A Behavioral Heterogeneous Agent New Keynesian Model

Oliver Pfäuti* Fabian Seyrich†

June 13, 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. Consistent with the data, monetary policy mainly works through indirect general equilibrium effects and fiscal multipliers on consumption are positive. In contrast to existing workhorse models, the model can account for these empirical facts while simultaneously resolving the forward guidance puzzle and remaining stable at the effective lower bound as the model features equilibrium determinacy even under an interest-rate peg. According to our model, central banks have to increase nominal interest rates much more strongly after an inflationary cost-push shock in order to stabilize inflation, leading to substantially higher government debt levels.

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, pfaeuti.oliver@gmail.com.
†Berlin School of Economics, DIW Berlin, and Freie Universität Berlin, fabian.seyrich@gmail.com.
We thank Klaus Adam, George-Marios Angeletos, Neele Balke, Nicholas Barberis, Christian Bayer, 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, the SMYE 2022, the University of Warwick, the 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 sheds new light on the transmission mechanisms and the effectiveness of monetary and fiscal policy. Government spending increases private consumption substantially. Monetary policy is transmitted to household consumption to a large extent through indirect general equilibrium effects which tends to amplify the effectiveness of contemporaneous monetary policy. Announcements of future monetary policy changes, however, have relatively weak effects on current economic activity. Despite these weak effects of forward guidance, advanced economies have not experienced large instabilities during long spells at the binding effective lower bound.[1] Accounting for all these facts within one framework, however, turns out to be challenging for existing workhorse models.

In this paper, we propose a new framework—*the behavioral heterogeneous agent New Keynesian model*—which can account for all these empirical facts *simultaneously*. While we keep the New Keynesian core, we allow for household heterogeneity and bounded rationality in the form of cognitive discounting. We do so using two complementary approaches: first, we rely on a limited heterogeneity set-up which allows us to derive all results in an analytical-tractable way and, thus, provides a clear understanding of the role of the two frictions and their interaction. 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, none of which, however, can account for all the listed empirical facts simultaneously. In the second approach, we relax the limited heterogeneity set-up and show that all our results carry over to a full-blown, quantitative behavioral HANK model.

We use the quantitative behavioral HANK model to revisit how monetary policy has to be implemented after an inflationary cost-push shock. Many advanced economies have recently experienced a dramatic surge in inflation and at least part of this is attributed to cost pressures coming from the supply side. We find that in order to stabilize inflation after a cost-push shock, central banks need to increase nominal interest rates much more strongly than in the rational model because in the behavioral model, expected future interest rate increases are less effective. This leads to substantially higher government debt levels. The effect on government debt is especially pronounced when initial debt levels are already high.

[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.

To arrive at our analytical framework, we assume that there are two groups of households. One group of households is "unconstrained", in the sense that they participate in financial markets so that they are on their Euler equation. Households in the other group 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 households have a precautionary-savings motive. We introduce bounded rationality in the form of cognitive discounting. Households anchor their expectations about future macroeconomic variables to the steady state and cognitively discount expected future deviations from it, as introduced in a representative agent set-up by Gabaix (2020). As a result, average expectations underreact to news, as we show to be the case empirically across all income groups.[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 who 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. Decomposing the total effect into indirect and direct effects, we show that the major share of the monetary policy transmission works through indirect effects. In addition, after an exogenous increase in government spending, these general equilibrium effects also ensure positive fiscal multipliers on consumption in the benchmark case of a constant real interest rate.

Despite the amplification through general equilibrium effects the behavioral HANK 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. On the one hand, 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. On the other

[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 can lead to observationally-equivalent outcomes. For further evidence on the underreaction of aggregate expectations to news, see, for example, Coibion and Gorodnichenko (2015), Bordalo et al. (2020) and Angeletos et al. (2021).

hand, 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. This second channel dominates, thus, the model resolves the forward guidance puzzle.

The fact that the behavioral HANK model can generate amplification through indirect effects and resolve the forward-guidance puzzle simultaneously is in stark contrast to its rational counterpart. The rational model generates either amplification or resolves the forward guidance puzzle but not both at the same time (see Werning (2015) and Bilbiie (2021)). Related to the resolution of the forward guidance puzzle, the behavioral HANK model remains stable during prolonged periods at the effective lower bound (ELB), even in cases in which output tends to implode in rational models. The behavioral HANK model remains determinate even in the limiting case of an ever-binding ELB, as the model features equilibrium determinacy under an interest-rate peg.

The insights of our tractable behavioral HANK model are not due to the limited heterogeneity set-up. To show this, we build a quantitative behavioral HANK model in which households face uninsurable idiosyncratic risk, incomplete markets and borrowing constraints and in which liquidity is in positive net supply. In line with our tractable model, households cognitively discount aggregate variables but are fully rational with respect to their individual state and their idiosyncratic risk. In its rational expectation limit, the quantitative behavioral HANK model 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. As in the tractable model, the quantitative behavioral HANK model resolves the forward guidance puzzle, and 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. In addition, the quantitative behavioral HANK model is also more stable at the effective lower bound than its rational counterpart and it produces positive consumption multiplier independent of the persistence of the fiscal spending shock.

The quantitative behavioral HANK model also allows us to consider heterogeneous degrees of bounded rationality. We show that the data suggests that households with higher income tend to deviate somewhat less from rational expectations than households with lower income. Incorporating heterogeneous degrees of bounded rationality along these lines, we then show that our results are robust. In particular, the forward guidance puzzle is still resolved, even though forward guidance is slightly more effective than in the model without heterogeneous degrees of rationality. The reason is that households with higher income—who are now more rational—are much more likely to be unconstrained and therefore more important for the transmission of announced future interest rate changes. In contrast, less rational households tend to be liquidity constrained and thus, do not respond to announced interest rate changes anyway.

We close by extending our tractable baseline framework along three dimensions. We first allow for positive savings which enables us to analytically derive a key statistic in HANK models for monetary and fiscal policy analysis: intertemporal MPCs—or iMPCs for short (Auclert et al. (2018), Wolf (2021), Kaplan and Violante (2020)). The behavioral HANK model matches the iMPCs estimated in the data. We find that 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.

Second, 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 expectations in our setup are purely forward looking.

Third, we show how to extend our framework to derive an equivalence result between heterogeneous-household models with bounded rationality and those of incomplete information and learning. If 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)) and thus, features myopia and anchoring in the aggregate IS equation.[3]

[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.

Related Literature. The literature so far treats the empirical facts laid out in the Introduction mostly independent from each other. The HANK and TANK (Two Agent New Keynesian) literature – both with quantitative and analytical models – has 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 and establish a trade-off inherent in models with household heterogeneity: if HANK models amplify contemporaneous monetary policy (and fiscal policy) through redistribution 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]

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.

Farhi and Werning (2019) show that the combination of incomplete markets and level k-thinking can resolve the forward guidance puzzle. In contrast to their paper, we employ

[4] Acharya and Dogra (2020) construct a pseudo-RANK model, in which they isolate and highlight the role of precautionary savings dynamics to explain 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 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.

cognitive discounting instead of level k-thinking and we do not only focus on the forward guidance puzzle but show that our behavioral HANK model can combine the resolution of the forward guidance puzzle with indirect, general-equilibrium amplification of monetary and fiscal policy and generating iMPCs as in the data. Auclert et al. (2020) show that the combination of heterogeneous agents and sticky information can produce MPCs that jump on impact whereas macroeconomic aggregates respond in a hump-shaped fashion to a monetary policy shock. Our extended model with sticky wages can be seen as a tractable complementary to their full-blown quantitative model. Other papers that share the combination of household heterogeneity and some deviation from FIRE but do not share our focus include 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 (including Farhi and Werning (2019) and Auclert et al. (2020)), 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 develop the quantitative behavioral HANK model and show that the results from the tractable model carry over to the quantitative model, and we use the quantitative model to study the policy implications of an inflationary supply-side shock. We discuss three extensions of the behavioral HANK model in Section 5. Section 6 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. For now, we focus on a limited heterogeneity set-up which is typical in the analytical HANK literature to ensure closed-form solutions (e.g., McKay et al. (2017), Bilbiie (2021)). We turn to a full-blown incomplete markets set-up in Section 4 to show that none of our results are driven by our simplifying assumptions in this section. 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 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 the relative risk aversion.

Households can save in government bonds B_{t+1}^i, 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. 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. 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.

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 those of U households. When households switch from being unconstrained to being hand-to-mouth, they only keep their government bonds. Shares, on the other hand, cannot be used to self-insure.

