debortoli et al. (2020) and Cochrane (2018) on the stability at the lower bound.

form of cognitive discounting. Households anchor their expectations about future macroeconomic variables to the steady state but cognitively discount expected future deviations from it. As a result, average expectations underreact to news, as empirically documented in Coibion and Gorodnichenko (2015), Bordalo et al. (2020) and Angeletos et al. (2021).[2]

In the behavioral HANK model, indirect general equilibrium effects account for large parts of how monetary policy is transmitted to consumption. Consistent with the data, hand-to-mouth households that exhibit high MPCs are more exposed to the business cycle. Thus, after an expansionary monetary policy shock (and likewise after a fiscal spending shock), high MPC households disproportionately benefit from the increase in output. This leads to an amplification of contemporaneous monetary policy through general equilibrium via a Keynesian-type multiplier. In addition, in the useful benchmark of a constant real interest rate, these general equilibrium effects also generate positive fiscal multipliers on consumption.

Even though the behavioral HANK model generates amplification of contemporaneous shocks through indirect effects the model does not suffer from the forward guidance puzzle. Announced changes of future interest rates have weaker effects on today's output than a current change in the interest rate and the effectiveness on today's output decreases with the horizon of the announcement. There are two competing forces shaping the effectiveness of a forward guidance shock on today's output. First, the general equilibrium amplification channel that is at work in response to contemporaneous monetary policy shocks is, *ceteris paribus*, compounded over time. The reason is that when unconstrained households expect lower interest rates in the future, they decrease their precautionary savings today as they would disproportionately benefit from the associated increase in output in the hand-to-mouth state. Yet, the behavioral agents cognitively discount both this indirect general equilibrium effect as well as the direct effects of the future interest rate changes which dampens the effects of forward guidance. We find that this second channel dominates, thus, the model resolves the forward guidance puzzle. Additionally, the behavioral HANK model remains stable during prolonged periods at the effective lower bound (ELB), as the model features equilibrium determinacy under an interest-rate peg for a large area of the parameter space.

A key statistic in HANK models for monetary and fiscal policy analyses are the intertemporal MPCs—or iMPCs for short (Auclert et al. (2018), Wolf (2021), Kaplan and Violante (2020)). We derive the iMPCs in the behavioral HANK model analytically. To the best of our knowledge, we are the first ones to do so in a HANK model featuring a departure from

[2]We show how to microfound cognitive discounting as a noisy-signal extraction problem of otherwise rational agents. Angeletos and Lian (2017) show how other forms of bounded rationality or lack of common knowledge lead to observationally-equivalent outcomes.

FIRE. The behavioral HANK model quantitatively matches the empirical iMPCs. Thanks to the closed-form solution the model also sheds light on how the iMPCs depend on household heterogeneity frictions and bounded rationality. Boundedly-rational households tend to save more than rational households out of the income windfall as they cognitively discount the decrease in their future marginal utility which lowers the current MPC. As time progresses, however, bounded rationality increases the aggregate MPC as the behavioral agents start to consume their (higher) savings. These dynamic effects are particularly pronounced when idiosyncratic risk is relatively high.

The behavioral HANK model can have *qualitatively different* policy implications than its rational counterpart. Consider an overheating economy which the monetary authority wants to tame by hiking interest-rates by a cumulative $x\%$. This rate hike can be implemented immediately or by raising the rate $\frac{x}{k}\%$ over k consecutive periods. A well-known feature of the RANK model is that monetary policy becomes more effective the more it is back-loaded. While this is also the case in the rational HANK model the opposite is true in the behavioral HANK model: monetary policy is more effective when it is completely *front-loaded*, i.e., when $k = 1$. The increased effectiveness is driven by the fact that the hand-to-mouth agents' incomes contract more strongly leading to a strong decrease in aggregate demand. Thus, the increased effectiveness of front-loading the policy comes at the cost of an increase in inequality.

We show that the main insights of our tractable behavioral HANK model carry over to a more quantitative version of our model. To this end, we construct a quantitative behavioral HANK model which, in its rational expectation limit, collapses to a standard one-asset HANK model. The same general equilibrium forces as in the tractable model lead to an amplification of contemporaneous monetary policy shocks. Yet, also in the quantitative behavioral HANK model, there is no forward guidance puzzle, but the effectiveness of a change in the interest rate declines in the horizon. This is in contrast to the rational HANK model, in which the forward guidance puzzle is aggravated relative to the representative-agent model.

We extend our tractable baseline framework in several dimensions. First, we allow for sticky wages and show how the interplay of sticky wages, household heterogeneity and bounded rationality leads to hump-shaped responses of macroeconomic variables in response to aggregate shocks, as documented empirically (see, e.g., Auclert et al. (2020)). When forming their expectations the behavioral households do not fully incorporate the implications of wage stickiness on future consumption in different states and, thus, on their idiosyncratic risk. As a consequence the economy grows stronger than expected during the first quarters after the shock which generates a hump-shaped response. We also show that the interaction of bounded rationality, sticky wages and household heterogeneity generates an initial underreaction of households' expectations about future output, followed by a delayed overshooting, which is consistent with recent findings from survey expectations data (see Angeletos et al. (2021) and Adam et al. (2022)). This is the case although in our setup expectations are purely forward looking.

Second, we show how to extend our framework to derive an equivalence result between models with bounded rationality and models of incomplete information and learning. To this end, we assume that behavioral agents anchor their beliefs to *past observations* of the respective variable instead of the respective steady state values. This extended behavioral HANK model is observationally equivalent to models featuring incomplete information and learning (see Angeletos and Huo (2021) and Gallegos (2021)) that induce myopia and anchoring in the aggregate IS equation.[3]

Related Literature. The literature so far treats the empirical facts laid out in the Introduction mostly independently from each other. The HANK and TANK literature – both with quantitative and analytical models – have highlighted the transmission of monetary policy through indirect, general equilibrium effects (Kaplan et al. (2018), Auclert (2019), Auclert et al. (2020), Bilbiie (2020)), positive fiscal multipliers on consumption (Auclert et al. (2018), Galí et al. (2007)), the role of iMPCs (Auclert et al. (2018), Cantore and Freund (2021), Kaplan and Violante (2020)), and potential resolutions of the forward guidance puzzle (McKay et al. (2016), McKay et al. (2017), Hagedorn et al. (2019)).

Werning (2015) and Bilbiie (2021) combine the themes of policy amplification and forward guidance puzzle in HANK. While these two papers focus on slightly different explanation mechanisms, both establish a trade-off inherent in models with household heterogeneity: if HANK models amplify contemporaneous monetary policy (and fiscal policy) through redistributing towards high MPC households, they also dampen precautionary savings desires after a forward guidance shock which aggravates the forward guidance puzzle.[4] One of our contributions is that our behavioral HANK model overcomes this so-called *Catch-22* (Bilbiie (2021)).[5]

[3]Angeletos and Huo (2021) derive an equivalence result between models with incomplete information and learning with models that impose behavioral myopia and an additional friction such as habit persistence or adjustment costs. We now complement their equivalence result with a behavioral model that solely relies on one behavioral friction.

[4]Acharya and Dogra (2020) construct a pseudo-RANK model, in which they isolate and highlight the role of precautionary savings dynamics highlight the role of precautionary savings dynamics in explaining the solution or aggravation of the forward guidance puzzle.

[5]Bilbiie (2021) provides two theoretical possibilities of how to sidestep the Catch-22. The first possibility is a pure risk channel which can, in theory, break the comovement of income risk and inequality. Yet to do so, it requires a calibration which seems highly at odds with the data. A second possibility is to drastically

A mostly-detached strand of the literature relaxes the assumption of full-information rational expectations (FIRE) to weaken the effectiveness of future monetary policies, thereby resolving the forward guidance puzzle (Wiederholt (2015), Angeletos and Lian (2018), Andrade et al. (2019), Gabaix (2020), Pfäuti (2021) and Roth et al. (2021)). We complement these papers by introducing household heterogeneity in terms of iMPCs, asset-market participation status, and exposure to the business cycle. This way, our model not only resolves the forward guidance puzzle (and other NK puzzles) but also generates amplification of contemporaneous monetary and fiscal policy through indirect GE channels, as well as it matches empirical estimates of iMPCs.

We share the combination of household heterogeneity and some deviation from FIRE with Farhi and Werning (2019), Auclert et al. (2020), Broer et al. (2021a), Angeletos and Huo (2021), Laibson et al. (2021), Gallegos (2021), and Bonciani and Oh (2022). In contrast to all these papers, we offer analytical insights into how the two frictions matter for policy analysis, and how the interaction of the two frictions is key to reconcile the model with recent empirical facts outlined above.

Outline. The rest of the paper is structured as follows. We present our behavioral HANK model in Section 2 and our main analytical results in Section 3. In Section 4, we discuss several model extensions, and Section 5 concludes.

2 A Behavioral HANK Model

In this section, we present our tractable New Keynesian model featuring household heterogeneity and bounded rationality (BR).

2.1 Structure of the Model

Households. The economy is populated by a unit mass of households, indexed by $i \in [0, 1]$. Households obtain utility from (non-durable) consumption, C_t^i, and dis-utility from working N_t^i. Households discount future utility at rate $\beta \in [0, 1]$. We assume a standard CRRA

narrow down the policy space: in a world in which monetary policy is described by Wicksellian price level targeting or fiscal policy follows a nominal bond rule, there would be no Catch-22. Hagedorn et al. (2019) use a similar description of fiscal policy to solve the forward guidance puzzle in a quantitative HANK model, in which contemporaneous monetary policy is amplified. Similarly, Kaplan et al. (2016) show that in their quantitative HANK model in Kaplan et al. (2018), there is no Forward Guidance puzzle, conditional on specific fiscal policy responses to a monetary policy shock. In contrast, in our model, there is no Catch-22 *independently* of the exact specification of monetary and fiscal policy.

utility function

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

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

Households can save in government bonds, paying nominal interest i_t, and they can acquire shares ι_t of intermediate monopolistic firms, introduced later. Households face an exogenous borrowing constraint which we set to zero. We allow for the possibility that households participate in financial markets infrequently. When they do participate, they can freely buy or sell bonds and shares and receive the intermediate firm profits, D_t. Otherwise, they simply receive the payoff from their previously acquired bonds. For now, asset-market participation is exogenous and can be interpreted, for example, as a shock to the household's taste or patience. In Section 3.7, we endogenize asset-market participation. We denote households participating in financial markets by U as, in equilibrium, they will be Unconstrained in the sense that they are on their Euler equation. We denote the non-participants by H as they will be off their Euler equation and, thus, $H$and-to-mouth. An unconstrained household remains unconstrained with probability s and becomes hand-to-mouth with probability $1-s$. Hand-to-mouth households remain hand-to-mouth with probability h and switch to being unconstrained with probability $1-h$. In what follows, we focus on stationary equilibria where $\lambda \equiv \frac{1-s}{2-s-h}$ denotes the constant share of hand-to-mouths. We relax the assumption of a constant λ in Section 3.7 in which the share of hand-to-mouth households will be endogenous.

We use the same simplifying assumptions as in Bilbiie (2021) to obtain an analytically-tractable solution. In particular, we assume that households belong to a family whose utilitarian intertemporal welfare is maximized by its family head. The head can only provide insurance within types but not across types, i.e., the head pools all the resources within types. Thus, in equilibrium every U household will consume and work the same amount and every H household will consume and work the same amount but the H households' consumption and labor supply is not necessarily the same as the U households' consumption and labor supply. When households switch from being unconstrained to the hand-to-mouth type, they only keep their government bonds. As a consequence, shares cannot be used to self-insure. In Section 3.7, we relax these assumptions and we allow for a non-degenerate asset distribution.

Let Z_{t+1}^i denote the bonds acquired by type $i \in \{U, H\}$ in period t before the realization of the idiosyncratic risk takes place. After the realization of the idiosyncratic risk and the in- and outflows between types at the end-of-period t, the bonds of type $i \in \{U, H\}$ at the beginning-of-period $t+1$ are denoted by B_{t+1}^i. Given the transition probabilities and the stationary equilibrium definition, it follows that

$$(1-\lambda)B_{t+1}^{U} = s(1-\lambda)Z_{t+1}^{U} + (1-h)\lambda Z_{t+1}^{H}$$
$$\lambda B_{t+1}^{H} = (1-s)(1-\lambda)Z_{t+1}^{U} + h\lambda Z_{t+1}^{H},$$

which, after using the definition of λ, can be re-written as

$$\begin{aligned} B_{t+1}^{U} &= sZ_{t+1}^{U} + (1-s)Z_{t+1}^{H} \\ B_{t+1}^{H} &= (1-h)Z_{t+1}^{U} + hZ_{t+1}^{H}. \end{aligned} \tag{2}$$

We allow for the possibility that the family head is boundedly rational in the way we describe in detail in Section 2.3.[6] The program of the family head is

$$V\left(B_t^U, B_t^H, \iota_t\right) = \max_{\{C_t^U, C_t^H, Z_{t+1}^U, Z_{t+1}^H, N_t^U, N_t^H, \iota_{t+1}\}} \Big[(1-\lambda)\mathcal{U}\left(C_t^U, N_t^U\right) + \lambda\mathcal{U}\left(C_t^H, N_t^H\right)\Big]$$
$$+\beta\mathbb{E}_t^{BR} V\left(B_{t+1}^U, B_{t+1}^H, \iota_{t+1}\right)$$

subject to the flow budget constraints of unconstrained households

$$C_t^U + Z_{t+1}^U + v_t\iota_{t+1} = W_t N_t^U + \iota_t(v_t + \tilde{D}_t) + \frac{1+i_{t-1}}{1+\pi_t}B_t^U + T_t^U, \tag{3}$$

and the hand-to-mouth households

$$C_t^H + Z_{t+1}^H = W_t N_t^H + T_t^H + \frac{1+i_{t-1}}{1+\pi_t}B_t^H, \tag{4}$$

as well as the borrowing constraints

$$Z_{t+1}^H, Z_{t+1}^U \geq 0,$$

where W_t is the real wage, ι_t are the shares of stocks traded at price v_t, and T_t^i are transfers to type-i households. As we will detail below, we assume that these transfers are financed by a proportional tax on profits, τ^D, such that they entail a redistribution from U households (who receive the profits) to H households. The family head takes these transfers as given. $\tilde{D}_t$ denotes the after-tax profits of the intermediate firms.

The optimality conditions are given by the Euler equations of unconstrained households

[6] We show in Appendix A.8 how the family head's expectation can be understood as an average expectation over all households' expectations within family where each household receives a noisy signal about the future state.

and the hand-to-mouths' households

$$\frac{\partial \mathcal{U}(C_t^U, N_t^U)}{\partial C_t^U} \geq \beta \mathbb{E}_t^{BR}\left[R_t\left(s\frac{\partial \mathcal{U}(C_{t+1}^U, N_{t+1}^U)}{\partial C_{t+1}^U} + (1-s)\frac{\partial \mathcal{U}(C_{t+1}^H, N_{t+1}^H)}{\partial C_{t+1}^H}\right)\right] \tag{5}$$

$$\frac{\partial \mathcal{U}(C_t^H, N_t^H)}{\partial C_t^H} \geq \beta \mathbb{E}_t^{BR}\left[R_t\left((1-h)\frac{\partial \mathcal{U}(C_{t+1}^U, N_{t+1}^U)}{\partial C_{t+1}^U} + h\frac{\partial \mathcal{U}(C_{t+1}^H, N_{t+1}^H)}{\partial C_{t+1}^H}\right)\right], \tag{6}$$

which hold with equality if the respective borrowing constraint does not bind and with inequality in case it binds. $R_t \equiv \frac{1+i_t}{1+\pi_{t+1}}$ denotes today's real interest rate. Furthermore, we obtain the demand for shares

$$\frac{\partial \mathcal{U}(C_t^U, N_t^U)}{\partial C_t^U} \geq \beta \mathbb{E}_t^{BR}\left[\frac{v_{t+1} + D_{t+1}}{v_t}\frac{\partial \mathcal{U}(C_{t+1}^U, N_{t+1}^U)}{\partial C_{t+1}^U}\right]. \tag{7}$$

The respective labor-leisure equations of both types are given by:

$$-\frac{\partial \mathcal{U}(C_t^U, N_t^U)}{\partial N_t^U} = W_t\frac{\partial \mathcal{U}(C_t^U, N_t^U)}{\partial C_t^U} \tag{8}$$

$$-\frac{\partial \mathcal{U}(C_t^H, N_t^H)}{\partial N_t^H} = W_t\frac{\partial \mathcal{U}(C_t^H, N_t^H)}{\partial C_t^H}. \tag{9}$$

In what follows, we focus on equilibria in which the H households are always off their Euler equation—as they are not participating in financial markets—such that equation (6) holds with strict inequality. In addition, we follow the tradition of analytical HANK models and assume a zero liquidity equilibrium to keep our model tractable.[7] We relax this later and show that our results do not hinge on this assumption. As shares cannot be transferred to the H state, equation (7) simply prices the shares. Thus, the bond Euler equation of unconstrained households (5) is the only Euler equation that is an equilibrium equation. Importantly, it features a self-insurance motive as unconstrained households demand bonds to self-insure their idiosyncratic risk of type-switching.

Firms. We assume a standard NK firm side. All households consume the same aggregate basket of individual goods, $j \in [0,1]$,

$$C_t = \left(\int_0^1 C_t(j)^{\frac{\epsilon-1}{\epsilon}} dj\right)^{\frac{\epsilon}{\epsilon-1}}$$

[7] See Krusell et al. (2011), McKay et al. (2017), Ravn and Sterk (2017), and Bilbiie (2021).

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

$$Y_t(j) = N_t(j).$$

The real marginal cost is given by W_t. We assume that the government pays a subsidy τ^S on revenues to induce marginal cost pricing. The 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. Given zero steady state profits, we have a full-insurance steady state, i.e., $C^H = C^U = C$. In the log-linear dynamics around this steady state, profits vary inversely with the real wage $\widehat{d}_t = -\widehat{w}_t$.[8] We allow for steady state inequality in Appendix C and show that our results are not driven by this assumption and are in fact barely affected even by substantial inequality in the steady state.

Government. Fiscal policy induces the optimal steady state subsidy financed by lump-sum taxation of firms and taxes profits at rate τ^D and rebates these taxes as a transfer to H households, such that

$$T^H = \frac{\tau^D}{\lambda} D_t.$$

As will become clear later the level of τ^D is key for the exposure of H households to the business cycle and thus for the cyclicality of inequality. We set $T_t^U = 0$ and we abstract from government spending for now, but introduce it in Section 3 to study fiscal multipliers.