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

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

[6]We show in Appendix A.9 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.

subject to the flow budget constraints of unconstrained households

$$C_t^U + B_{t+1}^U + v_t \iota_{t+1} = W_t N_t^U + \iota_t (v_t + \tilde{D}_t) + \frac{1 + i_{t-1}}{1 + \pi_t} \left(s B_t^U + (1 - h) \frac{\lambda}{1 - \lambda} B_t^H \right) + T_t^U, \quad (2)$$

and the hand-to-mouth households

$$C_t^H + B_{t+1}^H = W_t N_t^H + T_t^H + \frac{1 + i_{t-1}}{1 + \pi_t} \left((1 - s) \frac{1 - \lambda}{\lambda} B_t^U + h B_t^H \right), \quad (3)$$

as well as the borrowing constraints

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

where W_t is the real wage, v_t is the stock price, 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 budget constraints reflect our assumption that households keep their acquired government bonds when switching their type as well as the assumption of full-insurance within type, as the bonds are equally shared within types.

The optimality conditions are given by the Euler equations of unconstrained households and the hand-to-mouths' households

$$\frac{\partial \mathcal{U}(C_t^U, N_t^U)}{\partial C_t^U} \geq \beta \mathbb{E}_t^{BR} \left[R_t \left(s \frac{\partial \mathcal{U}(C_{t+1}^U, N_{t+1}^U)}{\partial C_{t+1}^U} + (1 - s) \frac{\partial \mathcal{U}(C_{t+1}^H, N_{t+1}^H)}{\partial C_{t+1}^H} \right) \right] \quad (4)$$

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

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

$$\frac{\partial \mathcal{U}(C_t^U, N_t^U)}{\partial C_t^U} \geq \beta \mathbb{E}_t^{BR} \left[\frac{v_{t+1} + \tilde{D}_{t+1}}{v_t} \frac{\partial \mathcal{U}(C_{t+1}^U, N_{t+1}^U)}{\partial C_{t+1}^U} \right]. \quad (6)$$

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

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

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 (5) 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] As shares cannot be transferred to the H state, equation (6) simply prices the shares. Thus, the bond Euler equation of unconstrained households (4) 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 New Keynesian firm side with sticky prices, as we detail below. All households consume the same aggregate basket of individual goods, $j \in [0, 1]$,

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

where $\epsilon > 1$ is the elasticity of substitution between the individual goods. Each firm faces demand

$$C_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} C_t,$$

where $P_t(j)/P_t$ denotes the individual price relative to the aggregate price index,

$$P_t^{1-\epsilon} = \int_0^1 P_t(j)^{1-\epsilon} dj,$$

and produces with the linear technology

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

The real marginal cost is given by W_t. We assume that the government pays a constant subsidy τ^S on revenues to induce marginal cost pricing in the steady state. 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

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

[8] Throughout the paper variables with a hat on top denote log-deviations from steady state.

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]

$$\widehat{i}_t = \phi \pi_t + \epsilon_t^{MP}, \tag{7}$$

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

$$B_{t+1}^U = 0$$

at all t.

2.2 Log-Linearized Model

We now focus on the log-linearized dynamics around the full-insurance, zero-liquidity steady state. In that case, we can write consumption of the hand-to-mouth households as

$$\widehat{c}_t^H = \chi \widehat{y}_t, \tag{8}$$

with

$$\chi \equiv 1 + \varphi \left(1 - \frac{\tau^D}{\lambda}\right) \tag{9}$$

[9]We study more general Taylor rules in Appendix A.

measuring the cyclicality of the H household's consumption (see appendix A.1). Auclert (2019) and Patterson (2019) document that households with higher MPCs tend to be more exposed to aggregate income fluctuations, which is the case when $\chi > 1$. In our model this reflects the assumption that hand-to-mouth households receive a smaller share of firm profits than unconstrained households.

Combining equation (8) with the goods market clearing condition yields

$$\widehat{c}_t^U = \frac{1 - \lambda\chi}{1 - \lambda}\widehat{y}_t, \tag{10}$$

which implies that consumption inequality is given by:

$$\widehat{c}_t^U - \widehat{c}_t^H = \frac{1 - \chi}{1 - \lambda}\widehat{y}_t. \tag{11}$$

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

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

$$\widehat{c}_t^U = s\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] + (1 - s)\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^H\right] - \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t^{BR}\pi_{t+1}\right). \tag{12}$$

For the case without type-switching, i.e., for $s = 1$, equation (12) 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. 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 (12). We show in Appendix C that our results go through with boundedly-rational real-rate expectations.

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

$$\pi_t = \kappa\widehat{y}_t, \tag{13}$$

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 (rational or behavioral). We discuss this case in Appendix C and show that a forward-looking Phillips Curve does not qualitatively affect our results.

We impose several assumptions in order to analytically characterize our main results as well as to generate analytical insights into how household heterogeneity and bounded rationality interact. In particular, we assume full insurance within types, exogenous type switching, a zero-liquidity equilibrium, no inequality in the steady state, a static Phillips Curve, and we linearize the economy around its steady state. We relax all these assumptions in Section 4 and show that our results presented in the following do not depend on these assumptions.

2.3 Bounded Rationality

We follow Gabaix (2020) and model bounded rationality in the form of cognitive discounting.[10] Let X_t be a random variable (or vector of variables) and let us define X_t^d as some default value the agent may have in mind and let $\tilde{X}_{t+1} \equiv X_{t+1} - X_t^d$ denote the deviation from this default value.[11] The behavioral agent's expectation about X_{t+1} is then defined as

$$\mathbb{E}_t^{BR}\left[X_{t+1}\right] = \mathbb{E}_t^{BR}\left[X_t^d + \tilde{X}_{t+1}\right] \equiv X_t^d + \bar{m}\mathbb{E}_t\left[\tilde{X}_{t+1}\right], \tag{14}$$

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 5.3.

While we present a way how to microfound $\bar{m}$ in Appendix A.9, 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

[10] 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 behavioral frictions.

[11] 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 5.3. Appendix A.8 derives our results following the approach in Gabaix (2020).

of common knowledge lead to observationally-equivalent expectations for the case in which X_t^d denotes the steady state.

Log-linearizing equation (14) around the steady state yields

$$\mathbb{E}_t^{BR}\left[\widehat{x}_{t+1}\right] = (1-\bar{m})\widehat{x}_t^d + \bar{m}\mathbb{E}_t\left[\widehat{x}_{t+1}\right] \tag{15}$$

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 F.1, we estimate $\bar{m}$ for different household groups based on their income and 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 that the empirical evidence points towards a $\bar{m}$ of 0.6 to 0.85.

3 Results

In this section, we derive the three-equation representation of the behavioral HANK model and show that the model is consistent with the discussed empirical facts. The model nests a wide spectrum of existing models—none of which can account for all the empirical facts simultaneously. We then illustrate how the behavioral HANK model leads to different policy implications than its rational counterpart.

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 (13), and a rule for monetary policy (equation (7)), 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 (8) with the consumption of unconstrained households (10) and their consumption Euler equation (12) (see appendix A for all the derivations).

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

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t \pi_{t+1}\right), \tag{16}$$

where

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

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

The second new coefficient in the behavioral HANK IS equation (16), 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 the two frictions. Given countercyclical income inequality, income risk is also countercyclical which manifests itself in $\delta > 1$. 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 (16) 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 $\bar{m} = 1$.

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 τ^D such that $\chi = 1.48$ 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, $\gamma = 1$, set $\beta = 0.99$, and 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, even when we vary certain parameters, we 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.

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

$$\widehat{y}_t = \frac{\psi_f + \psi_c \frac{\kappa}{\gamma}}{1 + \psi_c \phi \frac{\kappa}{\gamma}} \mathbb{E}_t \widehat{y}_{t+1} - \frac{\psi_c \frac{1}{\gamma}}{1 + \psi_c \phi \frac{\kappa}{\gamma}} \varepsilon_t^{MP}, \tag{17}$$

which we obtain by combining the IS equation (16) with the static Phillips Curve (13) and the Taylor rule (7).

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.[12] Such strong effects of future interest rate changes, however, seem puzzling and are not supported by the data (Del Negro et al. (2015), Miescu (2022), Roth et al. (2021)).

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 $\phi = 0$, as in Bilbiie (2021).[13]

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

$$\chi > 1, \tag{18}$$

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

[13]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 relaxes the condition to rule out the forward guidance puzzle.

and the forward guidance puzzle is ruled out if

$$\bar{m}\delta + \frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi}\kappa < 1. \tag{19}$$

The behavioral HANK model generates amplification of contemporaneous monetary policy with respect to RANK whenever $\chi > 1$, that is, when high-MPC households' consumption is relatively 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. To see the importance of GE or indirect effects, the following Lemma disentangles the direct and indirect effects (see appendix A.6).

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

$$\widehat{c}_t = [1 - \beta(1-\lambda\chi)]\,\widehat{y}_t - \frac{(1-\lambda)\beta}{\gamma}\widehat{r}_t + \beta\bar{m}\delta(1-\lambda\chi)\mathbb{E}_t\widehat{c}_{t+1}. \tag{20}$$

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

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

Given our calibration and assuming an AR(1) monetary policy shock with a persistence of 0.8, indirect effects account for about 77%, consistent with larger quantitative models as for example in Kaplan et al. (2018). For comparison, the representative agent model generates an indirect share of

$$\Xi^{GE} = \frac{1-\beta}{1-\beta\bar{m}\rho},$$

thus, about 3% in the behavioral RANK model and 5% in the rational RANK model.

Note, that in the case of an i.i.d. shock 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 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 (17) is smaller than 1. Given that we consider $\phi = 0$, this boils down to the condition stated in Proposition 2.

What determines whether condition (19) 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 ($\delta > 1$). Yet, as households are boundedly rational, they cognitively discount these effects taking place in the future. Importantly, unconstrained households cognitively discount both the 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 $\bar{m} < 0.93$ which is above the upper bounds for empirical estimates (see Section 2.3).

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), see also Werning (2015)). 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.

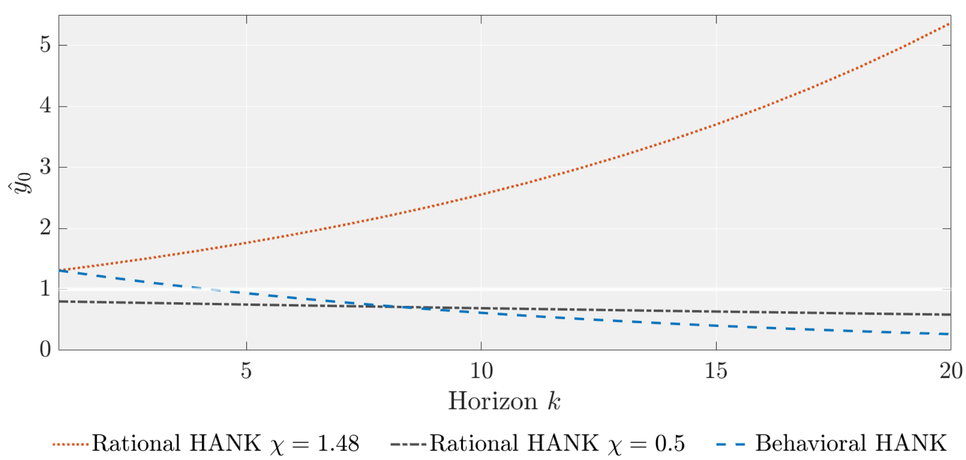
We graphically illustrate the Catch-22 of the rational model and the resolution of it in the behavioral HANK model in Figure 1. The figure shows on the vertical axis the response of contemporaneous output relative to the initial response in the RANK model with rational expectations for anticipated i.i.d. monetary policy shocks occurring at different times k on the horizontal axis.[14]

[14]Under fully-rigid prices (i.e., $\kappa = 0$) the RANK model would deliver a constant response for all k.

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)).

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).

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 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.

A direct consequence of Proposition 2 is that in the behavioral HANK model, highly

The same is true for two-agent NK models (TANK), i.e., tractable HANK models without type switching. Whether the constant response would lie above or below its RANK counterpart depends on $\chi \lesseqgtr 1$ in the same way the initial response depends on $\chi \lesseqgtr 1$.

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. As the persistence of the shocks approaches unity, an exogenous increase in the nominal interest rate becomes expansionary in the rational model. The behavioral HANK model, on the other hand, does not suffer from these paradoxical model predictions. We elaborate these points in more detail in Appendix C.2. A similar result applies to fiscal spending shocks, as we discuss below.

Determinacy in Behavioral HANK. According to the Taylor principle, monetary policy needs to respond sufficiently strongly to inflation in order to guarantee a determinate equilibrium. In the rational RANK model the Taylor principle is given by $\phi > 1$, where ϕ is the inflation-response coefficient in the Taylor rule (7). 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.[15]

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

$$\phi > \phi^* = 1 + \frac{\bar{m}\delta - 1}{\frac{\kappa}{\gamma}\frac{1-\lambda}{1-\lambda\chi}}. \tag{21}$$

We obtain Proposition 3 directly from the difference equation (17). For determinacy, we need that the coefficient in front of $\mathbb{E}_t \widehat{y}_{t+1}$ is smaller than 1 (the eigenvalues associated with any exogenous variables are assumed to be $\rho < 1$, which is stable). Solving this condition for ϕ yields Propositon 3. Appendix A.4 outlines the details and extends the result to more general Taylor rules.

To understand the condition in Proposition 3, consider first $\bar{m} = 1$ and, thus, focus solely on the role of household heterogeneity. With $\chi > 1$, it follows that $\phi^* > 1$ and, hence, the threshold is higher than the RANK Taylor principle states. This insufficiency of the Taylor principle in 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 impli-

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

cations 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 obtain $\phi^* = -3.07$. 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.[16]

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.[17]

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

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

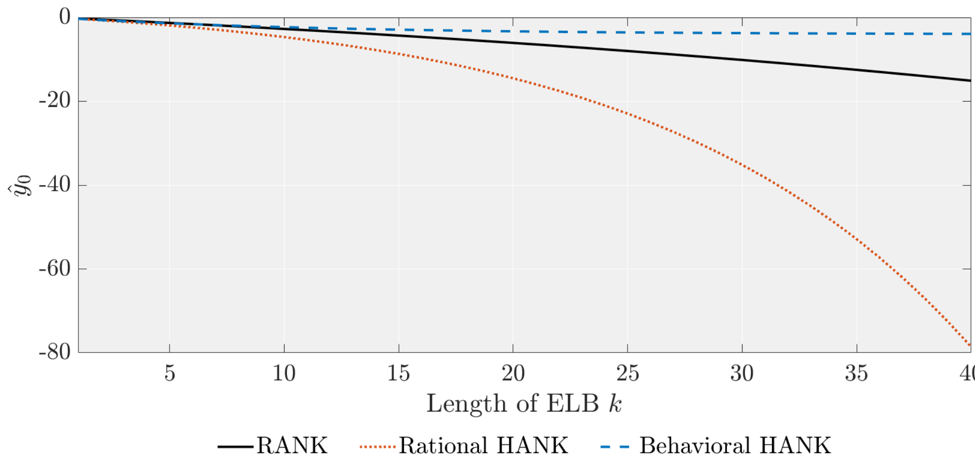
We assume that in period t the natural rate decreases to a value $\widetilde{r}^n$ that is sufficiently negative such that the natural rate in levels is below the ELB. The natural rate stays at $\widetilde{r}^n$ for $k \geq 0$ periods and after k periods the economy returns immediately back to steady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

$$\widehat{y}_t = -\frac{1}{\gamma}\psi_c \underbrace{\left(\widehat{i}_{ELB} - \widetilde{r}^n\right)}_{>0} \sum_{j=0}^{k}\left(\psi_f + \frac{\kappa}{\gamma}\psi_c\right)^j, \tag{22}$$

[16] 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.

[17] The intuition is directly related to our discussion about 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. 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.

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.

where the term $\left(\widehat{i}_{ELB} - \widetilde{r}^n\right) > 0$ captures the shortfall of the policy response due to the binding ELB. Under rational expectations, we have that $\delta > 1$ 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 $\psi_f = \psi_c = 1$. In the behavioral HANK model, however, this is not the case. As long as $\psi_f + \frac{\kappa}{\gamma}\psi_c < 1$ the output response in t is bounded even as $k \to \infty$. It follows that $\bar{m} < 0.93$ 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 model 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 (back-solid line) and even more so in the rational HANK model (orange-dotted line): an ELB that is expected to bind for 40 quarters would decrease today's output in the rational RANK 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%, as illustrated by the blue-dashed line.

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 estimated empirically. 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.