In most of the analysis, we assume that monetary policy follows a standard (log-linearized) Taylor rule[9]

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

[8]Throughout the paper variables with a hat on top denote log-deviations from steady state.
[9]We study more general Taylor rules in Appendix A.

with ϵ_t^{MP} being a monetary policy shock.

Market Clearing. Market clearing requires that the goods market clears

$$Y_t = C_t = \lambda C_t^H + (1-\lambda) C_t^U$$

and the labor market clears

$$N_t = \lambda N_t^H + (1-\lambda) N_t^U.$$

Bond market clearing implies

$$B_t^U = Z_t^U = 0$$

at all t.

2.2 Log-Linearized Model

We now focus on the log-linearized dynamics around the full-insurance, zero-liquidity steady state. The market clearing conditions yield $\widehat{y}_t = \widehat{c}_t = \lambda \widehat{c}_t^H + (1-\lambda)\widehat{c}_t^U$ and $\widehat{n}_t = \lambda \widehat{n}_t^H + (1-\lambda)\widehat{n}_t^U$. Importantly, we can write the consumption of the hand-to-mouth households as

$$\widehat{c}_t^H = \chi \widehat{y}_t, \tag{11}$$

with

$$\chi \equiv 1 + \varphi \left(1 - \frac{\tau^D}{\lambda}\right) \tag{12}$$

measuring the cyclicality of the H household's consumption.[10] Auclert (2019) and Patterson (2019) document that households with higher MPCs tend to be more exposed to aggregate income fluctuations, which is achieved by setting $\chi > 1$. As χ is a key coefficient in our model, we will vary χ throughout the paper to understand its role in shaping our results. Different levels of χ should then be thought of as different τ^D, thus, different redistributive tax-transfer systems.

Combining equation (11) with the goods market clearing condition yields

$$\widehat{c}_t^U = \frac{1-\lambda\chi}{1-\lambda}\widehat{y}_t, \tag{13}$$

which implies that consumption inequality is given by:

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

[10] See Appendix A.1 for the derivation of equation (11).

Thus, if $\chi > 1$, inequality is countercyclical as it varies negatively with total output, i.e., increases in recessions and decreases in booms. In line with the empirical evidence on the covariance between MPCs and business-cycle exposure the data also points towards $\chi > 1$ when looking at the cyclicality of inequality. Coibion et al. (2017), Mumtaz and Theophilopoulou (2017), Ampudia et al. (2018) and Samarina and Nguyen (2019) provide evidence of countercyclical inequality conditional on monetary policy shocks.

The log-linearized bond Euler equation of U households is given by

$$\widehat{c}_t^U = s\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] + (1-s)\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^H\right] - \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t^{BR}\pi_{t+1}\right). \tag{15}$$

We will, following the assumption in Gabaix (2020), often focus on the case in which households are rational with respect to the real rate, i.e., we replace $\mathbb{E}_t^{BR}\pi_{t+1}$ with $\mathbb{E}_t\pi_{t+1}$ in equation (15). We show in Appendix C that our results go through with boundedly-rational real-rate expectations. In fact, the model is even more likely to feature equilibrium determinacy with boundedly-rational real-rate expectations. For the case without type-switching, i.e., for $s = 1$, equation (15) boils down to a standard Euler equation. For $s \in [0, 1)$, however, the agent takes into account that she might switch her type and self-insures against becoming hand-to-mouth next period.

Supply Side. We distinguish between two set-ups for the supply side: For the main part, we work with a static Phillips Curve

$$\pi_t = \kappa\widehat{y}_t, \tag{16}$$

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)). The other setup considers a standard forward-looking New Keynesian Phillips Curve. We discuss this case in Appendix C and show that a forward-looking Phillips Curve barely affects our results.

2.3 Bounded Rationality

We follow Gabaix (2020) and model bounded rationality in the form of cognitive discounting.[11] Let X_t be a random variable (or vector of variables) and let us define X_t^d as some

[11] 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, 2016)) and a handbook treatment of behavioral inattention is given in Gabaix (2019). Benchimol and Bounader (2019) and Bonciani and Oh (2021) study optimal monetary policy in a RANK and TANK model, respectively, with this kind of

default value the agent may have in mind and $\tilde{X}_{t+1} \equiv X_{t+1} - X_t^d$ denotes the deviation from this default value.[12] The behavioral agent's expectation about X_{t+1} is then defined as

$$\mathbb{E}_t^{BR}\left[X_{t+1}\right] = \mathbb{E}_t^{BR}\left[X_t^d + \tilde{X}_{t+1}\right] \equiv X_t^d + \bar{m}\mathbb{E}_t\left[\tilde{X}_{t+1}\right], \tag{17}$$

where $\mathbb{E}_t\left[\cdot\right]$ is the rational expectations operator and $\bar{m} \in [0,1]$ is the behavioral parameter capturing the degree of rationality. A higher $\bar{m}$ denotes a smaller deviation from rational expectations and rational expectations are captured by $\bar{m} = 1$. The behavioral agent anchors her expectations to the default value and cognitively discounts expected future deviations from this default value. For now, we focus on the steady state as the default value but relax this assumption in Section 4.2.

While we present a way to microfound $\bar{m}$ in Appendix A.8, note, that the exact microfoundation or underlying behavioral friction is not crucial for the rest of our analysis. For example, Angeletos and Lian (2017) show how other forms of bounded rationality or lack of common knowledge lead to observationally-equivalent expectations for the case in which X_t^d denotes the steady state.

Log-linearizing equation (17) around the steady state yields

$$\mathbb{E}_t^{BR}\left[\widehat{x}_{t+1}\right] = (1-\bar{m})\widehat{x}_t^d + \bar{m}\mathbb{E}_t\left[\widehat{x}_{t+1}\right] \tag{18}$$

and when X_t^d is the steady state value, we obtain $\mathbb{E}_t^{BR}\left[\widehat{x}_{t+1}\right] = \bar{m}\mathbb{E}_t\left[\widehat{x}_{t+1}\right]$. In Appendix B, we discuss empirical estimates of $\bar{m}$ and how we can map recent evidence in Coibion and Gorodnichenko (2015) and Angeletos et al. (2021) to $\bar{m}$. As a benchmark, we follow Gabaix (2020) and set $\bar{m}$ to 0.85, which is a rather conservative choice, given the empirical evidence. As one goal of our paper is to understand the role of $\bar{m}$ for policy analysis and the interplay of $\bar{m}$ and household heterogeneity, we will also consider different values for $\bar{m}$.

3 Results

In this section, we derive the three-equation representation of the behavioral HANK model and show that the model is consistent with all the discussed empirical facts. The model nests a wide spectrum of existing models—none of which can account for all the empirical

behavioral frictions.

[12] 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. We, on the other hand, directly apply cognitive discounting to all variables. Given Lemma 1 in Gabaix (2020), our results would be unchanged, but our more direct method simplifies some of the derivations, especially in Section 4.2. Appendix A.7 derives our results following the approach in Gabaix (2020).

facts simultaneously. We then illustrate how the behavioral HANK model leads to different policy implications than its rational counterpart and we end the section by constructing a quantitative behavioral HANK model to show that the main insights from the tractable model carry over to more quantitative models.

3.1 The Three-Equation Representation

The behavioral HANK model can be summarized by three equations: a Phillips curve, representing the aggregate supply side captured by equation (16), and a rule for monetary policy (equation (10)), which together with the *behavioral HANK IS equation* determines aggregate demand. To obtain the behavioral HANK IS equation, we combine the hand-to-mouth households' consumption (11) with the consumption of unconstrained households (13) and their consumption Euler equation (15).[13]

Proposition 1. *The behavioral HANK IS equation is given by*

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t \pi_{t+1}\right), \tag{19}$$

where

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

and

$$\psi_c \equiv \frac{1-\lambda}{1-\lambda\chi}.$$

Compared to RANK, two extra 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 business-cycle exposure χ. As the H households are more exposed to the business cycle ($\chi > 1$), $\psi_c > 1$ and current output is more sensitive to changes in the real interest rate due to general equilibrium forces, as we discuss later.

The second new coefficient in the behavioral HANK IS equation (19), 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 cyclicality of income risk *and* the degree of bounded rationality of households as well as the interaction of

[13] All derivations are in Appendix A.

the two frictions. Given countercyclical income inequality, income risk is also countercyclical which manifests itself in $\delta > 1$. This 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—the 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 which is key to resolve the forward guidance puzzle as well as to obtain a determinate, locally unique equilibrium.

Equation (19) nests a wide range of existing IS equations: the IS equation in the standard rational-expectations RANK model by setting $\psi_f = \psi_c = 1$, RANK models deviating from FIRE by $\delta = \psi_c = 1$, TANK models by setting $\bar{m} = \psi_f = 1$, and rational HANK models by $\bar{m} = 1$.[14] We discuss this in more detail in section 3.4.

Baseline Calibration. We set the parameters close to the calibration in Bilbiie (2020) and Bilbiie (2021) which is set in order to replicate several findings on the New Keynesian cross coming from more quantitative HANK models. We set $\chi = 1.48$ which implies that H agents' income is relatively sensitive to aggregate fluctuations, in line with empirical findings in Auclert (2019) and Patterson (2019). We set the share of H agents to one third, $\lambda = 0.33$, and the probability of an 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, $\gamma = 1$, and set the slope of the Phillips Curve to $\kappa = 0.02$. The cognitive discounting parameter, $\bar{m}$ is set to 0.85, as explained in Section 2.3. Details on the calibration and a discussion of the robustness of our findings for different calibrations are presented in Appendix B. Note, that even when we vary certain parameters, we always keep $\lambda < \chi^{-1}$.

3.2 Monetary Policy

We now show how the behavioral HANK model generates amplification of contemporaneous monetary policy through indirect effects while solving 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.

[14] For the RANK model, see, for example, Woodford (2003) or Galí (2015), for the RANK models differing from FIRE, see, for example, Angeletos and Lian (2018), Woodford (2019), or Gabaix (2020), and for rational TANK or THANK models, see Bilbiie (2008), McKay et al. (2017) or Bilbiie (2021).

To derive these results, it is convenient to represent the model in a single first-order difference equation:

$$\widehat{y}_t = \frac{\psi_f + \psi_c \frac{\kappa}{\gamma}}{1 + \psi_c \phi \frac{\kappa}{\gamma}} \mathbb{E}_t \widehat{y}_{t+1} - \frac{\psi_c \frac{1}{\gamma}}{1 + \psi_c \phi \frac{\kappa}{\gamma}} \varepsilon_t^{MP}, \tag{20}$$

which we obtain by combining the IS equation (19) with the static Phillips Curve (16) and the Taylor rule (10).

General Equilibrium Amplification and Forward Guidance. We start by showing how the behavioral HANK model generates general equilibrium amplification of current monetary policy, 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. Miescu (2022) provides empirical evidence that conventional monetary policy is more effective than forward guidance. Consistent with these findings, Roth et al. (2021) combine experimental evidence with theory to show that forward guidance has relatively weak effects on consumption.

Let us now consider two different monetary policy experiments: (i) a contemporaneous monetary policy shock, i.e., a surprise decrease in the interest rate today, and (ii) a forward guidance shock, i.e., a news shock today about a decrease in the interest rate k periods in the future. In both cases, we focus on *i.i.d.* shocks and the case with $\phi = 0$, as in Bilbiie (2021). 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. Similarly, assuming completely fixed prices ($\kappa = 0$), as for example in Farhi and Werning (2019), or modelling forward guidance as changes in the *real* interest rate, as for example in McKay et al. (2016), would also leave the amplification condition unaltered but relax the condition to rule out the forward guidance puzzle.

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

$$\chi > 1, \tag{21}$$

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

and the forward guidance puzzle is ruled out if

$$\bar{m}\delta + \frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\kappa < 1. \tag{22}$$

The behavioral HANK model generates *amplification* of contemporaneous monetary policy with respect to the RANK model whenver $\chi > 1$, that is, when high-MPC households' consumption is relatively sensitive to aggregate income fluctuations. As discussed in Section 2.2, this is the case empirically. With $\chi > 1$ the income of H agents moves more than one for one with aggregate output. Hence, after a decrease in the interest rate, a disproportionate share of the extra income is received by H agents and, thus, the high-MPC households in the economy. As a result, $\psi_c > 1$ and the increase in output is amplified through general equilibrium. The behavioral friction leaves the relative importance of direct vs. indirect effects—i.e., amplification of contemporaneous monetary policy—unaltered, as amplification of a contemporaneous i.i.d. shock is solely determined by the static redistribution towards the high MPC households. It is thus through these indirect, general equilibrium, effects that monetary policy gets amplified as the H households do not directly respond to interest rate changes because they do not participate in asset markets.

Turning to forward guidance, note, that the forward guidance puzzle is ruled out if the term $\frac{\psi_f+\psi_c\frac{\kappa}{\gamma}}{1+\psi_c\phi\frac{\kappa}{\gamma}}$ in front of $\mathbb{E}_t\widehat{y}_{t+1}$ in the first-order difference equation (20) is smaller than 1. Given that we consider $\phi = 0$, this boils down to

$$\begin{aligned} & \psi_f + \psi_c\frac{\kappa}{\gamma} < 1, \\ \Leftrightarrow \quad & \bar{m}\delta + \frac{1-\lambda}{1-\lambda\chi}\frac{\kappa}{\gamma} < 1, \end{aligned} \tag{23}$$

which is the condition stated in Proposition 2.

What determines whether condition (23) holds or not? First, note that as in the discussion of contemporaneous monetary policy, it is still the case that 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 since if they become hand-to-mouth they would benefit more 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. Yet, as households are boundedly rational, they cognitively discount these effects taking place in the future. Importantly, unconstrained households cognitively discount both the future increase in output as well as the general equilibrium implications for their precautionary savings, thereby decreasing the effects of the forward guidance shock on today's consumption. Given our calibration there is no forward guidance puzzle in the behavioral HANK model as long as $\bar{m} < 0.93$.

We now compare the behavioral HANK model to its rational counterpart to show how the behavioral HANK model overcomes a major trade-off inherent in the rational HANK model – the *Catch-22* (Bilbiie (2021)). The Catch-22 describes the trade-off 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

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

which requires

$$\chi < 1,$$

as otherwise $\delta > 1$. Assuming $\chi < 1$, however, leads to *dampening* of contemporaneous monetary policy instead of amplification.

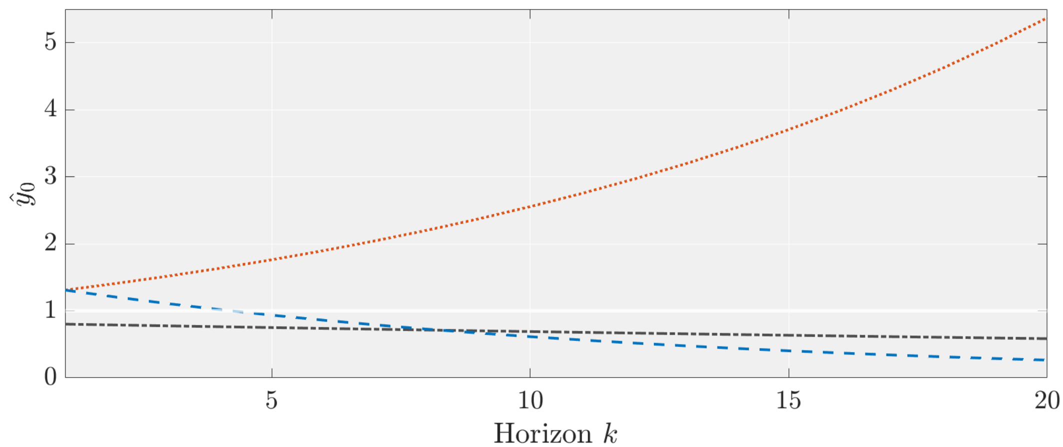
We graphically illustrate the Catch-22 of the rational model and the resolution of it in the behavioral HANK model in Figure 1. The figure shows the response of contemporaneous output relative to the initial response in the RANK model with rational expectations for anticipated i.i.d. monetary policy shocks occurring at different times k and a Taylor coefficient of 0.[16]

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 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 (Auclert (2019)).

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. Note that also rational TANK models (thus, turning off type switching) or the behavioral RANK model would not deliver

[16] Under fully-rigid prices (i.e., $\kappa = 0$) the RANK model would deliver a constant response for all k. The same is true for TANK, i.e., THANK 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 $\chi \lesseqgtr 1$.

Figure 1: 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).

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.

As a direct consequence of the resolution of the Catch-22 in the behavioral HANK model, highly persistent monetary policy shocks have smaller effects on contemporaneous output than in RANK whereas less persistent shocks have relatively larger effects in the behavioral HANK model. The reason is that persistent shocks also work through a forward guidance channel which is dampened in the behavioral HANK model. We elaborate this point in more detail in Appendix C.2.

Determinacy in Behavioral HANK. According to the Taylor principle, monetary policy needs to respond sufficiently strongly to changes in inflation in order to have 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.[17]

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

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

$$\phi > \phi^* = 1 + \frac{\bar{m}\delta - 1}{\frac{\kappa}{\gamma}\frac{1-\lambda}{1-\lambda\chi}}. \tag{24}$$

We obtain Proposition 3 directly from the difference equation (20). For determinacy, we need that the coefficient in front of $\mathbb{E}_t \widehat{y}_{t+1}$ is smaller than 1. Solving this condition for ϕ yields Propositon 3. Appendix A.5 outlines the details and extends the result to more general Taylor rules.

To understand condition (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 the rational HANK model has been shown by Bilbiie (2021) and in a similar way by 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 RANK Taylor principle in the behavioral HANK model is 0.95. What is more, given our baseline choice of $\bar{m} = 0.85$, we have $\phi^* = -3.07$. Thus, in the 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. In this sense the behavioral HANK model overcomes the famous result in Sargent and Wallace (1975) who have shown that an interest rate peg leads to equilibrium indeterminacy.[18]

Stability at the Effective Lower Bound. Related to the determinacy 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. The intuition is directly related to our discussion about

[18] Angeletos and Lian (2021) show (in a model without household heterogeneity) that small frictions in memory and intertemporal coordination lead to a unique equilibrium which is the same as the one selected by the Taylor principle but it does no longer depend on it.

determinacy under a peg: 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.[19]

We now show that the behavioral HANK model resolves these issues. To this end, let us add a demand shock to the IS equation:

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

Here, $\widehat{r}^n_t$ denotes the demand shock and we interpret it as a natural-rate shock.[20]

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

$$\widehat{y}_t = -\frac{1}{\gamma}\psi_c \underbrace{\left(\widehat{i}_{ELB} - \widetilde{r}^n\right)}_{>0} \sum_{j=0}^{k} \left(\psi_f + \frac{\kappa}{\gamma}\psi_c\right)^j, \tag{25}$$

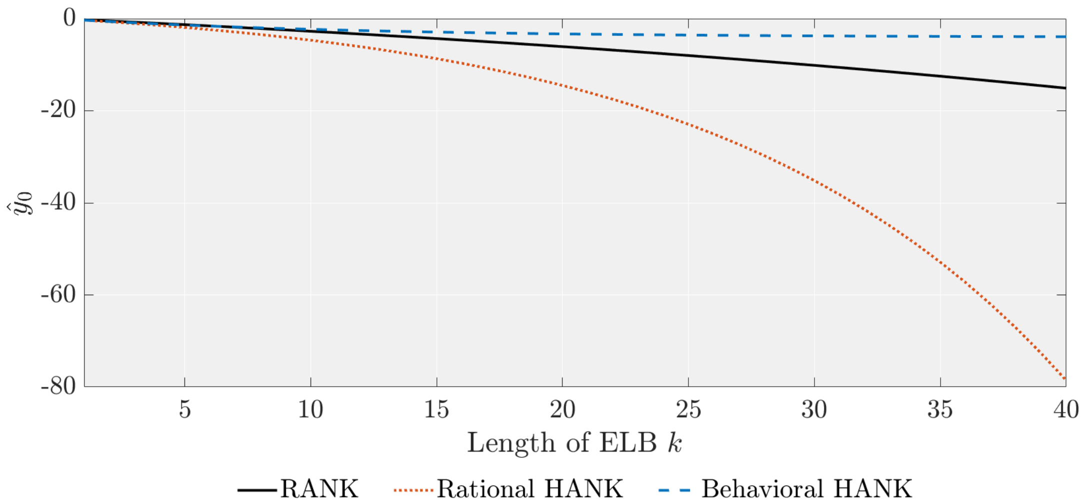
where the term $\left(\widehat{i}_{ELB} - \widetilde{r}^n\right) > 0$ captures the shortfall of the policy response due to the binding ELB. Under rational expectations and countercyclical inequality, $\chi > 1$ and, thus, $\delta > 1$ and $\psi_f > 1$, meaning that output implodes as $k \to \infty$. The same is true in the rational RANK model which is captured by $\chi = 1$ and, thus, $\delta = 1$. 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 when $k \to \infty$. The condition $\psi_f + \frac{\kappa}{\gamma}\psi_c < 1$ is the same as for determinacy under a peg. It follows that $\bar{m} < 0.93$ is enough to rule out unboundedly-severe recessions at the ELB even if the ELB is expected to persist forever.

We illustrate the stability of the behavioral HANK at the lower bound graphically in Figure 2. The figure 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 2 shows the implosion of output in the rational RANK and even more so in the rational HANK

[19]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.

[20]Note, that we write the IS equation in terms of output deviations from its steady state, not from its flexible-price counterpart. The interpretation and analysis in this section, however, would be the same.

Figure 2: 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.

model: an ELB that is expected to bind for 40 quarters would decrease today's output in the rational RANK by 15% and in the rational HANK model by 80%. 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 4%.

3.3 Fiscal Policy

We now show that the sufficient statistic for amplification of contemporaneous monetary policy is also a sufficient statistic to generate positive consumption multipliers of fiscal policy under constant real rates as in line with the data. Dupor et al. (2021) and Galí et al. (2007), for example, provide empirical evidence for positive effects of government spending on private consumption. Furthermore, Nakamura and Steinsson (2014), Ramey (2019) and Chodorow-Reich (2019) document fiscal multipliers above 1, which through the lens of our model is equivalent to saying that consumption responds positively to government spending.

To characterize fiscal multipliers, we follow Bilbiie (2021) and assume government spending g_t to follow an AR(1) with persistence $\rho_g \geq 0$, and to be 0 in steady state. The government taxes all agents uniformly to finance g_t.

We re-derive the behavioral HANK IS equation with government spending and obtain:[21]

$$\widehat{c}_t = \psi_f \mathbb{E}_t \widehat{c}_{t+1} - \psi_c \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t \pi_{t+1}\right) + \zeta\left[\frac{\lambda(\chi-1)}{1-\lambda\chi}\left(g_t - \bar{m}\mathbb{E}_t g_{t+1}\right) + (\psi_f - \bar{m})\mathbb{E}_t g_{t+1}\right],$$

where $\zeta \equiv \frac{\varphi}{\gamma\left(1+\frac{\varphi}{\gamma}\right)}$. The static Phillips Curve in this setting is given by $\pi_t = \kappa c_t + \kappa\zeta g_t$.

The following Proposition characterizes the fiscal multiplier in the behavioral HANK model.

Proposition 4. *The fiscal multiplier in the behavioral HANK model is given by*

$$\frac{\partial \widehat{c}_t}{\partial g_t} = \frac{1}{1-\nu\rho_g}\frac{\zeta}{1+\frac{1}{\gamma}\psi_c\phi\kappa}\left[\frac{\chi-1}{1-\lambda\chi}\left[\lambda(1-\bar{m}\rho_g) + \bar{m}\rho_g(1-s)\right] - \kappa\frac{1}{\gamma}\psi_c\left(\phi - \rho_g\right)\right],$$

where

$$\nu \equiv \frac{\psi_f + \kappa\frac{1}{\gamma}\psi_c}{1+\frac{1}{\gamma}\psi_c\phi\kappa}. \tag{26}$$

A corollary of Propositon 4 is that with persistent government spending, $\rho_g > 0$, and with $\chi > 1$, more bounded rationality, i.e., a lower $\bar{m}$, leads to a lower fiscal multiplier.[22] Bounded rationality decreases the fiscal multiplier as boundedly-rational agents discount the fact that an increase in government spending today has a positive effect on future spending as well. In the case of an i.i.d. spending shock the fiscal multiplier is independent of $\bar{m}$. Furthermore, the fiscal multiplier is bounded from above in the behavioral HANK model as $\nu\rho_g < 1$ even for highly persistent shocks. In the rational model, on the other hand, this is not the case. The fiscal multiplier approaches infinity as $\nu\rho_g \to 1$, which can occur because in the rational HANK model $\nu > 1$. As $\nu\rho_g > 1$ the multiplier even becomes negative. The behavioral HANK model, on the other hand, rules out these undesirable model implications.

To make the argument as clear as possible, we assume prices to be fully rigid, $\kappa = 0$, and assume that the real interest rate is held constant after the government spending shock. This is a useful benchmark as in this case the consumption response in RANK is 0 (see Bilbiie (2011) and Woodford (2011)).[23]

From Proposition 4, we can directly derive the constant-real-rate multiplier in the behavioral HANK model. It shows that with $\chi > 1$ the fiscal multiplier is bounded from below by 0 irrespective of the persistence ρ_g. In other words the constant-real-rate multiplier in the behavioral HANK model is strictly positive, regardless of the dampening of bounded

[21] See Appendix A.3 for the derivation.

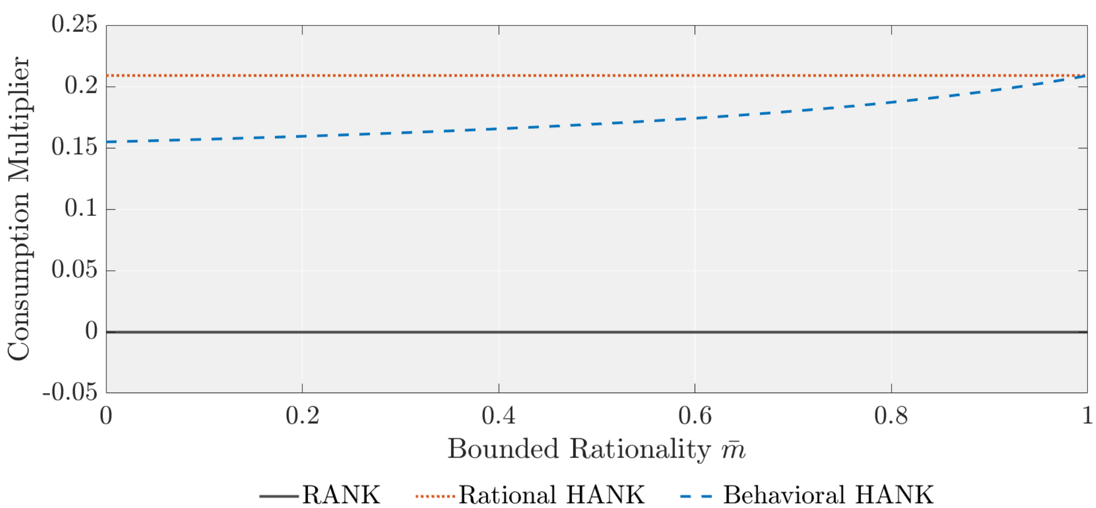
[22] We focus on the case in which $\nu\rho_g < 1$, which holds in the behavioral HANK model even for $\rho_g = 1$, and we assume $1 - s - \lambda < 0$, which holds under all reasonable parameterizations.

[23] Auclert et al. (2018) also use a constant real interest rate case to show that their HANK model can generate (output) fiscal multipliers larger than one.

rationality on the fiscal multiplier in the case of persistent spending. With $\chi > 1$ the high MPC households benefit disproportionately from the extra income out of the increase in government spending which increases the fiscal multiplier through a Keynesian type argument.

Figure 3 illustrates the effect of bounded rationality on the fiscal multiplier by plotting the fiscal multiplier in the behavioral HANK model for varying degrees of $\bar{m}$ (blue-solid line) and comparing it to the multiplier in the rational HANK model and RANK. For this exercise, we set the persistence parameter to an intermediate value $\rho_g = 0.6$. It shows that the fiscal multiplier decreases with decreasing $\bar{m}$. Yet, even for the extreme case $\bar{m} = 0$, in which households fully discount all future increases in government spending the fiscal multiplier is still substantially above zero even though it is somewhat weaker than under rational expectations. In fact, the behavioral HANK model generates consumption responses to fiscal spending that are quantitatively close to the empirical estimates in Dupor et al. (2021) who estimate the non-durable consumption response to lie between 0.2 and 0.29. Note, that we did not target this moment.

Figure 3: Consumption Response to Government Spending



Note: This figure shows the consumption multipliers (the consumption response to government spending) for different degrees of bounded rationality (blue-dashed line). The orange-dotted line plots the multiplier in the rational version of the model and the black-solid line shows the zero-multiplier in the RANK model.

It is noteworthy that the behavioral HANK model does not rely on a specific financing type to achieve positive consumption responses to fiscal spending. This is in contrast to the behavioral RANK model in Gabaix (2020). In the behavioral RANK model, bounded rationality can also increase the multiplier but only if the government delays taxing the agents to finance the contemporaneous spending as boundedly-rational agents will then discount the future increases in taxes. In HANK models, on the other hand, the fiscal multiplier can in principle be larger than one with $\chi < 1$ if the hand-to-mouth households pay relatively less of the fiscal spending's cost than unconstrained households (see Bilbiie (2020) or Ferriere and Navarro (2018)). Both of these channels would also push up the multiplier in the behavioral HANK model, yet it does not depend on any of these two to achieve (output) fiscal multipliers larger than 0.

3.4 Behavioral HANK as a Unifying Framework

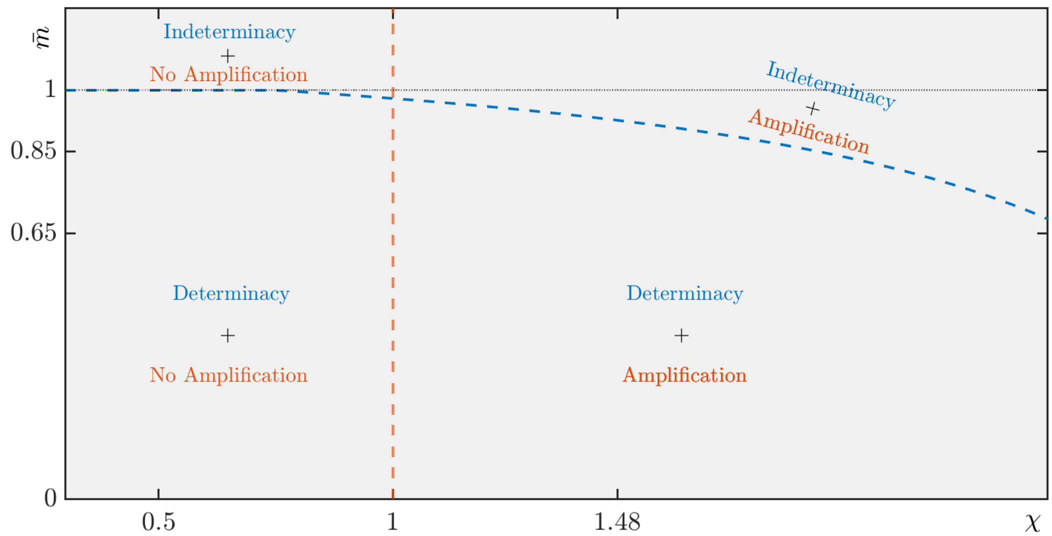
What allows the behavioral HANK model to be reconciled with the empirical facts on the transmission channels and effectiveness of monetary and fiscal policy is indeed the combination of bounded rationality and household heterogeneity. To see this, Figure 4 shows how the two frictions interact. The figure plots the parameter space for the two key parameters for household heterogeneity and bounded rationality, respectively, $(\chi, \bar{m})$. The blue and orange dashed lines split the parameter space in the following sense: the blue line denotes the cut-off values below which the model is determinate under an interest-rate peg while above it the model is indeterminate (with the line itself belonging to the indeterminacy region). Determinacy under a peg is a sufficient statistic to rule out the forward guidance puzzle as well as the instability issues that arise at the lower bound. The orange line denotes the cut-off values such that to the right of it the model generates amplification while left from it—again including the line—the model does not generate amplification. Here, *amplification* is a stand-in for the amplification of monetary and fiscal policies through indirect, general equilibrium, effects.

This split of the parameter space into four areas allows us to distinguish the models discussed so far and to show how the behavioral HANK can overcome the limitations inherent in existing model. The RANK model is located in the "indeterminacy + no amplification" region as $\bar{m} = 1$ and $\chi = 1$. The behavioral RANK can either be in "indeterminacy + no amplification" or in "determinacy + no amplification" depending on the degree of rationality.[24] Rational HANK models can either be in "indeterminacy + no amplification", "determinacy + no amplification" or in "indeterminacy + amplification" while rational TANK models can only be in "indeterminacy + no amplification" or in "indeterminacy + amplification". Importantly, both cannot be in "determinacy + amplification".[25] Only the behavioral HANK model can deliver "determinacy + amplification". Furthermore, the behavioral HANK model

[24] Note, this also applies to other models featuring deviations from FIRE that deliver equivalent reduced-form IS equations, e.g., Angeletos and Lian (2018) and Woodford (2019).

[25] Note that this also applies to the models in McKay et al. (2017), Werning (2015), Ravn and Sterk (2017), Debortoli and Galí (2018), Bilbiie (2020), Bilbiie (2021) and many more.

Figure 4: The Behavioral HANK as a Unifying Framework



Note: The figure characterizes four possible regions depending on whether the considered $(\chi, \bar{m})$-pair delivers determinacy under an interest-rate peg or not and whether the model generates amplification of contemporaneous monetary and fiscal policy or not (we only extend the y-axis above 1 for the sake of readability).

can in principle cover the whole parameter space as it nests all the aforementioned models as special cases.

Having discussed the aggregate implications of the model, we now zoom in closer into the model and derive the iMPCs and show how they depend on bounded rationality, household heterogeneity, and the interaction of the two.

3.5 Intertemporal MPCs

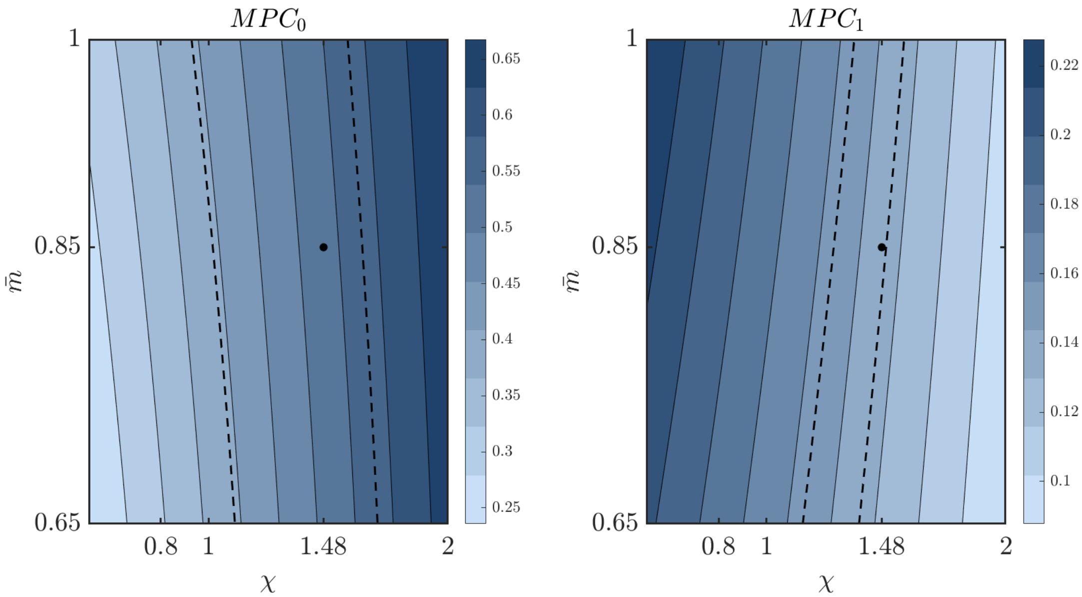
The HANK literature shows that the iMPCs are a key statistic for conducting policy analysis (see, e.g., Auclert et al. (2018), Auclert et al. (2020), and Kaplan and Violante (2020)).[26] We follow the tractable HANK literature and define the aggregate iMPCs in the behavioral HANK model as the partial derivative of aggregate consumption at time k, $\widehat{c}_k$, with respect to aggregate disposable income, $\tilde{y}_0$, keeping everything else fixed (see Bilbiie (2021), Cantore and Freund (2021), and Auclert et al. (2018)).

The following Proposition characterizes the iMPCs in the behavioral HANK model.[27]

Proposition 5. *The intertemporal MPCs in the behavioral HANK model, i.e., the aggregate consumption response in period k to a one-time change in aggregate disposable income in*

[26]See, e.g., Lian (2021) or Boutros (2022) for MPC analyses in models deviating from FIRE.
[27]See Appendix D for the derivation.