The behavioral HANK IS equation with government spending is given by:

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

where $\zeta \equiv \frac{\varphi}{\gamma\left(1+\frac{\varphi}{\gamma}\right)}$ (see appendix A.5). 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

$$\nu \equiv \frac{\psi_f + \kappa\frac{1}{\gamma}\psi_c}{1+\frac{1}{\gamma}\psi_c\phi\kappa}. \tag{23}$$

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.[18] 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

[18] 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.

is a useful benchmark as in this case the consumption response in RANK is 0 (see Bilbiie (2011) and Woodford (2011)).[19]

From Proposition 4, we derive the constant-real-rate multiplier in the behavioral HANK model:

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

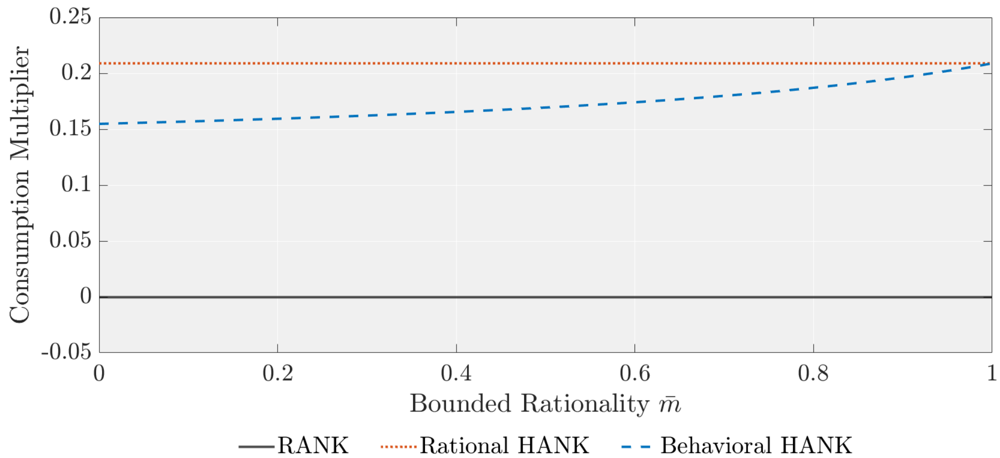
As $\chi > 1$ the fiscal multiplier is bounded from below by 0 irrespective of the persistence ρ_g. In other words the constant-real-rate multiplier in the behavioral HANK model is strictly positive, regardless of the dampening of bounded rationality on the fiscal multiplier in the case of persistent spending. With $\chi > 1$ the high MPC households benefit disproportionately more 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 compares 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 of $\bar{m} = 0$, 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.

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 fiscal multipliers larger than 0.

[19] 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.

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.

Comparison to Nested Models. The behavioral HANK model nests three classes of models in the literature: the representative-agent rational expectations (RANK) model for $\lambda = 0$ and $\bar{m} = 1$ (see Galí (2015), Woodford (2003)), representative agent models without FIRE for $\lambda = 0$ and $\bar{m} \in (0, 1)$ as, for example, in Gabaix (2019), Angeletos and Lian (2018), and Woodford (2019); and TANK and tractable HANK models as e.g. in Bilbiie (2008), Bilbiie (2021), McKay et al. (2017), or Debortoli and Galí (2018) for $\bar{m} = 1$. In contrast to these classes of models, the behavioral HANK model combines the indirect general equilibrium amplification of monetary and fiscal policy with a resolution of the forward guidance puzzle and stability at the ELB. In representative agent models monetary policy mainly works through direct intertemporal substitution channels and cannot have $\psi_c \neq 1$, rational HANK models on the other hand do not feature $\psi_f < 1$ and $\psi_c > 1$ simultaneously as discussed in Section 3.2 and in Bilbiie (2021).

4 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 do not depend on the simplifying assumptions we imposed to keep the model tractable. To this end, we replace the household set-up described in Section 2 by the typical incomplete markets set-up as in Bewley (1986), Huggett (1993), and Aiyagari (1994) which is standard in the quantitative HANK literature. There is a continuum of ex-ante identical households all subject to idiosyncratic productivity risk, incomplete markets, and exogenous borrowing constraints.

Households self-insure against their idiosyncratic risk by accumulating government bonds. Bonds are in positive net supply as the fiscal authority issues a constant amount of real bonds, B^G. To finance its interest payments, the fiscal authority collects tax payments from households. Given these assumptions, households differ ex-post in their current productivity level, e, and their wealth B. The households' utility function is the same as in the tractable model (equation (1)).

Household i faces the following budget and borrowing constraints:

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

where $\underline{B}$ denotes the (exogenous) borrowing limit. As in McKay et al. (2016), households pay taxes conditional on their productivity, $\tau_t(e)$, and, in particular, we also assume that only the most productive households pay taxes. Households receive a share of the dividends, $D_t d(e)$ also conditional on their productivity. 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 is 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 McKay et al. (2016) who assume that every household receives the same share of the dividends which leads to procyclical inequality.

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 such as wages) 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. We assume that households have perfect foresight about the path of the real interest rate.

Hence, the Euler equation of household i is given by

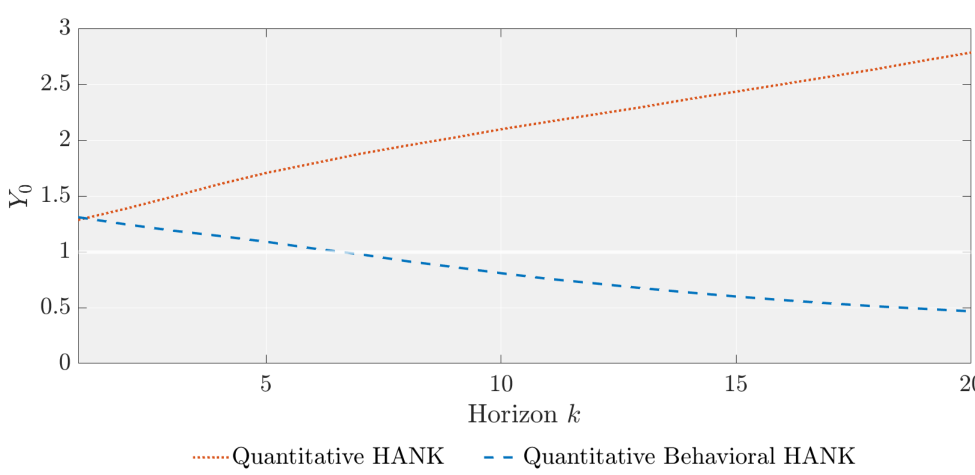
$$\begin{aligned} C_{i,t}^{-\gamma} &\geq \beta R_t \mathbb{E}_t^{BR}\left[C_{i,t+1}^{-\gamma}\right] \\ &= \beta R_t \mathbb{E}_t^{BR}\left[C_i^{-\gamma} + \left(C_{i,t+1}^{-\gamma} - C_i^{-\gamma}\right)\right] \\ &= \beta R_t \left[C_i^{-\gamma} + \bar{m}\mathbb{E}_t\left(C_{i,t+1}^{-\gamma} - C_i^{-\gamma}\right)\right], \end{aligned} \tag{24}$$

where $C_i^{-\gamma}$ denotes the marginal utility of household i (which depends on the household's individual states B and e) in the stationary equilibrium, i.e., when all aggregate variables and prices are constant. The Euler equation (24) holds with equality for non-constrained households, while it holds with strict inequality for households that are at 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 the details and the parameterization to Appendix F.

4.1 Monetary Policy

We now consider the same two monetary policy experiments as in the tractable model. First, how does the economy respond to an i.i.d. expansionary monetary policy shock compared to RANK and second, how do these effects change as the shock is announced today to take place at some point in the future? In particular, we assume that 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 such that that 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).

Figure 4: 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 (normalized to 1).

Figure 4 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 represents the response in the complete-markets model, i.e., in the rational RANK model.[20] The constant response in RANK is a consequence of the assumption that forward guidance is implemented through changes in the real rate.

The blue-dashed line shows the results for the quantitative behavioral HANK model. We see that contemporaneous monetary policy has stronger effects than in RANK. The intuition is the same as in the tractable model: as households with higher MPCs tend to be more exposed to the business cycle, monetary policy is amplified through indirect effects, Keynesian type general equilibrium effects. Turning again to an AR(1)-process with persistence of 0.8, we find that indirect effects account for 73% of the total effect in the quantitative behavioral HANK, consistent with other (larger) HANK models as well as our tractable behavioral HANK model. At the same time, the behavioral HANK model does not suffer from the forward guidance puzzle, as shown by the decline in the blue-dashed line. Interest rate changes announced to take place in the future have relatively weak effects on contemporaneous output and the effects decrease with the horizon. Overall, figure 4 shows that also in the quantitative behavioral HANK model, contemporaneous monetary policy is amplified through indirect, whereas announced future policy changes have relatively weaker effects on today's economy.

The orange-dotted line shows that this is not the case in the rational HANK model. Contemporaneous monetary policy is as strong as in the behavioral model, but with rational expectations the amplification through indirect effects extends intertemporally and results in an aggravation of the forward guidance puzzle. Indeed, we see 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.[21]

Stability at the ELB and positive fiscal multipliers. The quantitative behavioral HANK model is also consistent with the tractable model when considering the stability at the effective lower bound as well as the implied consumption responses to a fiscal spending shock. In particular, we employ a transitory shock to the discount factor which pushes the economy to the ELB for twelve periods, in the behavioral and the rational model. After that the shock jumps back to its steady state value. Consistent with the tractable model, the recession in the rational model is substantially more severe. While output drops only by 5.8% in the behavioral model, it drops by 9.8% in the rational model (see Appendix F for

[20] 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.

[21] We discuss other proposed resolution of the forward-guidance puzzle in HANK models in Appendix F and contrast those to our proposed resolution through bounded rationality.

details).