Figure 5: Intertemporal MPCs, Bounded Rationality and Household Heterogeneity



Note: This figure shows the aggregate intertemporal MPCs, i.e., the aggregate consumption response in year 0 (left) and year 1 (right) to a change in aggregate disposable income in year 0 for different χ (x-axis) and $\bar{m}$ (y-axis). The dashed lines show the range of empirically-estimated iMPCs and the black dot shows the model estimate given our baseline calibration. Darker colors represent higher MPCs, see the colorbars on the right side of the figures.

period 0, are given by

$$MPC_0 \equiv \frac{d\widehat{c}_0}{d\widetilde{y}_0} = 1 - \frac{1-\lambda\chi}{s\bar{m}}\mu_2^{-1}$$
$$MPC_k \equiv \frac{d\widehat{c}_k}{d\widetilde{y}_0} = \frac{1-\lambda\chi}{s\bar{m}}\mu_2^{-1}\left(\beta^{-1}-\mu_1\right)\mu_1^{k-1}, \qquad \textit{for } k>0,$$

where the parameters μ_1 and μ_2 depend on the underlying parameters, including $\bar{m}$ and χ and are explicitly spelled out in Appendix D.

Figure 5 graphically depicts how the interplay of bounded rationality $\bar{m}$ and household heterogeneity χ determines the size of the aggregate iMPCs. We calibrate the model annually as the empirical evidence on the iMPCs is annual (see Fagereng et al. (2021) and Auclert et al. (2018)). We set $s = 0.8$ and $\beta = 0.95$, and keep the rest of the calibration as above. The left panel depicts the aggregate MPCs within the first year (in period 0) and the right panel the aggregate MPCs within the second year (in period 1). Darker colors represent higher MPCs. First, note that with our baseline calibration—$\chi = 1.48$ and $\bar{m} = 0.85$ as shown by the black dots—the behavioral HANK model generates iMPCs within the first year of 0.55 and within the second year of 0.15. These values lie within the estimated bounds for the iMPCs in the data (Auclert et al. (2018)) which are between $0.42 - 0.6$ for the first and

$0.14 - 0.16$ for the second year (see dashed lines). Away from our baseline calibration, an increase in χ increases the MPCs in the first year but decreases them in the second year.[28] In contrast, an increase in $\bar{m}$ increases the aggregate MPC in the first year and in the second year.

Let us first turn to the role of χ for the iMPCs: Recall, the higher χ the more sensitive is the income of the H households to a change in aggregate income. Thus, with higher χ, H households gain weight in relative terms for the aggregate iMPCs while unconstrained households loose weight in relative terms. This pushes up the aggregate MPC within the first year as the H households spend all of their income windfall, but pushes down the aggregate MPC within the second year as households that were hand-to-mouth in the period of the income windfall have a MPC of 0 in the second year.

Bounded rationality, captured by $\bar{m}$, affects only the MPCs of unconstrained households as these are the only households who intertemporarily optimize. Their Euler equation dictates that the decrease in today's marginal utility of consumption—following the increase in consumption—is equalized by a decrease in tomorrow's expected marginal utility. For behavioral households, however, the decrease in tomorrow's marginal utility needs to be more substantial as they cognitively discount the future decrease. Hence, behavioral households save relatively more out of the income windfall. This pushes down the aggregate MPCs in $t = 0$. The same is true for the aggregate MPC in $t = 1$, in which there are now two opposing forces at work: on the one hand, unconstrained households again cognitively discount the expectations about the future decrease in her marginal utility which depresses their consumption. On the other hand, unconstrained households have accumulated more wealth from period $t = 0$ which tends to increase consumption. Given our calibration, in $t = 1$ the former dominates. Figure 11 in Appendix D shows that, beginning in $k = 3$, the latter effect starts to dominate. For a higher idiosyncratic risk of becoming hand-to-mouth, i.e., an increase in the transition probability $1 - s$, the aggregate MPC is already higher in $t = 1$ for lower $\bar{m}$. The reason is that a smaller fraction of initial unconstrained households remains unconstrained which pushes upwards consumption in $k = 1$ (see Figure 12 in Appendix D).

The effects of a change in $\bar{m}$ are more pronounced at lower levels of χ. Combining our discussion about the role of χ and $\bar{m}$, this is intuitive: the lower χ, the higher is the relative importance of unconstrained households for the aggregate iMPCs and, in turn, the stronger is the effect of $\bar{m}$ on the aggregate iMPCs. These interaction effects are quite substantial:

[28] Note, that when considering micro moments like the iMPCs, $\chi = 1$ is not sufficient anymore for the model to collapse to RANK. More precisely, with $\chi = 1$ the model collapses to a HANK model which behaves in the aggregate exactly like RANK (see the incomplete-markets irrelevance result in Werning (2015)). Hence, the RANK iMPCs cannot directly be seen in Figure 5. Yet, Proposition 5 still nests RANK for $\chi = 1$ and $\lambda = 0$.

at $\chi = 1.48$, a decrease of $\bar{m}$ from 1 to 0.65 decreases the MPC_0 by 7% and the MPC_1 by more than 11%.

3.6 Policy Implications: The Timing of Monetary Policy

What are the policy implications from the behavioral HANK model? In this section, we examine how the timing of monetary policy affects its effectiveness and its distributional consequences and show that the behavioral HANK model differs *qualitatively* from its rational counterpart.

Assume the central bank wants to increase the nominal interest rate by a cumulative $x\%$, for example, to counteract a negative supply shock. The central bank decides whether to implement this policy within a single quarter or to gradually raise the interest rate by $\frac{x}{k}\%$ for k consecutive quarters.

Lemma 1. *The effect of a $\frac{x}{k}\%$ interest rate hike over k consecutive periods decreases current output by*

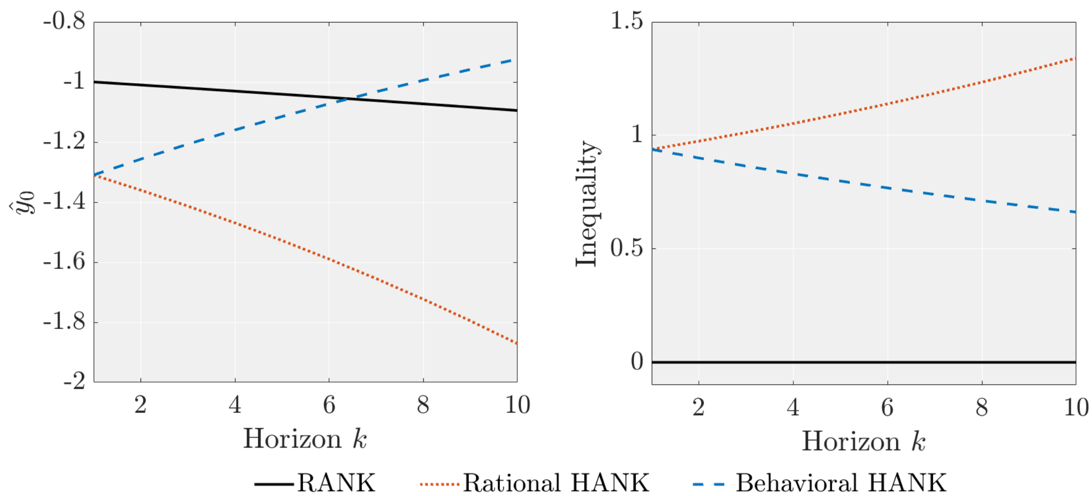
$$\widehat{y}_t = \frac{\psi_c}{\gamma} \left[\sum_{j=0}^{k-1} \left(\psi_f + \frac{\psi_c}{\gamma} \kappa \right)^j \right] \frac{x}{k}.$$

The left panel of Figure 6 depicts the result in Lemma 1 for the behavioral HANK model and compares it to its rational counterpart and the rational RANK model. The figure shows the response of current output on the vertical axis to the policy change stretched over different number of periods (horizontal axis). The solid-black line shows the well-known feature of RANK that the effects of monetary policy on current output become stronger when monetary policy is back-loaded: the further the interest hike is stretched out the higher is the response on current output. The orange-dotted line shows that this feature is even more pronounced in the rational HANK model as the line is steeper than in RANK.

In contrast, the blue-dashed line representing the behavioral HANK model is increasing instead of decreasing in k. Thus, back-loading monetary policy decreases its effect on current output. To put it differently, monetary policy is most effective if it is completely front-loaded. Hence, if the central bank wants to fight an overheating of the economy as effectively as possible, the behavioral HANK model implies front-loading the interest rate hike, while its rational counterpart suggests to back-load the hike.

The right panel of Figure 6 depicts the effects of the different timing of the monetary policy hikes on consumption inequality, as defined in equation (14). It shows that, according to the behavioral HANK, if monetary policy front-loads the interest rake hike, it increases inequality the most whereas a more gradual increase in the interest rate would have weaker effects on inequality. This illustrates a trade-off for the central banker: the more effectively monetary policy combats the overheating, the more it increases inequality.

Figure 6: Monetary Policy Timing: Effectiveness and Distributional Consequences



Note: This figure shows the response of current output (left panel) of a cumulative interest-rate hike by $x\%$ implemented over k consecutive periods. The right panel shows the corresponding response of inequality, defined as $\widehat{c}_t^U - \widehat{c}_t^H$.

3.7 A Quantitative Behavioral HANK Model

In this section, we develop a quantitative behavioral heterogeneous agent New Keynesian model and show that the main insights of our tractable three-equation model carry over to this more quantitative version of our model.

Our quantitative model employs the standard HANK set-up in which households are ex-ante identical but face an idiosyncratic productivity risk. Households self-insure against this idiosyncratic risk by accumulating bonds issued by the government. The fiscal authority issues a constant amount of real debt, B^G, and collects tax payments from households to finance its interest rate payments. Hence, ex-post households differ in their current productivity level, e, and their wealth B. Households' utility function is the same as considered in the tractable model (equation (1)).

The budget constraint of household i is given by

$$C_{i,t} + \frac{B_{i,t+1}}{1+r_t} = B_{i,t} + W_t e_{i,t} N_{i,t} + D_t d(e) - \tau_t(e)$$
$$B_{i,t+1} \geq \underline{B},$$

where the second equation denotes the household's (exogenous) borrowing constraint. As in McKay et al. (2016), we assume that households pay taxes conditional on their productivity level, $\tau_t(e)$. In addition, we assume that households receive a share of the dividends, $D_t d(e)$ conditional on their productivity level. Similar to the setup in the tractable model, we assume that the high productivity households receive a larger share of the dividends than low-productivity households. As dividends are countercyclical in the model, this assumption makes sure that households with higher MPCs (which are highly correlated with the low-productivity state) tend to be more exposed to the business cycle, in line with the tractable model and the empirical evidence (Auclert (2019), Patterson (2019)). This is different from the model in McKay et al. (2016) who assume that every household receives the same share of the dividends which leads to procyclical inequality.[29]

We introduce bounded rationality in the same way as in our tractable model. Households are fully rational with respect to their idiosyncratic risk, but they cognitively discount the expected deviations of future aggregates (including prices) from their respective values in the stationary equilibrium. As a household's individual consumption depends on these aggregates, she cognitively discounts expected future *deviations* of her marginal utility in each state from its stationary equilibrium counterpart. Households have perfect foresight about the path of the real interest rate.

The Euler equation of a household with current state $i = (B, e)$ is given by

$$C_{i,t}^{-\gamma} \geq \beta R_t \mathbb{E}_t^{BR} \left[C_{i',t+1}^{-\gamma} \right], \tag{27}$$

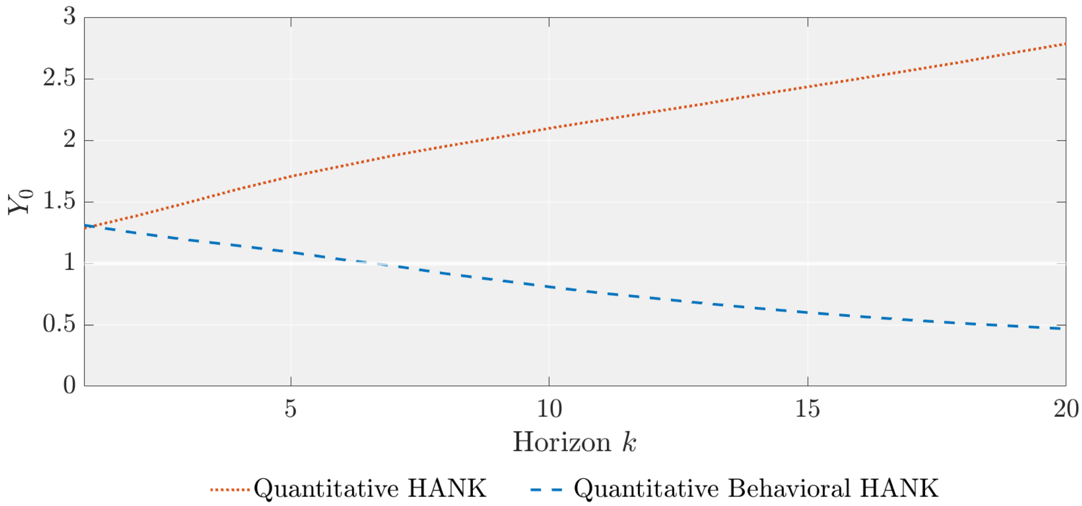
which holds with equality for non-constrained households, while it holds with strict inequality for households that are pushed to their borrowing constraint. The labor-leisure condition is identical to the one in the tractable model and holds for every household. In the case of rational expectations, the model collapses to a standard one-asset HANK model, similar to McKay et al. (2016), Hagedorn et al. (2019), or Debortoli and Galí (2018). We relegate further details of the model and the parameterization to Appendix F.

Monetary Policy Experiment. Let us now consider the following experiment. The monetary authority announces in period 0 to decrease the nominal interest rate by 10 basis points in period k and keeps the nominal rate at its steady state value in all other periods. Following Farhi and Werning (2019), we focus on the case with fully rigid prices and thus, the change in the nominal rate translates one for one to changes in the real rate and is thus also consistent with the exercise in McKay et al. (2016). What is the effect of such an

[29] We show how our quantitative behavioral HANK nests McKay et al. (2016) in Appendix F.

interest-rate change on total output in period 0?

Figure 7: Monetary Policy in the Quantitative Model



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, relative to the response in the RANK model under rational expectations (equal to 1).

Figure 7 provides the answer. It shows on the vertical axis the response of output in period 0, Y_0, to an announced real rate change implemented in period k (horizontal axis). The white, horizontal line at 1 represents the response in the complete-markets model, i.e., in the rational RANK model.[30] The constant response in RANK is a consequence of the assumption that forward guidance is implemented through changes in the real rate.

The orange-dotted line shows the results for the rational HANK model. We see how the fact that households with higher MPCs tend to be more exposed to the business cycle leads to an increase in the effectiveness of contemporaneous monetary policy. This amplification through indirect effects, however, extends intertemporally and results in an aggravation of the forward guidance puzzle. Indeed, we see that the further away the announced interest rate change takes place, the stronger the response of output today. A change that is announced to take place in twenty quarters leads to a response of today's output that is almost three times as strong as in RANK.

The behavioral HANK model, on the other hand, does not suffer from the forward guidance puzzle, as shown by the blue-dashed line. While the effect of a contemporaneous interest rate change is almost identical as in the rational HANK model, interest rate changes announced to take place in the future have relatively weak effects on contemporaneous output

[30] Note that for an easier interpretation, we normalized the y-axis by dividing through the response in the rational RANK model which is 0.05% after a shock of 10 basis points.

and the effects decrease with the horizon. In Appendix F, we discuss how our resolution of the forward guidance puzzle contrasts with other resolutions in the HANK literature.

Overall, Figure 7 illustrates that the main insights of the tractable behavioral HANK model carry over to more quantitative models. Contemporaneous monetary policy is amplified through indirect, general equilibrium, channels whereas announced future policies have relatively weaker effects on today's economy.

4 Extensions

We now extend our baseline tractable model along two dimensions. First, we allow for sticky wages and show how the interplay of sticky wages, household heterogeneity and bounded rationality leads to hump-shaped responses of macroeconomic variables in response to aggregate shocks, as well as forecast-error dynamics consistent with recent findings from survey data. Second, we derive an equivalence result between HANK models with bounded rationality and HANK models with incomplete information and learning.

4.1 Sticky Wages

Recent HANK models have relaxed the assumption of fully-flexible wages and rather assume wages to be sticky, bringing these models closer to some dimensions in the data (see, e.g., Auclert et al. (2020) or Broer et al. (2020)).

Modelling sticky wages. To introduce sticky wages, we follow Colciago (2011) and assume a centralized labor market in which a labor union allocates the hours of households to firms and makes sure that U and H households work the same amount. The labor union faces the typical Calvo (1983) constraint, such that it can re-optimize the wage within a given period only with a certain probability, giving rise to a wage Phillips Curve. We assume that the labor union sets wages based on rational expectations to focus on the effects of bounded rationality solely on the household side.

The wage Phillips curve is given by

$$\pi_t^w = \beta \mathbb{E}_t \pi_{t+1}^w + \kappa_w \widehat{\mu}_t^w,$$

where π_t^w denotes wage inflation, κ_w the slope of the wage Phillips curve and $\widehat{\mu}_t^w$ is a time-varying wage markup, given by

$$\widehat{\mu}_t^w = \gamma \widehat{c}_t + \varphi \widehat{n}_t - \widehat{w}_t.$$

We set $\kappa_w = 0.075$ as in Bilbiie et al. (2021).

We follow Auclert et al. (2020) and introduce interest-rate smoothing in the Taylor rule:

$$\widehat{i}_t = \rho_i \widehat{i}_{t-1} + (1-\rho_i)\phi\pi_t + \varepsilon_t^{MP}$$

and we set $\rho_i = 0.89$ and $\phi = 1.5$ as estimated by Auclert et al. (2020) and assume the shocks ε_t^{MP} to be completely transitory. Similar to the wage setters, we assume price-setting firm managers to be fully rational, giving rise to the standard New Keynesian Phillips curve

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa_\pi \widehat{mc}_t,$$

where $\widehat{mc}_t$ denotes the time-varying price markup. The rest of the model is as above. We relegate the details and the parameterization to Appendix E.

Hump-shaped responses to monetary policy shock. Figure 8 shows the impulse-response functions of output, real wages and consumption of the two household types to a monetary policy shock for the behavioral HANK model (blue-dashed lines), the behavioral RANK model (orange-dashed-dotted lines), and the rational HANK model (black-solid lines).