Turning to fiscal policy, we also confirm that the quantitative behavioral HANK model generates positive consumption multipliers under a constant real rate. To this end, we run the same exercise—a temporary increase in government consumption financed by lump-sum transfers—as in Section 3.3. To such a fiscal policy shock, private consumption increases independent of the persistence of the fiscal shock (see Appendix F for details). Overall, we conclude that our main insights of the tractable behavioral HANK model carry over to a quantitative behavioral HANK model.

4.2 Heterogeneous Cognitive Discounting

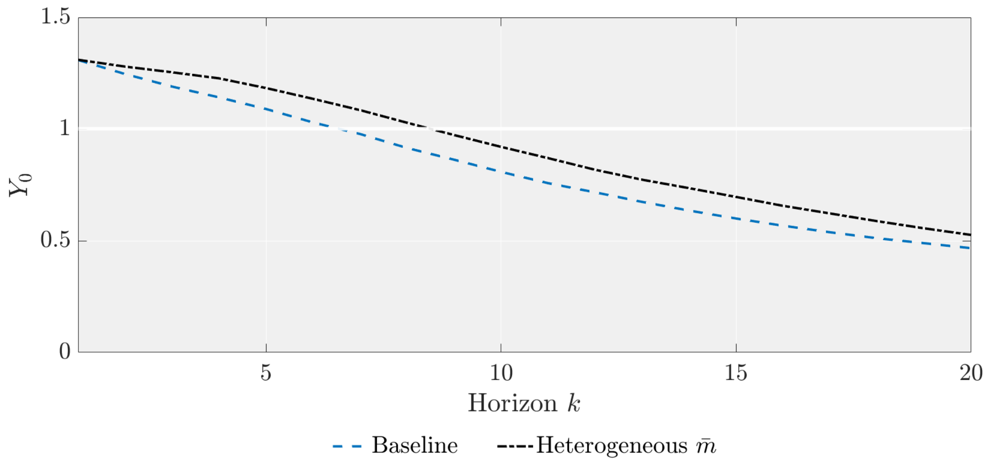
So far, we have assumed that all households exhibit the same degree of rationality. In reality, however, there might be heterogeneity with respect to the degree of cognitive discounting. Indeed, as we show in Appendix F.1, while cognitive discounting is found across all income groups, the data suggests that higher income households deviate somewhat less from rational expectations.

We incorporate this result into our theoretical framework by allowing for different degrees of rationality. To capture the positive correlation between households income and the degree of rationality, we assume that a household's rationality is a function of her productivity level e: $\bar{m}(e = e_1) = 0.8$, $\bar{m}(e = e_2) = 0.85$ and $\bar{m}(e = e_3) = 0.9$.[22] This parameterization serves three purposes: first, in line with the data, the lowest-productivity households exhibit the largest deviation from rational expectations and the degree of rationality increases monotonically with productivity. Second, the average degree of bounded rationality remains 0.85 such that we can isolate the effect of heterogeneous degrees of bounded rationality. And third, this is a very conservative parameterization—both in terms of the degree of heterogeneity and in the level of rationality—compared to the results in the data which points more towards lower level of rationality across all households and less dispersion.

Figure 5 compares the model with heterogeneous degrees of bounded rationality to our baseline quantitative behavioral HANK model for the same monetary policy experiments as above. It shows that the effectiveness of monetary policy at different horizons changes when households are heterogeneous in their degree of cognitive discounting. The effect of a contemporaneous monetary policy shock is practically identical across the two scenarios consistent with the insight that amplification of a contemporaneous monetary policy shock is barely affected by the degree of rationality. At longer horizons, however, monetary policy is more effective in the economy in which households differ in their degrees of rationality.

[22]We discuss other scenarios in Appendix F.1.

Figure 5: Heterogeneous $\bar{m}$ and Monetary Policy



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

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

4.3 Policy Implications: Stabilizing Inflation

Many advanced economies have recently experienced a dramatic surge in inflation and at least part of this is attributed to cost pressures coming from the supply side. We now use the quantitative behavioral HANK model to revisit how monetary policy has to be implemented after such an inflationary cost-push shock. In particular, we show that according to the behavioral HANK model, central banks need to hike interest rates much more strongly if they want to stabilize inflation after an inflationary supply-side shock. Increasing the nominal interest rate more strongly, however, poses a new policy trade-off after supply side shocks as this comes at the cost of pushing up the government debt level substantially.

We now assume that the desired mark-up of firms, μ_t follows the following AR(1)-process:

$$\mu_t = (1 - \rho_\mu)\bar{\mu} + \rho_\mu \mu_{t-1} + \varepsilon_t^\mu$$

where ε_t^μ is an i.i.d. shock, $\bar{\mu}$ the steady-state level of the desired markup and ρ_μ the persistence of the shock process which we set to $\rho_\mu = 0.9$. Each firm can adjust its price with probability 0.15 in a given quarter and we assume that firms have rational expectations to fully focus on the role of bounded rationality on the household side.

Monetary policy follows a strict inflation-targeting rule and implements a zero inflation rate in all periods. Fiscal policy is such that government debt is time-varying and total tax payments, T_t, follow a standard debt feedback rule:

$$T_t - \bar{T} = \vartheta \frac{B_{t+1} - \bar{B}}{\bar{Y}}, \tag{25}$$

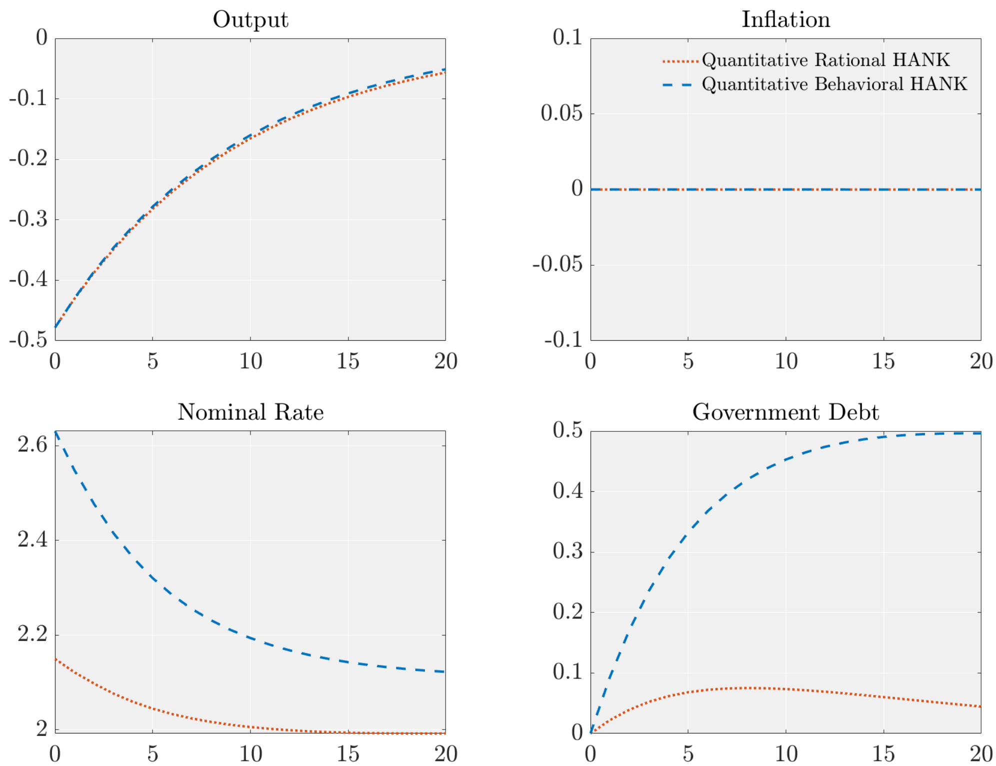
where we set $\vartheta = 0.05$.

Figure (6) shows the impulse-response functions of output, inflation, nominal interest rates and government debt following an inflationary cost-push shock. The blue-dashed lines show the responses in the behavioral HANK model and the orange-dotted lines for the rational HANK model. In both cases, monetary policy fully stabilizes inflation by assumption. Output drops in both cases, with the responses being practically identical across the two models. The required response of the nominal interest rate, however, differs substantially across the two models. In the behavioral HANK the monetary authority increases the nominal rate much more strongly and more persistently. The reason for this strong response is that households cognitively discount future (expected) interest rate hikes making them less effective for stabilizing the economy today. Thus, in order to achieve the same stabilization outcome in every period, the interest rate needs to increase by more.

Increasing the interest rates increases the cost of debt for the government which it finances in short-run by issuing more debt. The bottom-right panel in Figure 6 shows that government debt in the behavioral model increases about five times as much as in the rational model. As we show in Figure 18 in Appendix F, this is even more pronounced in a "post-Covid" world, in which initial debt levels are already high. Stabilizing inflation after a cost-push shock increases government debt (as a share of annual GDP) by 1.0 percentage points whereas it would only increase it by 0.2 percentage points in the rational model.