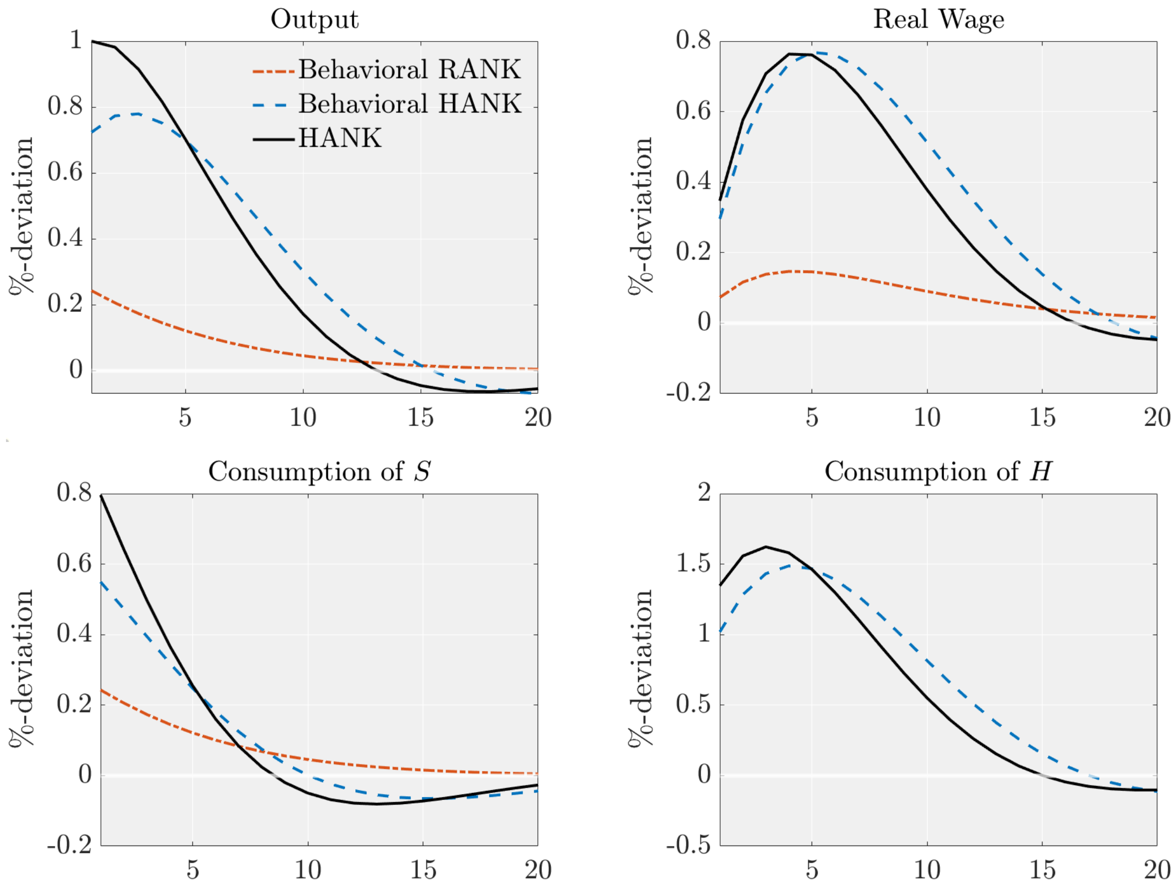
The output response to a monetary policy shock is hump-shaped in the behavioral HANK model but neither in its representative agent nor in its rational counterpart. The hump-shaped response is quite remarkable as the model does not feature backward-looking expectations or any form of consumption habits, investment or investment adjustment costs. So, where do these hump-shaped responses come from in the behavioral HANK model?

First, note that the introduction of wage rigidity leads to a hump-shaped response in real wages, which is the case in all three models. Since wages determine the H households' income in the rational and the behavioral HANK, their consumption also follows a hump-shape (see lower right figure).[31] Crucial for the overall response, however, is not only the response of H households but also the response of unconstrained households.

Under rational expectations, unconstrained households perfectly understand how the consumption of H agents responds and what this implies for their idiosyncratic risk induced by type switching. In particular, they understand already on impact that their self-insurance motive will be relaxed for some periods. Thus, unconstrained households immediately cut back on precautionary savings and, thus, their consumption responds strongly on impact. Under bounded rationality, however, unconstrained households cognitively discount the future and thus, underreact to the expected increase in wages and, thus, the relaxation of

[31] We show the H households' consumption response only for the rational and behavioral HANK model, as the representative agent model does not feature H agents.

Figure 8: Monetary Policy Shock



Note: This figure shows the impulse-response functions of output, real wages and consumption of the two household types to a monetary policy shock in the behavioral HANK model for different $\bar{m}$ and the rational HANK model with sticky wages. The shock size is normalized such that output in the rational model increases by 1pp on impact.

their idiosyncratic risk. Hence, on impact, they do not cut back on precautionary savings as strong as a rational household would. Going forward, they learn that their self-insurance motive is still (or even more) relaxed. As a consequence, their consumption decreases slower inducing a flatter consumption profile compared to a rational unconstrained household. It is the combination of the flatter consumption profile of unconstrained households and the hump-shaped consumption profile of the hand-to-mouth that generates the hump-shaped response of consumption in the aggregate.

The model with a representative (behavioral) agent does not generate the hump-shaped response. The reason is that without constrained agents, the wage profile does not translate into hump-shaped consumption of (a sub population of) households to begin with. It is thus indeed the *interaction* of household heterogeneity and bounded rationality that produces these hump-shaped responses.

Auclert et al. (2020) argue that many macroeconomic models fail to generate the *micro jumps and macro humps* that we observe in the data, i.e., iMPCs that respond strongly on impact and hump-shaped responses of macroeconomic variables to aggregate shocks. Our results on iMPCs in Section 3.5 as well as the results presented in Figure 8 show how the behavioral HANK model offers a tractable analogue to the full-blown HANK model presented in Auclert et al. (2020).[32]

Forecast-errors dynamics. We now show that the sticky-wage behavioral HANK model generates dynamic forecast errors as observed in survey data. We focus on the dynamics of the forecast errors from the behavioral agents' perspective after a monetary policy shock and focus on one-period ahead forecast errors and how they evolve over time. For a variable $\widehat{x}$, the one-period ahead forecast error is defined as

$$FE^{\widehat{x}}_{t+h+1|t+h} \equiv \widehat{x}_{t+h+1} - \bar{m}\mathbb{E}_{t+h}\left[\widehat{x}_{t+h+1}\right].$$

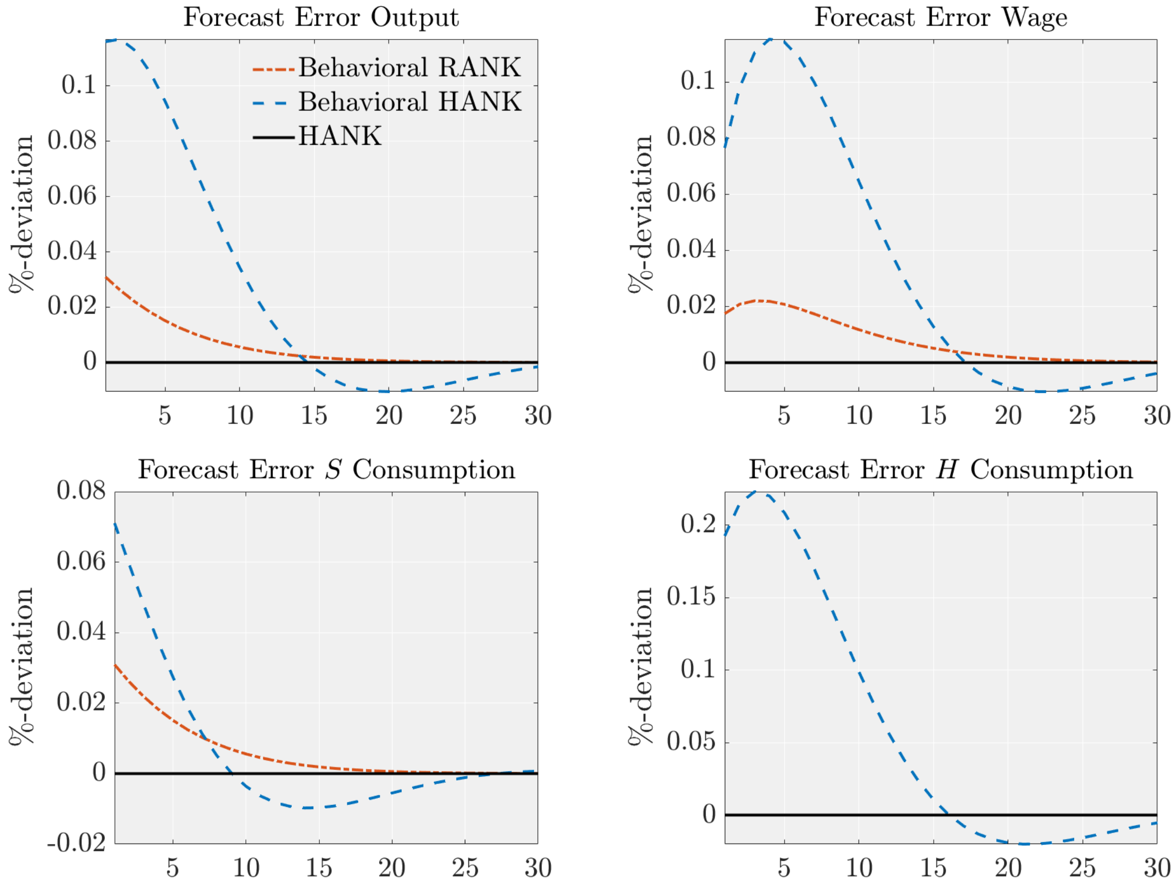
A positive forecast error thus means that the agent's forecast was lower than the actual outcome.

Figure 9 shows the forecast errors of output, the real wage and consumption of the two household types, starting in the first period after the shock. For completeness, the black line at zero shows that under rational expectations, i.e., $\bar{m} = 1$, forecast errors are equal to 0. In the behavioral model, however, this is not the case. In fact, forecast errors are positive in the first few quarters after the shock, illustrating the underreaction of the agents' expectations to the shock. This underreaction is what drives the hump-shaped response to a large degree and what is responsible for the dynamics of the consumption response of unconstrained households discussed in Figure 8.

After about 10-15 quarters, however, forecast errors turn negative. Put differently, the behavioral agents' expectations show patterns of delayed overreaction. These dynamic patterns of initial underreaction followed by a delayed overreaction has recently been documented empirically in Angeletos et al. (2021) for unemployment and inflation and in Adam et al. (2022) for housing prices. In fact, Angeletos et al. (2021) argue that looking at the dynamics of forecast errors in response to structural shocks is more informative than other tests of FIRE. The dynamic responses reconcile seemingly conflicting evidence on underreaction (as in Coibion and Gorodnichenko (2015)) and overreaction (as in Adam et al. (2017) or Kohlhas and Walther (2021)). In contrast to Angeletos et al. (2021) or Adam et al.

[32]Another way to generate hump-shaped responses of output to monetary policy shocks in the behavioral HANK model is to keep wages fully flexible and to allow for persistence in the monetary policy shocks. In this way, the iMPCs presented in Figure 5 are completely unaltered.

Figure 9: Forecast Error Dynamics



Note: This figure shows the forecast error dynamics of output, the real wage, consumption of unconstrained households and of hand-to-mouth households after an expansionary monetary policy shock.

(2022), the behavioral HANK model with sticky wages generates these dynamic patterns of forecast errors even though the behavioral agents' expectations are purely forward looking. Note that the behavioral RANK model as in Gabaix (2020) does not generate these delayed overreactions. Hence, the interaction of household heterogeneity, cognitive discounting and sticky wages cannot only generate hump-shaped responses of macroeconomic aggregates but also forecast error dynamics that are consistent with recent evidence from household survey expectations.

4.2 Bounded Rationality and Incomplete Information with Learning: An Equivalence Result

In this section, we derive an equivalence result of heterogeneous-household models featuring bounded rationality and those featuring incomplete information with learning. In particular, we show how a change in the default value in the behavioral setup leads to an observationally equivalent IS equation as in models with incomplete information and learning (see Angeletos and Huo (2021) and Gallegos (2021)).

To this end, we now assume that behavioral agents anchor their expectations to their *last observation* instead of the steady state values. Models featuring some form of backward-looking behavior indeed tend to match the expectations data coming from household surveys quite well (see, for example, Adam et al. (2017), Adam et al. (2022), Angeletos and Huo (2021), and Angeletos et al. (2021)). The backward-looking components in these models usually arise from an incomplete or noisy information setting as well as some form of (Bayesian) learning. We now show how our bounded rationality setup generates expectations that resemble these aforementioned expectations models.

Proposition 6. *Set the boundedly-rational agents' default value to the variable's past value*

$$X_t^d = X_{t-1}. \tag{28}$$

In this case, the boundedly-rational agent's expectations of X_{t+1} becomes

$$\mathbb{E}_t^{BR}\left[X_{t+1}\right] = (1-\bar{m})X_{t-1} + \bar{m}\mathbb{E}_t\left[X_{t+1}\right]. \tag{29}$$

These backward-looking expectations introduce a backward-looking component into the behavioral IS equation as shown in the following Proposition.

Proposition 7. *In case the behavioral agents' default value is the past value of the respective variable, i.e., $X_t^d = X_{t-1}$, the behavioral HANK IS equation is given by*

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t \pi_{t+1}\right) + (1-\bar{m})\delta\widehat{y}_{t-1}. \tag{30}$$

Proposition 7 shows that the change in the agents' default value does not change the existing behavioral and heterogeneity coefficients ψ_f and ψ_c. Yet, anchoring to past realizations introduces an additional backward-looking term in the IS equation, similar to models relying on habit persistence.[33] The IS equation thus features *myopia* and *anchoring* as in Angeletos and Huo (2021) and Gallegos (2021) who derive an IS equation with the same reduced form. Their setup, however, is based on incomplete-information and learning. We

[33]In Appendix G, we discuss how we can calibrate the model to match recent evidence from survey expectations and that the backward-looking model features determinacy under an interest-rate peg and delivers hump-shaped responses of macroeconomic aggregates to monetary shocks through a behavioral channel.

complement their findings by showing how we can generate the equivalent outcome based on a *behavioral* relaxation of FIRE.

5 Conclusion

We develop a framework that accounts for recent empirical facts on the transmission channels and effectiveness of monetary and fiscal policy. To arrive at this framework, we introduce bounded rationality in the form of cognitive discounting and household heterogeneity into a sticky price model. We show that the interaction of both frictions—household heterogeneity and bounded rationality—enables the model to be reconciled with the data. The presence of both frictions is thus crucial to arrive at our results. The resulting behavioral HANK model is analytically tractable and nests a wide range of existing models—none of which can account for all the empirical patterns. What is more, we show that the behavioral HANK model can have different policy implications than its rational counterpart, e.g., when it comes to the timing of monetary policy. The main insights carry over to a quantitative version of our behavioral HANK model. Extending the model by allowing for sticky wages generates hump-shaped responses of macroeconomic aggregates to monetary policy shocks and delivers forecast error dynamics that are consistent with recent survey evidence. We also show how our framework can be used to arrive at an equivalence result of models featuring bounded rationality and models of incomplete information and learning. Altogether, the behavioral HANK model offers a tractable framework to study a broad array of questions in future research.

References

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

Adam, K., A. Marcet, and J. Beutel (2017): "Stock price booms and expected capital gains," *American Economic Review*, 107, 2352–2408.

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

Almgren, M., J. E. Gallegos, J. Kramer, and R. Lima (2019): "Monetary policy and liquidity constraints: Evidence from the euro area," *Available at SSRN 3422687*.

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. and Z. Huo (2021): "Myopia and anchoring," *American Economic Review*, 111, 1166–1200.

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 (2017): "Dampening general equilibrium: From micro to macro," Tech. rep., National Bureau of Economic Research.

——— (2018): "Forward guidance without common knowledge," *American Economic Review*, 108, 2477–2512.

——— (2021): "Determinacy without the Taylor Principle," Tech. rep., National Bureau of Economic Research.

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

Auclert, A., M. Rognlie, and L. Straub (2018): "The intertemporal keynesian cross," Tech. rep., National Bureau of Economic Research.

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

Benchimol, J. and L. Bounader (2019): *Optimal monetary policy under bounded rationality*, International Monetary Fund.

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

——— (2011): "Nonseparable preferences, frisch labor supply, and the consumption multiplier of government spending: One solution to a fiscal policy puzzle," *Journal of Money, Credit and Banking*, 43, 221–251.

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

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

Bilbiie, F. O., D. Känzig, and P. Surico (2021): "Capital and income inequality: An aggregate-demand complementarity," .

Bonciani, D. and J. Oh (2021): "Optimal monetary policy mix at the zero lower bound," .

——— (2022): "Unemployment risk, liquidity traps and monetary policy," .

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

Boutros, M. (2022): "Windfall Income Shocks with Finite Planning Horizons," .

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 (2021a): "Information and Wealth Heterogeneity in the Macroeconomy," .

Broer, T., P. Krusell, and E. Öberg (2021b): "Fiscal Multipliers: A Heterogenous-Agent Perspective," Tech. rep., National Bureau of Economic Research.

Calvo, G. A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of monetary Economics*, 12, 383–398.

Cantore, C. and L. B. Freund (2021): "Workers, capitalists, and the government: fiscal policy and income (re) distribution," *Journal of Monetary Economics*, 119, 58–74.

Chodorow-Reich, G. (2019): "Geographic cross-sectional fiscal spending multipliers: What have we learned?" *American Economic Journal: Economic Policy*, 11, 1–34.

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

Colciago, A. (2011): "Rule-of-thumb consumers meet sticky wages," *Journal of money, credit and banking*, 43, 325–353.

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

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," .

Dupor, B., M. Karabarbounis, M. Kudlyak, and M. Mehkari (2021): "Regional consumption responses and the aggregate fiscal multiplier," .

D'Acunto, F., D. Hoang, and M. Weber (2020): "Managing households' expectations with unconventional policies," Tech. rep., National Bureau of Economic Research.

Erceg, C. J., D. W. Henderson, and A. T. Levin (2000): "Optimal monetary policy with staggered wage and price contracts," *Journal of monetary Economics*, 46, 281–313.

Fagereng, A., M. B. Holm, and G. J. Natvik (2021): "MPC heterogeneity and household balance sheets," *American Economic Journal: Macroeconomics*, 13, 1–54.

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

Ferriere, A. and G. Navarro (2018): "The Heterogeneous Effects of Government Spending: It's All About Taxes," *FRB International Finance Discussion Paper*.

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.

——— (2016): "Behavioral macroeconomics via sparse dynamic programming," Tech. rep., National Bureau of Economic Research.

——— (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.

Galí, J., J. D. López-Salido, and J. Vallés (2007): "Understanding the effects of government spending on consumption," *Journal of the european economic association*, 5, 227–270.

Gallegos, J.-E. (2021): "HANK beyond FIRE," .

Giannoni, M., C. Patterson, M. Del Negro, et al. (2015): "The forward guidance puzzle," in *2015 Meeting Papers*, Society for Economic Dynamics, 1529.

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

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

Jappelli, T. and L. Pistaferri (2020): "Reported MPC and unobserved heterogeneity," *American Economic Journal: Economic Policy*, 12, 275–97.

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

Kaplan, G., B. Moll, and G. L. Violante (2016): "A note on unconventional monetary policy in HANK," *unpublished paper, University of Chicago*.

Kaplan, G. and G. L. Violante (2020): "The Marginal Propensity to Consume in Heterogeneous Agents Models," Tech. rep., Working Paper. Princeton University.

Kohlhas, A. N. and A. Walther (2021): "Asymmetric attention," *American Economic Review*, 111, 2879–2925.

Krusell, P., T. Mukoyama, and A. A. Smith Jr (2011): "Asset prices in a Huggett economy," *Journal of Economic Theory*, 146, 812–844.

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," Tech. rep., National Bureau of Economic Research.

Lian, C. (2021): "Mistakes in future consumption, high MPCs now," Tech. rep., National Bureau of Economic Research.

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.

Miescu, M. S. (2022): "Forward Guidance Shocks," .

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.

Nakamura, E. and J. Steinsson (2014): "Fiscal stimulus in a monetary union: Evidence from US regions," *American Economic Review*, 104, 753–92.

Patterson, C. (2019): "The matching multiplier and the amplification of recessions," *Unpublished Manuscript, Northwestern University.*

Perotti, R. (2007): "In search of the transmission mechanism of fiscal policy [with comments and discussion]," *NBER macroeconomics Annual*, 22, 169–249.

Pfäuti, O. (2021): "Inflation–who cares? Monetary Policy in Times of Low Attention," *arXiv preprint arXiv:2105.05297.*

Ramey, V. A. (2019): "Ten years after the financial crisis: What have we learned from the renaissance in fiscal research?" *Journal of Economic Perspectives*, 33, 89–114.

Ravn, M. O. and V. Sterk (2017): "Job uncertainty and deep recessions," *Journal of Monetary Economics*, 90, 125–141.

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

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

Sargent, T. J. and N. Wallace (1975): "" Rational" expectations, the optimal monetary instrument, and the optimal money supply rule," *Journal of political economy*, 83, 241–254.

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.

Werning, I. (2015): "Incomplete markets and aggregate demand," 2015 Meeting Papers 932, Society for Economic Dynamics.

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," Tech. rep., Working paper, University of Chicago.

Woodford, M. (2003): "Interest and prices," .

——— (2011): "Simple analytics of the government expenditure multiplier," *American Economic Journal: Macroeconomics*, 3, 1–35.

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

A Model Details and Derivations

A.1 Derivation of χ

In Section 2, we stated that

$$\widehat{c}_t^H = \chi \widehat{y}_t, \tag{31}$$

where $\chi \equiv 1 + \varphi\left(1 - \frac{\tau^D}{\lambda}\right)$ is *the* crucial statistic coming from the household heterogeneity friction. We now show how we arrive at equation (31) from the H-households' budget constraint, optimality conditions and market clearing.

The labor-leisure condition ofthe H households is given by

$$(N_t^H)^{\varphi} = W_t (C_t^H)^{-\gamma}, \tag{32}$$

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.

Thus, log-linearizing the labor-leisure conditions yields

$$\begin{aligned} \varphi \widehat{n}_t^H &= \widehat{w}_t - \gamma \widehat{c}_t^H \\ \varphi \widehat{n}_t^U &= \widehat{w}_t - \gamma \widehat{c}_t^U. \end{aligned}$$

Since both households work for the same wage, we obtain

$$\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H = \varphi \widehat{n}_t^U + \gamma \widehat{c}_t^U \tag{33}$$

Log-linearizing the market clearing conditions yields

$$\begin{aligned} \widehat{n}_t &= \lambda \widehat{n}_t^H + (1-\lambda)\widehat{n}_t^U \\ \widehat{c}_t &= \lambda \widehat{c}_t^H + (1-\lambda)\widehat{c}_t^U, \end{aligned}$$

which can be re-arranged to (using $\widehat{y}_t = \widehat{c}_t = \widehat{n}_t$)

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

Replacing $\widehat{n}_t^U$ and $\widehat{c}_t^U$ in equation (33) then gives

$$\varphi\widehat{n}_t^H + \gamma\widehat{c}_t^H = (\varphi + \gamma)\widehat{y}_t. \tag{34}$$

The budget constraint of H households (accounting for the fact that bond holdings are zero in equilibrium) is given by

$$C_t^H = W_t N_t^H + \frac{\tau^D}{\lambda} D_t, \tag{35}$$

where we replaced T_t^H with $\frac{\tau^D}{\lambda} D_t$. In log-linearized terms, we get

$$\widehat{c}_t^H = \widehat{w}_t + \widehat{n}_t^H + \frac{\tau^D}{\lambda}\widehat{d}_t, \tag{36}$$

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

$$\widehat{c}_t^H = \left(\varphi\widehat{n}_t^H + \gamma\widehat{c}_t^H\right)\left(1 - \frac{\tau^D}{\lambda}\right) + \widehat{n}_t^H. \tag{37}$$

Using (34) to solve for $\widehat{n}_t^H$ and plugging it into (37) yields

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

Grouping terms, we obtain

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

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

A.2 Derivation of Proposition 1.

Combining equations (11) and (13) with the bounded-rationality setup in equation (18) for $\widehat{x}_t^d = 0$ as X_t^d is given by the steady state, we have

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

Plugging these two equations as well as equation (13) into the Euler equation of unconstrained households (15) yields

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

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 $\mathbb{E}_t\left[\widehat{y}_{t+1}\right]$:

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

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 IS Curve with Government Spending

Since the government spending is financed by uniform taxes, $\tau_t^H = \tau_t^U = G_t$, household h's net income is:

$$\widehat{c}_t^H = \widehat{w}_t + \widehat{n}_t^H + \frac{\tau^D}{\lambda}\widehat{d}_t - g_t, \tag{38}$$

where $g_t = log(G_t/Y)$.

We first derive households h consumption as a function of total income $\widehat{y}_t$. The good markets clearing condition is now

$$\widehat{y}_t = \lambda\widehat{c}_t^H + (1-\lambda)\widehat{c}_t^U + g_t. \tag{39}$$

Plugging this and the labor market clearing condition into (33), yields:

$$\varphi\widehat{n}_t^H + \gamma\widehat{c}_t^H = (\varphi+\gamma)\widehat{y}_t - \gamma g_t. \tag{40}$$

Replace wages and the dividends in the households' budget constraint yields:

$$\widehat{c}_t^H = \left(\varphi\widehat{n}_t^H + \gamma\widehat{c}_t^H\right)\left(1-\frac{\tau^D}{\lambda}\right) + \widehat{n}_t^H - g_t. \tag{41}$$

and using (40) yields:

$$\widehat{c}_t^H = \left(\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H\right)\left(1 - \frac{\tau^D}{\lambda}\right) + \widehat{n}_t^H - g_t. \tag{42}$$

Finally, consumption of h is given by:

$$\widehat{c}_t^H = \chi \widehat{y}_t - \left[\frac{\chi - 1}{1 + \frac{\varphi}{\gamma}} + 1\right] g_t \tag{43}$$

which is

$$\widehat{c}_t^H = \chi \widehat{y}_t^d + \left[\frac{\chi - 1}{1 + \frac{\gamma}{\varphi}}\right] g_t \tag{44}$$

where $\widehat{y}_t^d$ is defined as the disposable income of all households.

The consumption of unconstrained households is then given by (using the market clearing condition and inserting disposable income):

$$\widehat{c}_t^U = \frac{1 - \lambda\chi}{1 - \lambda} \widehat{y}_t^d - \frac{\lambda}{1 - \lambda} \frac{\chi - 1}{1 + \frac{\gamma}{\varphi}} g_t. \tag{45}$$

The IS curve in terms of aggregate consumption is then obtained by plugging the consumption of the hand-to-mouth and of unconstrained households into the Euler equation of unconstrained households and using $\widehat{c}_t = \widehat{y}_t^d$.

$$\begin{aligned} \frac{1 - \lambda\chi}{1 - \lambda} \widehat{c}_t - \frac{\lambda}{1 - \lambda} \frac{\chi - 1}{1 + \frac{\gamma}{\varphi}} g_t = s\mathbb{E}_t^{BR} &\left[\frac{1 - \lambda\chi}{1 - \lambda} \widehat{c}_{t+1} - \frac{\lambda}{1 - \lambda} \frac{\chi - 1}{1 + \frac{\gamma}{\varphi}} g_{t+1}\right] \\ +(1 - s)\mathbb{E}_t^{BR} &\left[\chi \widehat{c}_{t+1} + \left[\frac{\chi - 1}{1 + \frac{\gamma}{\varphi}}\right] g_{t+1}\right] - \frac{1}{\gamma} \mathbb{E}_t(\widehat{i}_t - \pi_{t+1}), \end{aligned}$$

using similar derivations as in Appendix A.2, we have

$$\begin{aligned} \widehat{c}_t &= \psi_f \mathbb{E}_t \widehat{c}_{t+1} - \frac{1}{\gamma} \psi_c \mathbb{E}_t(\widehat{i}_t - \pi_{t+1}) + \frac{\lambda}{1 + \frac{\gamma}{\varphi}} \frac{\chi - 1}{1 - \lambda\chi} g_t - \left[s \frac{\lambda}{1 - \lambda\chi} \frac{\chi - 1}{1 + \frac{\gamma}{\varphi}} + (1 - s) \frac{\chi - 1}{1 + \frac{\gamma}{\varphi}} \frac{1 - \lambda}{1 - \chi\lambda}\right] \mathbb{E}_t^{BR} g_{t+1} \\ &= \psi_f \mathbb{E}_t \widehat{c}_{t+1} - \frac{1}{\gamma} \psi_c \mathbb{E}_t(\widehat{i}_t - \pi_{t+1}) + \zeta \left[\frac{\lambda(\chi - 1)}{1 - \lambda\chi}(g_t - \bar{m}\mathbb{E}_t g_{t+1}) + (\delta - 1)\bar{m}\mathbb{E}_t g_{t+1}\right] \end{aligned}$$

with $\zeta = \frac{1}{1 + \frac{\gamma}{\varphi}}$.

A.4 Derivation of Proposition 2.

The first part comes from the fact that amplification is defined as

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

which requires $\chi > 1$.

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 (16), i.e., $\pi_t = \kappa \widehat{y}_t$, we can re-write the behavioral HANK IS equation from Proposition 1 as

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

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 the condition stated in Proposition 2:

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

Solving this for $\bar{m}$ yields

$$\bar{m} < \frac{1 - \frac{1-\lambda}{\gamma(1-\lambda\chi)}\kappa}{\delta},$$

which completes Proposition 2.

A.5 Derivation of Proposition 3.

Replacing $\widehat{i}_t$ by $\phi\pi_t = \phi\kappa\widehat{y}_t$ and $\mathbb{E}_t\pi_{t+1} = \kappa\mathbb{E}_t\widehat{y}_{t+1}$ in the IS equation (19), 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

$$\phi > \phi^* = 1 + \frac{\bar{m}\delta - 1}{\frac{\kappa}{\gamma}\frac{1-\lambda}{1-\lambda\chi}}, \tag{46}$$

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 (46) as

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

Proposition 3 can 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

$$\widehat{y}_t = \frac{\omega_f + \frac{\kappa}{\gamma}\omega_c \bar{m}^r}{1 + \frac{\omega_c}{\gamma}\left(\kappa\phi_\pi + \phi_y\right)} \mathbb{E}_t \widehat{y}_{t+1}. \tag{47}$$

From equation (47), it follows that we need

$$\phi_\pi > \bar{m}^r - \phi_y + \frac{\omega_f - 1}{\omega_c \frac{\kappa}{\gamma}} = \bar{m}^r - \phi_y + \frac{\bar{m}\delta - 1}{\frac{1-\lambda}{1-\chi\lambda}\frac{\kappa}{\gamma}} \tag{48}$$

for the model to feature a determinate, locally unique equilibrium. Condition (48) 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.6 Derivation of Proposition 7

To prove Proposition 7, we start from the Euler equation (15). Plugging in for $\widehat{c}_t^U$, $\widehat{c}_{t+1}^U$ and $\widehat{c}_{t+1}^H$ from equations (11) and (13), we get

$$\widehat{y}_t = s\mathbb{E}_t^{BR}\left[\widehat{y}_{t+1}\right] + (1-s)\frac{1-\lambda}{1-\lambda\chi}\mathbb{E}_t^{BR}\left[\widehat{y}_{t+1}\right] - \psi_c\left(\widehat{i}_t - \mathbb{E}_t\pi_{t+1}\right),$$

which can be re-written as

$$\widehat{y}_t = \delta\mathbb{E}_t^{BR}\left[\widehat{y}_{t+1}\right] - \psi_c\left(\widehat{i}_t - \mathbb{E}_t\pi_{t+1}\right).$$

Now, using the expectations setup from Proposition 6, we get $\delta\mathbb{E}_t^{BR}\left[\widehat{y}_{t+1}\right] = (1-\bar{m})\delta\widehat{y}_{t-1} + \bar{m}\delta\mathbb{E}_t\left[\widehat{y}_{t+1}\right]$ which proves Proposition 7.

A.7 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

$$X_{t+1} = G^X\left(X_t, \varepsilon_{t+1}\right), \tag{49}$$

where G^X denotes the transition function of X in equilibrium and ε are zero-mean innovations. Linearizing equation (49) yields

$$X_{t+1} = \Gamma X_t + \varepsilon_{t+1}, \tag{50}$$

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

$$X_{t+1} = \bar{m} G^X(X_t, \varepsilon_{t+1}), \tag{51}$$

or in linearized terms

$$X_{t+1} = \bar{m}\left(\Gamma X_t + \varepsilon_{t+1}\right). \tag{52}$$

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

$$\mathbb{E}_t^{BR}\left[\widehat{c}^U(X_{t+1})\right] = \bar{m}\mathbb{E}_t\left[\widehat{c}^U(X_{t+1})\right], \tag{53}$$

which we denote in the main text as

$$\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] = \bar{m}\mathbb{E}_t\left[\widehat{c}_{t+1}^U\right]. \tag{54}$$

Now, take the linearized Euler equation (15) of unconstrained households:

$$\widehat{c}_t^U = s\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] + (1-s)\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^H\right] - \frac{1}{\gamma}\widehat{r}_t, \tag{55}$$

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

$$\widehat{c}^U(X_t) = s\mathbb{E}_t^{BR}\left[\widehat{c}^U(X_{t+1})\right] + (1-s)\mathbb{E}_t^{BR}\left[\widehat{c}^H(X_{t+1})\right] - \frac{1}{\gamma}\widehat{r}(X_t). \tag{56}$$

Now, applying the results above, we obtain

$$\widehat{c}^U(X_t) = s\bar{m}\mathbb{E}_t\left[\widehat{c}^U(X_{t+1})\right] + (1-s)\bar{m}\mathbb{E}_t\left[\widehat{c}^H(X_{t+1})\right] - \frac{1}{\gamma}\widehat{r}(X_t), \tag{57}$$

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 behavioral HANK IS equation in Proposition 1.

A.8 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 this signal-extraction problem generates a set-up in which the family head behaves as if she was boundedly rational.

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 every agent j within the family of unconstrained households (the expectations of the hand-to-mouth agents are irrelevant) receives a noisy signal of X_{t+1}, S_{t+1}^j, given by

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

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 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 conditional mean of X_{t+1}, given the signal S_{t+1}^j, is given by

$$X_{t+1}^e \equiv \mathbb{E}\left[X_{t+1} | S_{t+1} = s_{t+1}^j\right] = p \cdot s_{t+1}^j.$$ [34]

The intuition is that the signal distribution is such that the agent either receives a perfectly

[34] To see this, note that the joint distribution of (X_{t+1}, S_{t+1}^j) is

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

where $g(X_{t+1})$ denotes the distribution of X_{t+1} and δ is the Dirac function. Given that the unconditional precise signal or a completely uninformative signal. As the perfectly-precise signal arrives with probability p and the unconditional mean is zero, it follows that the agent puts a weight p on the signal.

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} within the family is given by

$$\begin{aligned}\mathbb{E}\left[X^{e}_{t+1}(S_{t+1})|X_{t+1}\right] &= \mathbb{E}\left[p \cdot S_{t+1}|X_{t+1}\right] \\ &= p \cdot \mathbb{E}\left[S_{t+1}|X_{t+1}\right] \\ &= p^2 X_{t+1}.\end{aligned}$$

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

$$X_{t+1} = \bar{m}\left(\Gamma X_t + \varepsilon_{t+1}\right), \tag{58}$$

as imposed in equation (52). The boundedly-rational expectation of X_{t+1} is then given by

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

mean of X_{t+1} is 0, i.e., $\int x_{t+1} g(x_{t+1}) dx_{t+1} = 0$, it follows that

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

B Calibration

Parameter	Value	Source/Target
	HANK Parameters	
γ	1	Bilbiie (2020)
κ	0.02	Bilbiie (2020)
χ	1.48	Bilbiie (2020)
λ	0.33	Bilbiie (2020)
s	$0.8^{1/4}$	Bilbiie (2020)
	Behavioral Parameter	
$\bar{m}$	0.85	Gabaix (2020)

Table 1: Baseline calibration.

Our baseline calibration is summarized in Table 1. The values for γ and κ are directly taken from Bilbiie (2021, 2020) 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.

The household heterogeneity parameters, χ, λ and s are also standard in the analytical HANK literature (see Bilbiie (2020)). The most important assumption for our qualitative results in Section 3 is $\chi > 1$, which is consistent with the data. Patterson (2019) provides empirical evidence for the countercyclicality of inequality. Coibion et al. (2017), Mumtaz and Theophilopoulou (2017), Ampudia et al. (2018) and Samarina and Nguyen (2019) provide evidence of countercyclical inequality conditional on monetary policy shocks. Almgren et al. (2019) show that output in countries with higher shares of hand-to-mouth households responds more strongly to monetary policy shocks which, through the lens of the model, calls for $\chi > 1$.