If the monetary authority follows a simple Taylor rule instead of fully stabilizing inflation, the nominal rate increases again more strongly in the behavioral model, even though the differences are somewhat smaller. Inflation, however, now also increases by more in the behavioral model and government debt increases more strongly as well. We discuss this case in more detail in Appendix F.

Figure 6: Inflationary cost-push shock



Note: This figure shows the impulse responses after a cost-push shock that increases the desired mark-up by 2% in the inflation-stabilizing monetary policy regime. Output is shown as percentage deviations from steady state, inflation and nominal interest rate as annualized percentage points and government debt level as percentage point deviations in debt-per annual GDP level.

5 Model Extensions

We now extend our baseline tractable model along three dimensions. First, we allow for positive savings which enables us to derive the so-called iMPCs and show how they depend on bounded rationality, heterogeneity and the interaction of the two. Second, 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. Third, we derive an equivalence result between HANK models with bounded rationality and HANK models with incomplete information and learning.

5.1 Intertemporal MPCs

The HANK literature shows that the intertemporal marginal propensities to consume are a key statistic for conducting policy analysis (see, e.g., Auclert et al. (2018), Auclert et al. (2020), and Kaplan and Violante (2020)).[23] 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 (see Appendix D).

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 period 0, are given by*

$$MPC_0 \equiv \frac{d\widehat{c}_0}{d\widetilde{y}_0} = 1 - \frac{1-\lambda\chi}{s\bar{m}}\mu_2^{-1}$$

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

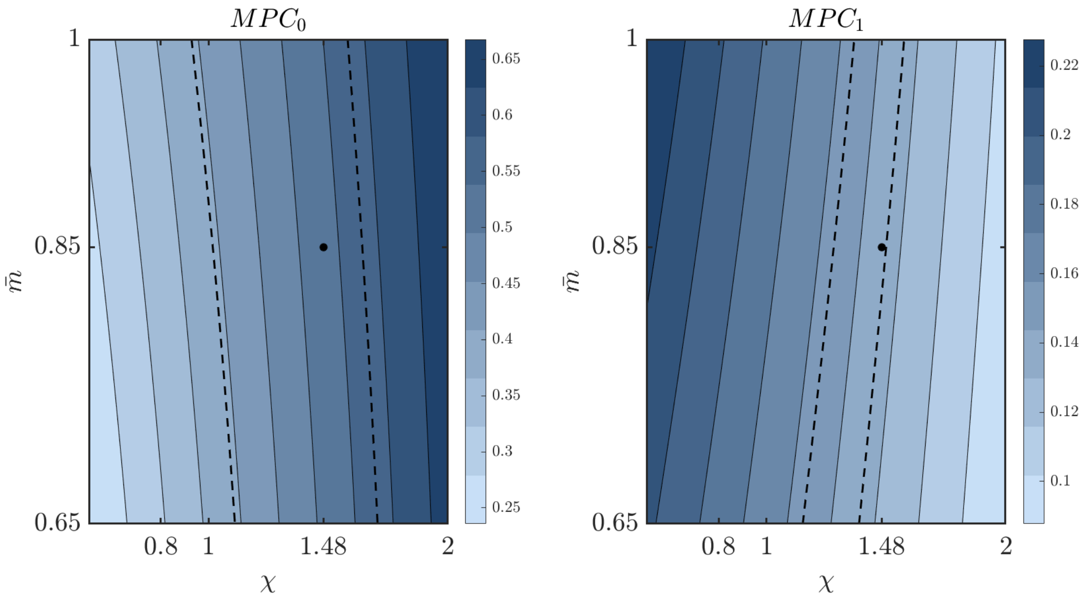
where the parameters μ_1 and μ_2 depend on the underlying parameters, including $\bar{m}$ and χ and are explicitly spelled out in Appendix D.

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. Figure 7 graphically depicts how the interplay of bounded rationality $\bar{m}$ and household heterogeneity χ determines the size of the aggregate iMPCs. 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. 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

[23] See, e.g., Lian (2021) or Boutros (2022) for MPC analyses in models deviating from FIRE.

Figure 7: 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.

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 the initially-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—due to 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 their 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 initially-unconstrained households remains unconstrained which pushes consumption upwards 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. This follows directly from our discussion about the role of χ and $\bar{m}$: 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: 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%.

5.2 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 the data (see, e.g., Auclert et al. (2020) or Broer et al. (2020)). 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) friction, 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:

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

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

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

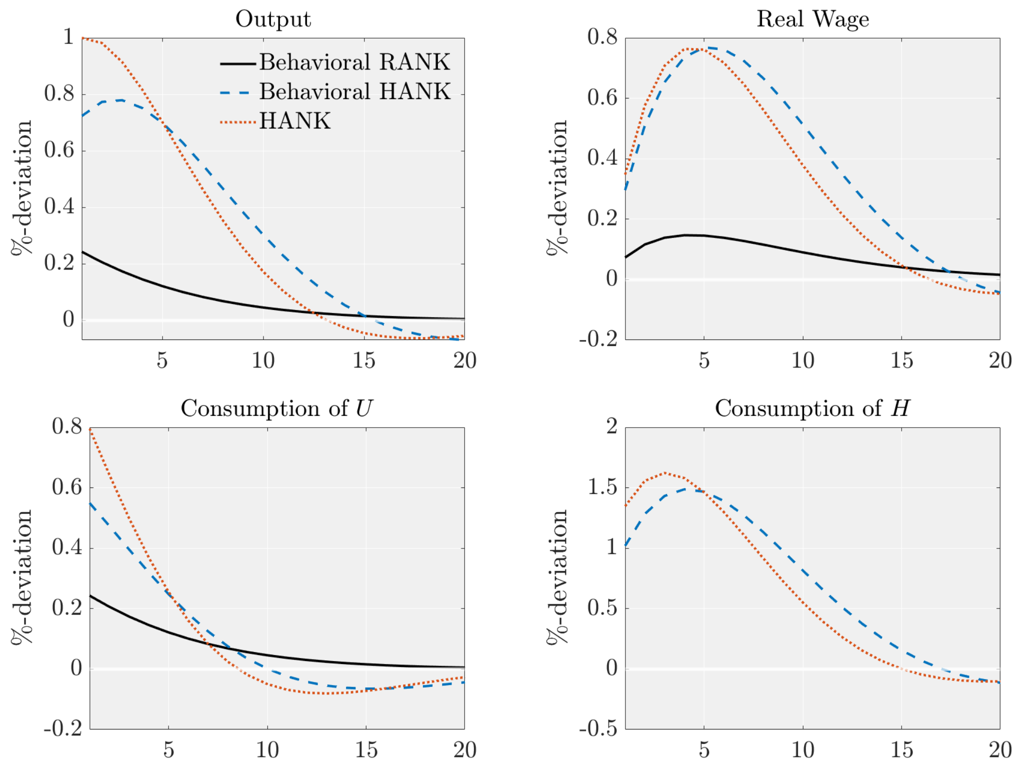
where $\widehat{mc}_t$ denotes the time-varying price markup. The rest of the model is as above. We relegate the details and the parameterization to Appendix E.

Hump-shaped responses to monetary policy shocks. 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 with sticky wages (blue-dashed lines). Importantly, the figure shows that the output response to a monetary policy shock is hump-shaped in the behavioral HANK model but neither in its representative agent (black-solid lines) nor in its rational counterpart (orange-dotted lines).

Where does the hump-shaped response come from? 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). 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 will respond in the future 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 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 house-

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, the behavioral RANK model 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.

holds 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 hand-to-mouth 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 5.1 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).[24]

[24]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

Forecast-errors dynamics. We now show that the sticky-wage behavioral HANK model generates dynamic forecast errors as observed in survey data. In particular, households' expectations initially underreact followed by delayed overreaction as recently documented empirically in Angeletos et al. (2021) for unemployment and inflation and in Adam et al. (2022) for housing prices.[25] Consistent with the empirical exercise in Angeletos et al. (2021), we focus on three-quarter ahead forecasts. For a variable $\widehat{x}$, the three-period ahead forecast error is defined as

$$FE^{\widehat{x}}_{t+h+3|t+h} \equiv \widehat{x}_{t+h+3} - \bar{m}^3 \mathbb{E}_{t+h}\left[\widehat{x}_{t+h+3}\right].$$

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 in which the expectations start to change which in this case corresponds to the fourth period after the shock. For completeness, the orange-dashed lines at zero show that under rational expectations, i.e., $\bar{m} = 1$, forecast errors are equal to 0. In the behavioral HANK 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.