For figure 5, i.e., to compute the iMPCs we choose a yearly calibration with $s = 0.8$ and $\beta = 0.95$ (this calibration is close to the iMPC exercise in Bilbiie (2021) but while he fixes χ to match the empirically-observed iMPCs, we vary χ together with $\bar{m}$ to examine their joint effects on iMPCs).

The Cognitive Discounting Parameter $\bar{m}$. The cognitive discounting parameter $\bar{m}$ is set to 0.85, as in Gabaix (2020) and Benchimol and Bounader (2019). Fuhrer and Rudebusch (2004), for example, estimate an IS equation and find that $\psi_f \approx 0.65$, 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. Note, that the calibration of the backward-looking behavioral HANK model in Section 4.2, which is based on household survey expectations and taken from Angeletos and Huo (2021), is close to the estimation results from Fuhrer and Rudebusch (2004).

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 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 $\bar{m}$.

C Extensions

C.1 Allowing for Steady State Inequality

So far, 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

$$\widehat{y}_t = \bar{m}\tilde{\delta}\mathbb{E}_t \widehat{y}_{t+1} - \frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\left(\widehat{i}_t - \mathbb{E}_t \pi_{t+1}\right), \tag{59}$$

with

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

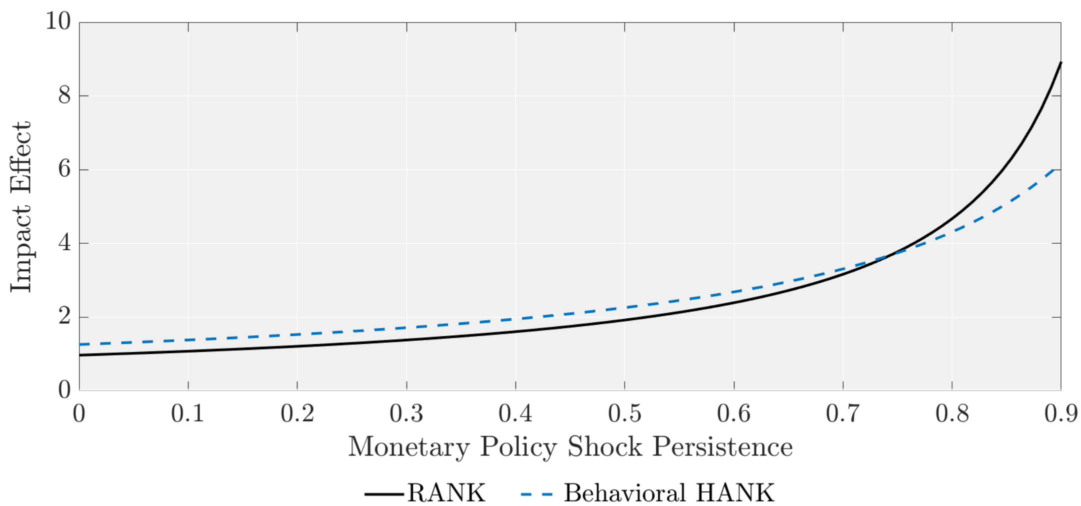
From a qualitative perspective, the whole analysis in the paper could be carried out with $\tilde{\delta}$ instead of δ. Quantitatively the differences are small as well. For example, if we set $\Omega = 1.5$, we get $\tilde{\delta} = 1.074$ instead of $\delta = 1.051$. Thus, we need $\bar{m} < 0.91$ instead of $\bar{m} < 0.93$ for determinacy under a peg.

C.2 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 10 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 10: 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.

C.3 Forward-Looking NKPC and Real Interest Rates

In the main part of the paper, we made the assumption that agents are rational with respect to real interest rates (as in Gabaix (2020)) and assumed a static Phillips Curve (as in Bilbiie (2021)). 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:

$$\begin{pmatrix} \mathbb{E}_t \pi_{t+1} \\ \mathbb{E}_t \widehat{y}_{t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{\beta M^f} & -\frac{\kappa}{\beta M^f} \\ \frac{\psi_c}{\gamma \psi_f} \left(\phi - \frac{\bar{m}}{\beta M^f} \right) & \frac{1}{\psi_f} \left(1 + \frac{\psi_c \bar{m} \kappa}{\gamma \beta M^f} \right) \end{pmatrix}}_{\equiv A} \begin{pmatrix} \pi_t \\ \widehat{y}_t \end{pmatrix}. \tag{60}$$

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 $M^f = 0$. In this case, it follows that we have a uniquely-determined (bounded) equilibrium for $\phi > -3.22$. Thus, the condition is even weaker than in the main part of the paper.

If we allow for a forward-looking Phillips curve and using the same calibration as in the main text and relying on Gabaix (2020) for the newly-introduced parameters, $\theta = 0.875$, it follows that we again have determinacy under an interest rate peg for our baseline calibration with $\bar{m} = 0.85$.

D Details on Intertemporal MPCs

In this section, we derive the iMPCs discussed in Section 3.5. Defining Y_t^j as type j's disposable income, we can write the households' budget constraints as

$$C_t^H = Y_t^H + \frac{1-s}{\lambda} R_t B_t$$
$$C_t^U + \frac{1}{1-\lambda} B_{t+1} = Y_t^U + \frac{s}{1-\lambda} R_t B_t,$$

where R_t denotes the real interest rate and B_t real bonds. Log-linearizing the two budget constraints around the zero-liquidity steady state and $R = \beta^{-1}$ yields

$$\widehat{c}_t^H = \widehat{y}_t^H + \frac{1-s}{\lambda} \beta^{-1} b_t \tag{61}$$
$$\widehat{c}_t^U + \frac{1}{1-\lambda} b_{t+1} = \widehat{y}_t^U + \frac{s}{1-\lambda} \beta^{-1} b_t, \tag{62}$$

where b_t denotes real bonds in shares of steady state output. Aggregating (61) and (62) delivers

$$\widehat{c}_t = \tilde{y}_t + \beta^{-1} b_t - b_{t+1}, \tag{63}$$

where $\tilde{y}_t$ denotes aggregate disposable income.

By plugging equations (61) and (62) into the Euler equation of unconstrained households (15), we can derive the dynamics of liquid assets b_t (ignoring changes in the real rate as this is a partial equilibrium exercise):

$$\begin{aligned} &\mathbb{E}_t b_{t+2} - b_{t+1} \left[\frac{1}{s\bar{m}} + \beta^{-1} s + \frac{(1-s)^2 \beta^{-1} (1-\lambda)}{s\lambda} \right] + \frac{\beta^{-1}}{\bar{m}} b_t = \\ &(1-\lambda) \mathbb{E}_t \widehat{y}_{t+1}^U + \frac{1-s}{s} (1-\lambda) \mathbb{E}_t \widehat{y}_{t+1}^H - \frac{1-\lambda}{s\bar{m}} \widehat{y}_t^U . \end{aligned} \tag{64}$$

Note that a change in total disposable income by one changes the hand-to-mouth households' disposable income by χ and the disposable income of unconstrained households by $\frac{1-\lambda\chi}{1-\lambda}$.

Let us denote the right-hand side of equation (64) by $-\mathbb{E}_t \widehat{z}_t$. Factorizing the left-hand side and letting F denote the forward-operator, it follows that

$$(F - \mu_1)(F - \mu_2) \mathbb{E}_t b_t = -\mathbb{E}_t \widehat{z}_t, \tag{65}$$

where μ_1 and μ_2 denote the roots of the characteristic equation

$$\mathbb{E}_t b_{t+2} - \phi_1 b_{t+1} - \phi_2 b_t = 0, \tag{66}$$

where

$$\phi_1 \equiv \left[\frac{1}{s\bar{m}} + \beta^{-1}s + \frac{(1-s)^2\beta^{-1}(1-\lambda)}{s\lambda}\right] \tag{67}$$

and

$$\phi_2 \equiv -\frac{\beta^{-1}}{\bar{m}}. \tag{68}$$

Thus, the roots are given by

$$\mu_{1,2} = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}. \tag{69}$$

It follows that

$$\begin{aligned} b_{t+1} &= \mu_1 b_t - (F - \mu_2)^{-1}\mathbb{E}_t\widehat{z}_t \\ &= \mu_1 b_t + \frac{\mu_2^{-1}}{1 - F\mu_2^{-1}}\mathbb{E}_t\widehat{z}_t. \end{aligned}$$

Note that $\mathbb{E}_t\widehat{z}_t$ can be written as $\frac{1-\lambda\chi}{s}\left(\delta\mathbb{E}_t\widehat{y}_{t+1} - \frac{1}{\bar{m}}\widehat{y}_t\right)$. Without loss of generality, we let $\mu_2 > \mu_1$ and we have $\mu_2 > 1$. We have $(1 - F\mu_2^{-1})^{-1} = \sum_{l=0}^{\infty}\mu_2^{-l}F^l$. Thus, we end up with

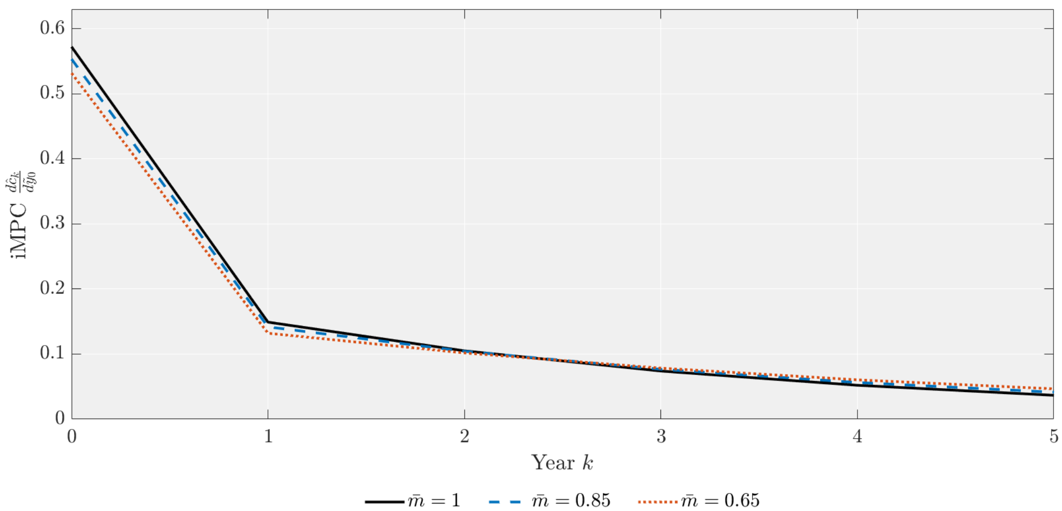
$$b_{t+1} = \mu_1 b_t + \frac{1-\lambda\chi}{s}\sum_{l=0}^{\infty}\mu_2^{-(l+1)}\mathbb{E}_t\left(\frac{1}{\bar{m}}\widehat{y}_{t+l} - \delta\widehat{y}_{t+1+l}\right). \tag{70}$$

Taking derivatives with respect to $\widehat{y}_{t+k}$ yields Proposition 5.

iMPCs for more than two periods. Figure 11 plots the MPCs for the year of the income windfall as well as the five consecutive years for different degrees of rationality. As discussed in section 3.5, under our benchmark calibration, the rational model predicts somewhat larger initial MPCs as behavioral, unconstrained households save relatively more. Over time, however, the MPCs in the behavioral model lie above their rational counterparts due to the fact that more and more of the initial unconstrained households become hand-to-mouth and start consuming their (higher) savings. As Figure 12 shows, the probability of type switching, $1-s$, matters for when exactly the behavioral model starts to generate larger MPCs compared to the rational model.

iMPCs and the Role of Idiosyncratic Risk. In Figure 12, we plot he MPCs in the year of the income windfall (left panel) and the first year after the windfall (right panel) for a relatively high idiosyncratic risk of $1-s=0.5$. The high probability of becoming hand-to-mouth flips the role of $\bar{m}$ for the MPC_1 compared to our baseline calibration as discussed in Section 3.5. The reason being that the behavioral, unconstrained households save a relatively large amount of the received income windfall in period 0 as they cognitively

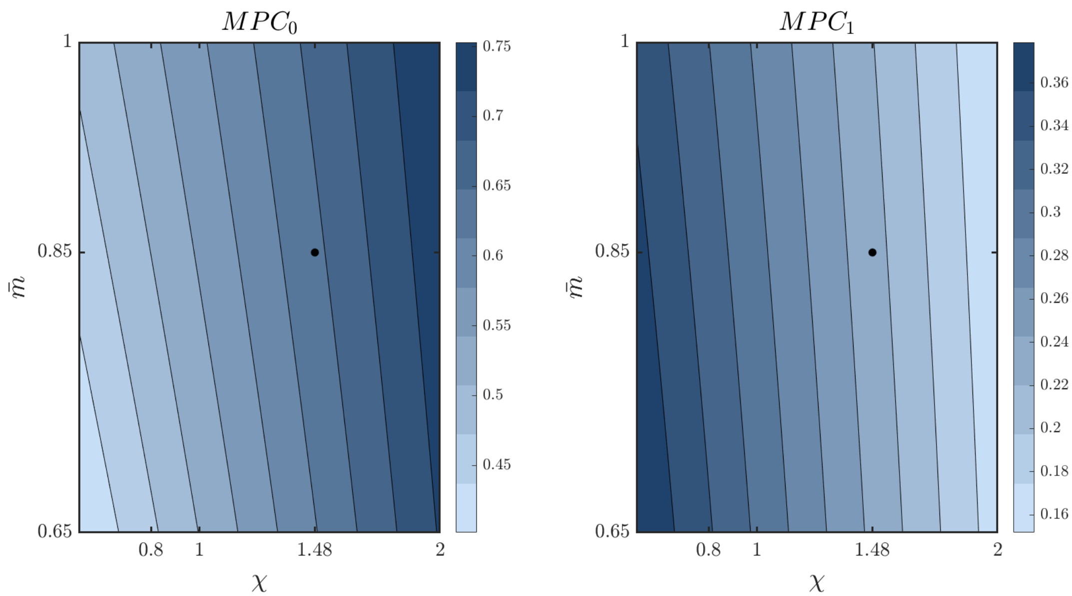
Figure 11: Intertemporal MPCs



Note: This figure shows the aggregate intertemporal MPCs, i.e., the aggregate consumption response in year k to a change in aggregate disposable income in year 0 for different $\bar{m}$.

discount the decrease in their future marginal utility. Thus, they end up with relatively more disposable income in year 1. Now, given the relatively high probability of type switching, there are many unconstrained households who end up being hand-to-mouth in year 1 after the income windfall. As they are hand-to-mouth, they consume their previously-accumulated savings which increases the MPC_1. The more behavioral unconstrained households are, i.e., the lower $\bar{m}$ is, the more pronounced this effect and hence, a lower $\bar{m}$ increases the MPC_1 in the case of a relatively high $1 - s$.

Figure 12: Intertemporal MPCs, Bounded Rationality and Household Heterogeneity



Note: This figure shows the aggregate intertemporal MPCs, i.e., the aggregate consumption response in year 0 (left) and year 1 (right) to a change in aggregate disposable income in year 0 for a transition probability $1 - s = 0.5$.

E Sticky Wages

In this section, we provide details on the sticky-wage extension presented in Section 4.1 as well as the calibration used to produce Figures 8 and 9. The way we introduce sticky wages follows Colciago (2011) and recently adopted by Bilbiie et al. (2021).[35]

In the household block, the only difference to our benchmark model is that we assume that there is a labor union pooling labor and setting wages on behalf of households. This leads to a condition similar to the labor-leisure conditions in Section 2. But instead of individual conditions, the condition is the same for every household:

$$\varphi \widehat{n}_t = \widehat{w}_t - \gamma \widehat{c}_t,$$

and $\widehat{n}_t = \widehat{n}_t^U = \widehat{n}_t^H$.

The labor union, however, is subject to wage rigidities. The nominal wage can only be re-optimized with a constant probability, which leads to a time-varying wage markup

$$\widehat{\mu}_t^w = \varphi \widehat{n}_t - \widehat{w}_t + \gamma \widehat{c}_t,$$

and a wage Phillips Curve

$$\pi_t^w = \beta \mathbb{E}_t \pi_{t+1}^w + \kappa_w \widehat{\mu}_t^w.$$

Wage inflation is given by

$$\pi_t^w = \widehat{w}_t - \widehat{w}_{t-1} + \pi_t.$$

The firm side is exactly the same as in the main text but we focus on the case with rational firms, which gives rise to a standard Phillips Curve:

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa_\pi \widehat{mc}_t,$$

where $\widehat{mc}_t$ is a time-varying price markup. Table 2 summarizes all equilibrium equations.

The calibration of this extended model is presented in Table 3. The parameters γ, φ, s, β and $\bar{m}$ are as in our baseline calibration. The parameters of the Taylor rule, ρ_i and ϕ, are set as estimated in Auclert et al. (2020).

The slope of the wage Phillips curve, κ_w, is set as in Bilbiie et al. (2021) and we focus on the *no-redistribution* case $\tau^D = 0$. Note, that this leads to impact responses of consumption of the two household types that are very close to the ones in our baseline model: $\widehat{c}_t^H$ increases by about 1.42, whereas output increases by 1. The baseline calibration of $\chi = 1.48$ would

[35]See also Erceg et al. (2000). Broer et al. (2020) and Broer et al. (2021b) discuss the role of sticky wages in (rational) TANK models for the analysis of monetary and fiscal policy, respectively.

Table 2: Sticky Wages, Equilibrium Equations

Name	Equation
Wage Markup	$\widehat{\mu}_t^w = \gamma\widehat{c}_t + \varphi\widehat{n}_t - \widehat{w}_t$
Wage Phillips Curve	$\pi_t^w = \beta\mathbb{E}_t\pi_{t+1}^w + \kappa_w\widehat{\mu}_t^w$
Wage Inflation	$\pi_t^w = \widehat{w}_t - \widehat{w}_{t-1} + \pi_t$
Bond Euler	$\widehat{c}_t^U = s\bar{m}\mathbb{E}_t\widehat{c}_{t+1}^U + (1-s)\bar{m}\mathbb{E}_t\widehat{c}_{t+1}^H - \frac{1}{\gamma}(\widehat{i}_t - \mathbb{E}_t\pi_{t+1})$
H Budget Constraint	$\widehat{c}_t^H = \widehat{w}_t + \widehat{n}_t + \widehat{t}_t^H$
H Transfer	$\widehat{t}_t^H = \frac{\tau^D}{\lambda}D_t$
Profits	$\widehat{d}_t = \widehat{y}_t - (\widehat{w}_t + \widehat{n}_t)$
Labor Demand	$\widehat{w}_t = \widehat{mc}_t + \widehat{y}_t - \widehat{n}_t$
Phillips Curve	$\pi_t = \beta\mathbb{E}_t\pi_{t+1} + \kappa_\pi\widehat{mc}_t$
Production	$\widehat{y}_t = \widehat{n}_t$
Consumption	$\widehat{c}_t = \lambda\widehat{c}_t^H + (1-\lambda)\widehat{c}_t^U$
Resource Constraint	$\widehat{y}_t = \widehat{c}_t$
Taylor Rule	$\widehat{i}_t = \rho_i\widehat{i}_{t-1} + (1-\rho_i)\phi\pi_t + \varepsilon_t^{MP}$

Table 3: Sticky Wage Model Calibration.

Parameter	γ	κ_π	λ	s	φ	τ^D	κ_w	β	ρ_i	ϕ
Value	1	0.01	0.37	$0.8^{1/4}$	1	0	0.075	0.99	0.89	1.5

predict that in the model without sticky wages, $\widehat{c}_t^H$ increases by 1.48 when output increases by 1. We focus on a relatively stable inflation and set κ_π to 0.01.

The only parameter that we change with respect to our baseline calibration is λ which we set to 0.37 instead of 0.33. A value of 0.37 is still in the range of often used values (see, for example Bilbiie (2020)). We increase λ somewhat compared to our baseline calibration in order to increase the role of hand-to-mouth households in the response to monetary policy shocks and thus, allows the model to generate the pronounced hump-shaped responses. Setting $\lambda = 0.33$ still produces hump-shaped responses but those are somewhat less pronounced.

F A Quantitative Behavioral HANK Model

Table 4 shows how we calibrate the quantitative model introduced in Section 3.7.

The calibration closely follows the parameterization in McKay et al. (2016). As in McKay et al. (2016), we assume that high productivity households pay all the taxes. The main difference to their calibration is that they assume that every household receives an equal share of the dividends whereas we assume that the high productivity households receive 80% of the dividend payments, while the middle productivity class receive 20% of it. The low productivity households do not receive any dividend payments. We choose this calibration such that the contemporaneous amplification in the quantitative HANK model matches the one from the tractable model, outlined in Section 2. Note that this dividend distribution is in line with empirical findings in Kuhn et al. (2020).

Parameter	Description	Value
R	Steady State Real Rate	0.5%
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	0.0384
$\tau(e)$	Tax shares	[0, 1, 1]
$d(e)$	Dividend shares	[0, 0.2, 0.8]
$\frac{B^G}{4Y}$	Total wealth	0.625

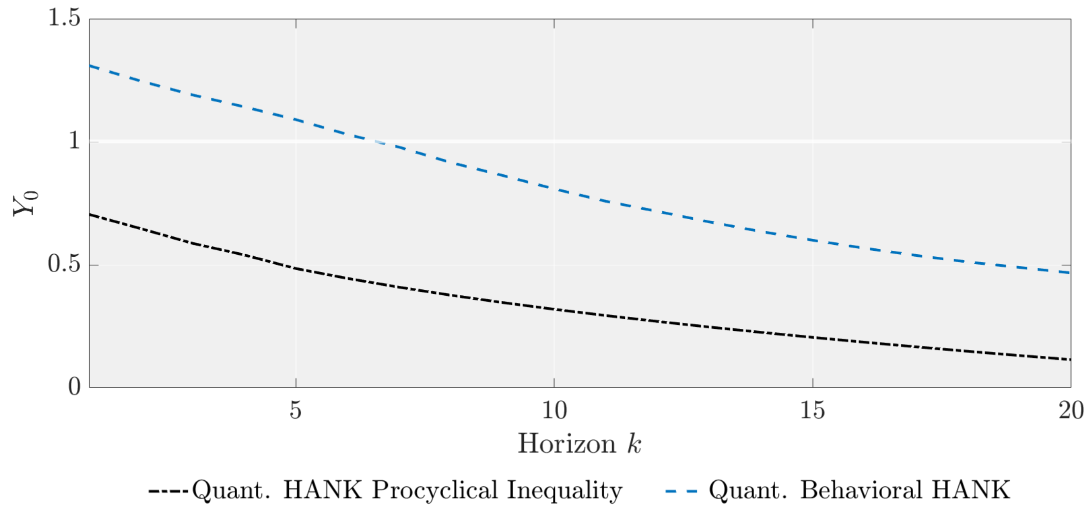
Table 4: Baseline calibration of quantitative HANK model.

Other resolutions of the forward-guidance puzzle in quantitative HANK model. How does our quantitative behavioral HANK model compare to other resolutions of the forward guidance puzzle within one-asset HANK models? McKay et al. (2016) resolve the forward guidance puzzle by assuming that every household receives an equal share of the dividends, leading to pro-cyclical inequality. Thus, the low-productivity households—who also exhibit larger MPCs on average than households with higher productivities—are less exposed to monetary policy. Therefore, the effectiveness of monetary policy is dampened overall, leading to a resolution of the forward guidance puzzle but also ruling out amplification of contemporaneous shocks, as shown Figure 13.

Second, Hagedorn et al. (2019) solve the forward guidance puzzle by introducing a nominal anchor into their model. In particular, they impose a nominal steady state government debt level, which implies that the model has a steady state price level. This allows them to resolve the forward guidance puzzle and generate amplification of contemporaneous monetary policy. We show how introducing bounded rationality also sidesteps the Catch-22 without relying on a nominal anchor.

Third, Farhi and Werning (2019) suggest a similar resolution to the forward guidance puzzle as our model by combining incomplete markets and bounded rationality. Our behavioral HANK model differs from theirs in two dimension: first, we introduce bounded rationality in the form of cognitive discounting whereas Farhi and Werning (2019) assumes level-k think-

Figure 13: Resolving the Forward Guidance Puzzle in HANK



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, relative to the response in the RANK model under rational expectations (equal to 1).

ing. Second, in our model contemporaneous monetary policy is amplified whereas it is not in Farhi and Werning (2019).

G Details on the Backward-Looking Behavioral HANK Model

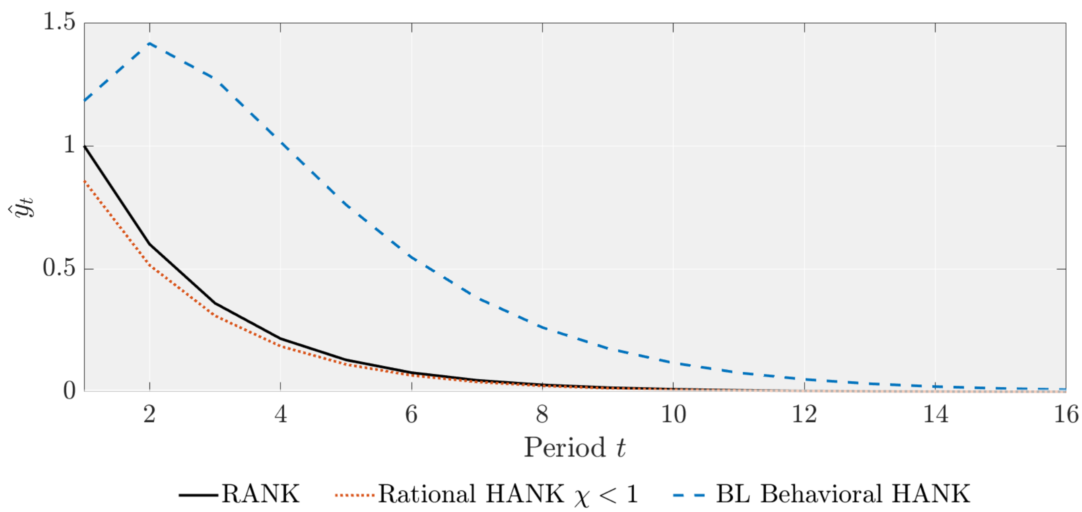
Here, we discuss how we can calibrate the backward-looking behavioral HANK model from Section 4.2 to match data coming from survey expectations. To do so, we follow Angeletos and Huo (2021) who calibrate the coefficients in front of $\mathbb{E}_t\widehat{y}_{t+1}$ and $\widehat{y}_{t-1}$ to match exactly this kind of evidence from survey expectations data. By following their calibration, we can back out the implied $\bar{m}$ and χ. We get $\bar{m} = 0.59$ and $\chi = 0.72$, thus, relatively low values compared to the calibration above. We leave the other parameters as in Section 3. We complement the backward-looking behavioral HANK IS equation with the static Phillips Curve (16).

Determinacy. We numerically verify that the backward-looking behavioral HANK model restores the Taylor principle. In fact, the equilibrium is determinate even under an interest-rate peg. Thus, also the backward-looking behavioral HANK model overturns the Sargent and Wallace (1975) result with this calibration.

Impulse-Response Functions. We now show how the backward-looking behavioral HANK model generates hump-shaped impulse responses and a novel behavioral amplification channel. To this end, we examine how output in the backward-looking behavioral HANK model responds to an expansionary monetary policy shock and compare the response to its rational counterpart and the RANK version of the model. We set the Taylor coefficient to 1.5, thus, guaranteeing determinacy also in the rational models and the persistence of the shock to an intermediate value, $\rho^{MP} = 0.6$.

Figure 14 shows the corresponding impulse-response functions. The blue-dashed line shows the results of our behavioral HANK, the orange-dotted line of its rational counterpart (THANK) and the black-solid line of RANK.

Figure 14: Output Response to a Monetary Policy Shock



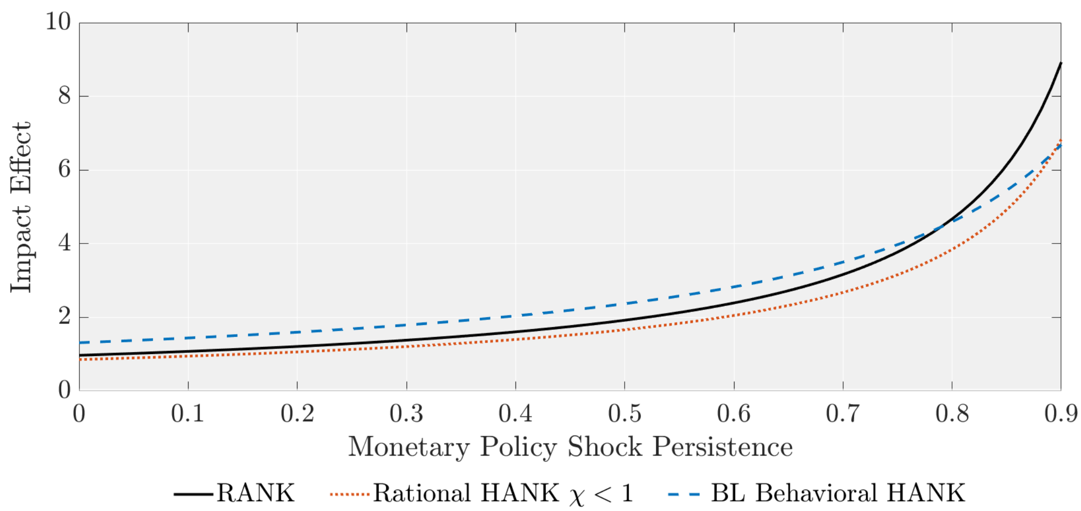
Note: This figure shows the output response to a monetary policy shock for different models.

Two things stand out. First, the behavioral HANK model delivers amplification compared to RANK—even in the first period—and second, the backward-looking anchor generates hump-shaped responses. As the latter has been highlighted in Angeletos and Huo (2021), we here focus on the amplification. Figure 14 shows that the amplification stems from a *behavioral amplification channel*: the initial output response is amplified although the model features procyclical inequality ($\chi < 1$) and, thus, the heterogeneity frictions themselves would generate dampening.

Where does the behavioral amplification come from? Given the backward-looking component in households' expectations, the increase in today's output is expected to persist as it becomes tomorrow's default value for the household's expectations. The behavioral anchor induces *endogenous* persistence which further increases today's output response through more optimistic expectations. Yet, there is an opposing channel at work: an *exogenously* persistent shock not only decreases interest rates today but also expected future interest rates. Behavioral households congitively discount these future changes and, thus, perceive the shock as less expansionary compared to a rational agent which dampens the initial response.[36] Given our calibration, the first channel dominates, thereby generating amplification as depicted in Figure 14.

Given the two opposing forces at work, the degree of initial amplification depends on the persistence of the shock. Figure 15 shows the initial response of all three models for different degrees of persistence of the shock. As the persistence declines, the initial response becomes relatively stronger in the backward-looking behavioral HANK model compared to RANK. As a consequence, the relative amplification is largest for an i.i.d. shock.

Figure 15: 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.

In addition, comparing the backward-looking behavioral HANK model to its rational counterpart shows that for $\rho^{MP} < 0.9$, there is behavioral amplification while for more persistent shocks, there is behavioral dampening. The comparison with RANK shows that for $\rho^{MP} < 0.80$, the behavioral amplification dominates the heterogeneity dampening which arises because $\chi < 1$.

[36] This is the same channel through which the fiscal multiplier of persistent government spending is dampened in our baseline model in Section 3.

Behavioral Amplification and Forward Guidance. We now analyze analytically the behavioral-amplification mechanism and its implications for forward guidance. In the backward-looking behavioral HANK model, the output response to an interest rate change depends on the (expected) infinite future even when the shock is completely transitory.

Consider the following. The monetary authority decreases the nominal interest rate in period t to $\widetilde{i}_t < 0$ but will keep it at steady state thereafter (the argument extends to changes of the interest rate in the future). Output and inflation would be expected to go back to zero in $t+1$ under rational expectations. This is, however, not true for the backward-looking behavioral HANK model.

To understand this, combine the static Phillips Curve (a static Phillips curve is again not crucial for the argument but facilitates the derivations) with the behavioral HANK IS equation to arrive at

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

If households expect future output to be back to steady state – as would be the case in the rational model or the behavioral model in which the households' default value equals the steady state – a one-time, completely transitory decrease in the nominal interest rate changes contemporaneous output by

$$\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi} > 0. \tag{71}$$

Yet, in the backward-looking model, expectations in $t+1$ of output in $t+2$ will be above steady state when output in t increases. The more optimistic expectations feed back into output already in t.

This becomes apparent when we write the IS equation as

$$\begin{aligned}
&\widehat{y}_t\left[1-(1-\bar{m})\delta\left[\delta\bar{m}+\kappa\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\right]\right] = \\
&(1-\bar{m})\delta\widehat{y}_{t-1} - \frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\left[\widetilde{i}_t + \left[\delta\bar{m}+\kappa\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\right]\mathbb{E}_t\left[\widetilde{i}_{t+1}\right]\right] \\
&+\left[\delta\bar{m}+\kappa\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\right]^2\mathbb{E}_t\widehat{y}_{t+2}.
\end{aligned}$$

Thus, if households would assume that $\widehat{y}_{t+2}$ will be zero but not $\widehat{y}_{t+1}$, the discussed interest-rate change in t increases output in t by

$$\frac{\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}}{1-(1-\bar{m})\delta\left[\delta\bar{m}+\kappa\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\right]},$$

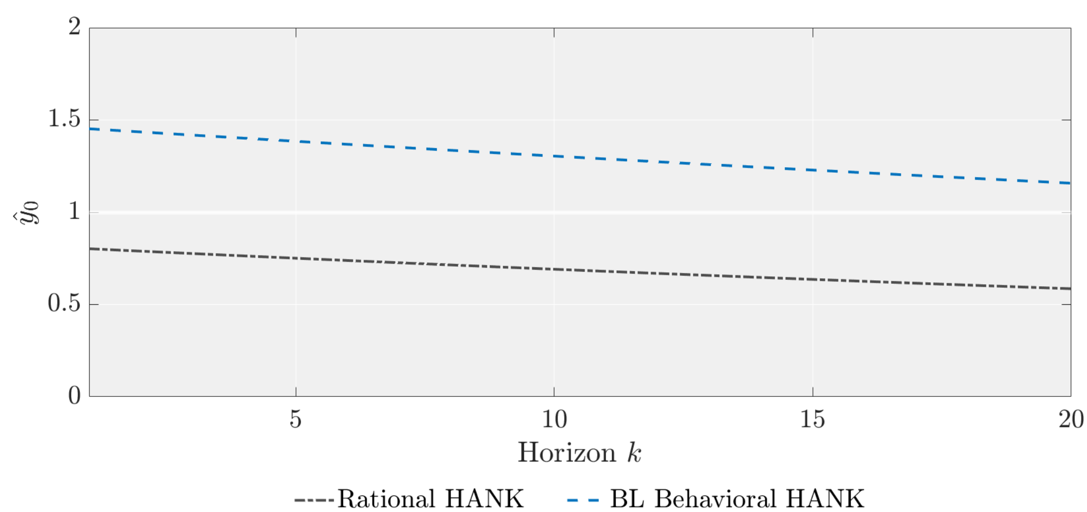
which is larger than the effect for models without a backward-looking anchor as can be seen by comparing it to equation (71). Put differently, the initial output response is amplified through a behavioral channel. Iterating forward in this fashion shows how the effect increases with each iteration. However, the response is bounded, as we will see below.

Turning to forward guidance, an expected change in the nominal interest rate in period $t+1$, affects output in t by

$$-\frac{\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\left[\delta\bar{m}+\kappa\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\right]}{1-(1-\bar{m})\delta\left[\delta\bar{m}+\kappa\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\right]},$$

if we assume output in $t{+}2$ to be back to zero. Given our calibration, the term $\left[\delta\bar{m}+\kappa\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\right]$ is smaller than 1. Thus, an interest rate change tomorrow has a smaller effect on output today than a contemporaneous interest rate change such that there is no forward guidance puzzle in the backward-looking behavioral HANK model. We can continue in this fashion to show that the effects increase with the iteration but decrease with the period of the shock.

Figure 16: Forward Guidance with Backward-Looking Anchor



Note: This figure shows the period-t output response to an anticipated i.i.d. monetary policy shock in period $t+k$ for three different economies.

Figure 16 shows these patterns graphically. First, the behavioral amplification channel discussed above is reflected in the contemporaneous effect ($k = 0$) which is stronger than without the backward-looking expectations —reflected in the black-dashed-dotted line. Second, increasing the horizon k shows that there is no forward guidance puzzle in the backward-looking behavioral HANK model. To sum it up, also the backward-looking behavioral HANK model amplifies contemporaneous monetary policy (even for $\chi < 1$) while it simultaneously dampens the effects of forward guidance.