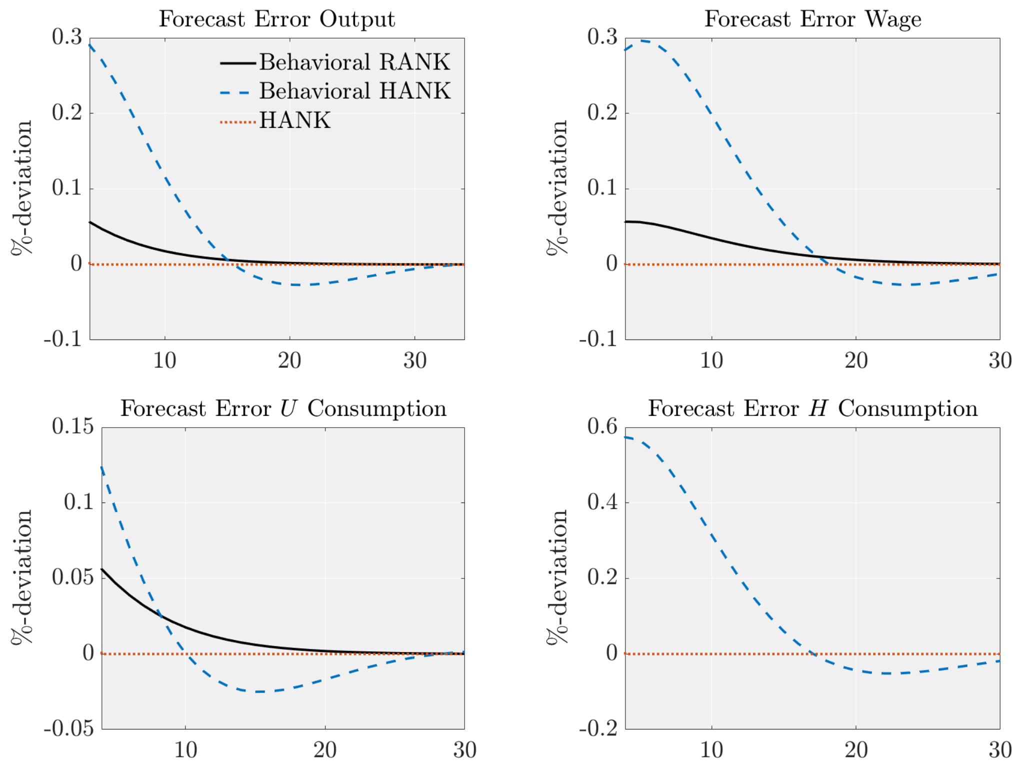
After about 10-15 quarters, however, forecast errors turn negative. Put differently, the behavioral agents' expectations show patterns of delayed overreaction. In contrast to Angeletos et al. (2021) or Adam et al. (2022), the behavioral HANK model with sticky wages generates these dynamic patterns of forecast errors even though the behavioral agents' expectations are purely forward looking.

Where does the delayed overshooting come from? As figure 8 shows, output falls below its steady-state level after some periods in the HANK models. The reason is that with sticky wages, wages increase very persistently. In HANK, this makes the consumption of the H households very persistent which, ceteris paribus, makes the increase in aggregate demand more persistent. Monetary policy reacts to this by increasing the nominal interest rate more strongly and more persistently. Due to inertia in the Taylor rule, however, the interest rate stays high even as aggregate demand returns to its steady state level, generating a mild recession after about 15 quarters (consistent with larger HANK models, see, for example, Auclert et al. (2020)). The behavioral agents then not only underestimate the boom after the monetary policy shock in the short-run, but also underestimate the mild recession in the

this way, the iMPCs presented in Figure 7 are completely unaltered.

[25] 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)).

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.

medium-run, which causes the delayed overshooting in their expectations.

Note that the behavioral RANK model (black solid lines) does not generate these delayed overreactions. Only when allowing for both—household heterogeneity and bounded rationality—the model is able to generate hump-shaped responses of macroeconomic aggregates and forecast error dynamics that are consistent with recent evidence from household survey expectations.

5.3 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 which induces a backward-looking component in the expectations as well as in the IS equation:

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

$$X_t^d = X_{t-1}. \tag{26}$$

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

$$\mathbb{E}_t^{BR}\left[X_{t+1}\right] = (1-\bar{m})X_{t-1} + \bar{m}\mathbb{E}_t\left[X_{t+1}\right] \tag{27}$$

and the behavioral HANK IS equation is then given by

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \frac{1}{\gamma}\left(\widehat{i}_t - \mathbb{E}_t \pi_{t+1}\right) + (1-\bar{m})\delta \widehat{y}_{t-1}. \tag{28}$$

Proposition 6 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.[26] 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 complement their findings by showing how we can generate the equivalent outcome based on a *behavioral* relaxation of FIRE.

6 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 these two frictions enables the model to be reconciled with the data. The presence of both frictions is thus crucial to arrive at our results. The behavioral HANK model is analytically tractable and nests a wide range of existing models—none of which can account for all the empirical patterns. The main insights from the tractable model carry over to a quantitative version of our behavioral HANK model including a set-up with heterogeneous degrees of bounded rationality. The behavioral HANK

[26] 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.

model predicts that central banks that want to stabilize inflation after an inflationary supply shock need to hike the nominal interest rate much more strongly and for longer than under rational expectations. Hiking interest rates, however, leads to a more pronounced increase in public debt, especially when the initial debt levels are already high. 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," .

Aiyagari, S. R. (1994): "Uninsured idiosyncratic risk and aggregate saving," *The Quarterly Journal of Economics*, 109, 659–684.

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

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.

Bewley, T. (1986): "Stationary monetary equilibrium with a continuum of independently fluctuating consumers," *Contributions to mathematical economics in honor of Gérard Debreu*, 79.

Bhandari, A., J. Borovička, and P. Ho (2019): "Survey data and subjective beliefs in business cycle models," *Available at SSRN 2763942*.

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.

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

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.

Huggett, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of economic Dynamics and Control*, 17, 953–969.

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.

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

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

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

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.

Appendix For Online Publication

A Model Details and Derivations

A.1 Derivation of χ

In Section 2, we stated that

$$\widehat{c}_t^H = \chi \widehat{y}_t, \tag{29}$$

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 (29) from the H-households' budget constraint, optimality conditions and market clearing.

The labor-leisure condition of the H households is given by

$$(N_t^H)^{\varphi} = W_t (C_t^H)^{-\gamma}, \tag{30}$$

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

$$\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H = \varphi \widehat{n}_t^U + \gamma \widehat{c}_t^U \tag{31}$$

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 as (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 (31) then gives

$$\varphi\widehat{n}_t^H + \gamma\widehat{c}_t^H = (\varphi + \gamma)\widehat{y}_t. \tag{32}$$

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

$$C_t^H = W_t N_t^H + \frac{\tau^D}{\lambda} D_t, \tag{33}$$

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

$$\widehat{c}_t^H = \widehat{w}_t + \widehat{n}_t^H + \frac{\tau^D}{\lambda}\widehat{d}_t, \tag{34}$$

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

$$\widehat{c}_t^H = \left(\varphi\widehat{n}_t^H + \gamma\widehat{c}_t^H\right)\left(1 - \frac{\tau^D}{\lambda}\right) + \widehat{n}_t^H. \tag{35}$$

Using (32) to solve for $\widehat{n}_t^H$ and plugging it into (35) 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 (8) and (10) with the bounded-rationality setup in equation (15) 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 (10) into the Euler equation of unconstrained households (12) 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 $\mathbb{E}_t[\widehat{y}_{t+1}]$ terms and dividing by $\frac{1-\lambda\chi}{1-\lambda}$ yields the following coefficient in front of $\mathbb{E}_t[\widehat{y}_{t+1}]$:

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

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 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 (13), 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.4 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 (16), 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

$$\phi > \phi^* = 1 + \frac{\bar{m}\delta - 1}{\frac{\kappa}{\gamma}\frac{1-\lambda}{1-\lambda\chi}}, \tag{36}$$

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 (36) 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

$$\widehat{y}_t = \frac{\omega_f + \frac{\kappa}{\gamma}\omega_c \bar{m}^r}{1 + \frac{\omega_c}{\gamma}\left(\kappa\phi_\pi + \phi_y\right)} \mathbb{E}_t \widehat{y}_{t+1}. \tag{37}$$

From equation (37), it follows that we need

$$\phi_\pi > \bar{m}^r - \phi_y + \frac{\omega_f - 1}{\omega_c \frac{\kappa}{\gamma}} = \bar{m}^r - \phi_y + \frac{\bar{m}\delta - 1}{\frac{1-\lambda}{1-\chi\lambda}\frac{\kappa}{\gamma}} \tag{38}$$

for the model to feature a determinate, locally unique equilibrium. Condition (38) shows that both, $\bar{m}^r < 1$ and $\phi_y > 0$, weaken the condition in Proposition 3. Put differently, bounded rationality with respect to the real rate or a Taylor rule that responds to changes in output, both relax the condition on ϕ_π to yield determinacy.

A.5 IS Curve with Government Spending

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

$$\widehat{c}_t^H = \widehat{w}_t + \widehat{n}_t^H + \frac{\tau^D}{\lambda}\widehat{d}_t - g_t, \tag{39}$$

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

$$\widehat{y}_t = \lambda \widehat{c}_t^H + (1-\lambda)\widehat{c}_t^U + g_t. \tag{40}$$

Plugging this and the labor market clearing condition into (31), yields:

$$\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H = (\varphi + \gamma)\widehat{y}_t - \gamma g_t. \tag{41}$$

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

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

and using (41) yields:

$$\widehat{c}_t^H = \left(\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H\right)\left(1 - \frac{\tau^D}{\lambda}\right) + \widehat{n}_t^H - g_t. \tag{43}$$

Finally, consumption of H is given by:

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

which is

$$\widehat{c}_t^H = \chi\left(\widehat{y}_t - g_t\right) + \left[\frac{\chi - 1}{1 + \frac{\gamma}{\varphi}}\right] g_t. \tag{45}$$

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

$$\widehat{c}_t^U = \frac{1 - \lambda\chi}{1 - \lambda}\left(\widehat{y}_t - g_t\right) - \frac{\lambda}{1 - \lambda}\frac{\chi - 1}{1 + \frac{\gamma}{\varphi}} g_t. \tag{46}$$

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 $\widehat{c}_t = (\widehat{y}_t - g_t)$.

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

which can be re-written as (using similar derivations as in Appendix A.2)

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

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

A.6 Derivation of Lemma 1

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

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

and for i periods ahead:

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

Furthermore, we have:

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

and the stochastic discount factor has the property

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

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

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

or put differently

$$\frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t,t+i}^U + \bar{m}^i\mathbb{E}_t\widehat{c}_{t+i}^U = \widehat{c}_t^U + (1-s)\mathbb{E}_t\sum_{k=1}^{i}\bar{m}^k\left(\widehat{c}_{t+k}^U - \widehat{c}_{t+k}^H\right). \tag{47}$$

The (linearized) budget constraint can be written as

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

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

Now, focus on the left-hand side and notice that the sum $\mathbb{E}_t\sum_{i=0}^{\infty}(\beta\bar{m})^i\,\widehat{c}_{t+i}^U$ cancels with the $\bar{m}^i\mathbb{E}_t\widehat{c}_{t+i}^U$ terms in equation (47) when summing them up. The left-hand side of the budget constraint can thus be written as

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

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

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

and similarly,

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

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

Combining these results, we see that

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

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

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

or written recursively

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

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

$$\widehat{c}_t = [1 - \beta(1-\lambda\chi)] \widehat{y}_t - \frac{(1-\lambda)\beta}{\gamma} \widehat{r}_t + \beta \bar{m} \delta (1-\lambda\chi) \mathbb{E}_t \widehat{c}_{t+1}, \tag{48}$$

as stated in the main text.

To obtain the share of indirect effects, note that the model does not feature any endogenous state variables and hence, endogenous variables inherit the persistence of the exogenous variables, ρ. Thus, $\mathbb{E}_t \widehat{c}_{t+1} = \rho \widehat{c}_t$. Plugging this into the consumption function (48), we get

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

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

A.7 Derivation of Proposition 6

To prove Proposition 6, we start from the Euler equation (12). Plugging in for $\widehat{c}_t^U$, $\widehat{c}_{t+1}^U$ and $\widehat{c}_{t+1}^H$ from equations (8) and (10), we get

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

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 6.

A.8 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 (12) 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.9 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^j_{t+1}, given by

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

where X'_{t+1} is an i.i.d. draw from the unconditional distribution of X_{t+1}, which has an unconditional mean of zero. In words, with probability p the agent j receives perfectly precise information 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^j_{t+1}, is given by

$$X^e_{t+1} \equiv \mathbb{E}\left[X_{t+1} | S_{t+1} = s^j_{t+1}\right] = p \cdot s^j_{t+1}.$$ [27]

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

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

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

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

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

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

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

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

B 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.

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)
β	0.99	Bilbiie (2020)
	Behavioral Parameter	
$\bar{m}$	0.85	Gabaix (2020)

Table 1: Baseline calibration.

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 7, 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 5.3, 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 $\bar{m}$. We provide direct evidence on $\bar{m}$ for households (of different income groups) in Section 4. We find that households are somewhat less rational than professional forecasters.

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

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

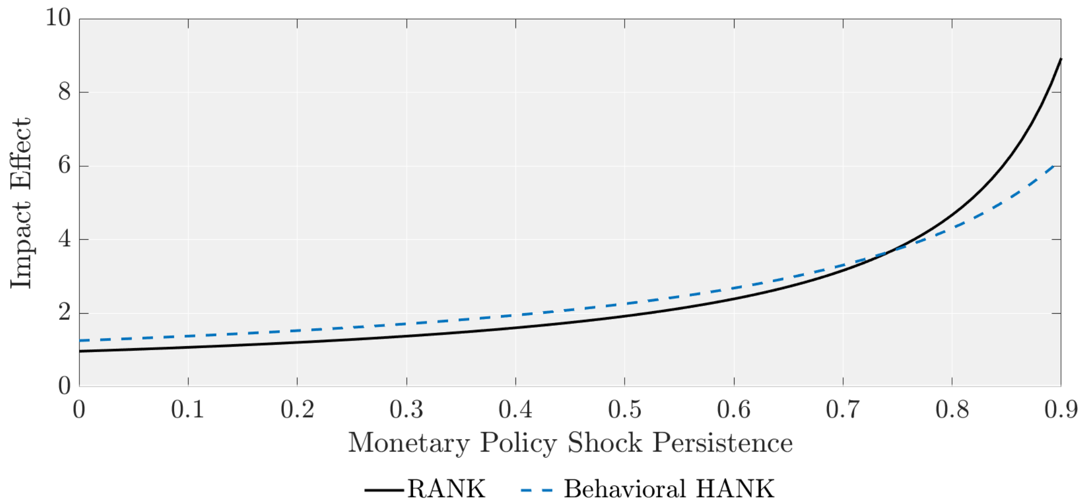
From a qualitative perspective, the whole analysis in the paper could be carried out with $\tilde{\delta}$ instead of δ. Quantitatively the differences are small as well. For example, if we set $\Omega = 1.5$, we get $\tilde{\delta} = 1.074$ instead of $\delta = 1.051$. Thus, we need $\bar{m} < 0.91$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

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

$$\widehat{y}_t = -\frac{\frac{\psi_c}{\gamma}}{1 + \frac{\psi_c}{\gamma}\phi\kappa - \left(\psi_f + \psi_c \frac{\kappa}{\gamma}\right)\rho} \varepsilon_t^{MP}. \tag{60}$$

Given our baseline calibration and a Taylor coefficient of $\phi = 1.5$, the rational model would produce these paradoxical findings for $\rho > 0.97$. The behavioral HANK model, on the other hand, does not suffer from this as the denominator is always positive, even when $\phi = 0$ and $\rho = 1$.

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:

$$\begin{pmatrix} \mathbb{E}_t \pi_{t+1} \\ \mathbb{E}_t \widehat{y}_{t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{\beta M^f} & -\frac{\kappa}{\beta M^f} \\ \frac{\psi_c}{\gamma\psi_f}\left(\phi - \frac{\bar{m}}{\beta M^f}\right) & \frac{1}{\psi_f}\left(1 + \frac{\psi_c \bar{m}\kappa}{\gamma\beta M^f}\right) \end{pmatrix}}_{\equiv A} \begin{pmatrix} \pi_t \\ \widehat{y}_t \end{pmatrix}. \tag{61}$$

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 5.1. Defining Y_t^j as type j's disposable income, we can write the households' budget constraints as

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

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

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

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

where b_t denotes real bonds in shares of steady state output. Aggregating (62) and (63) delivers

$$
\widehat{c}_t = \tilde{y}_t + \beta^{-1} b_t - b_{t+1}, \tag{64}
$$

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

By plugging equations (62) and (63) into the Euler equation of unconstrained households (12), we can derive the dynamics of liquid assets b_t (ignoring changes in the real rate as this is a partial equilibrium exercise):

$$
\begin{aligned}
&\mathbb{E}_t b_{t+2} - b_{t+1} \left[\frac{1}{s\bar{m}} + \beta^{-1} s + \frac{(1-s)^2 \beta^{-1} (1-\lambda)}{s\lambda} \right] + \frac{\beta^{-1}}{\bar{m}} b_t = \\
&(1-\lambda) \mathbb{E}_t \widehat{y}_{t+1}^U + \frac{1-s}{s} (1-\lambda) \mathbb{E}_t \widehat{y}_{t+1}^H - \frac{1-\lambda}{s\bar{m}} \widehat{y}_t^U.
\end{aligned} \tag{65}
$$

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

$$
(F - \mu_1)(F - \mu_2) \mathbb{E}_t b_t = -\mathbb{E}_t \widehat{z}_t, \tag{66}
$$

where μ_1 and μ_2 denote the roots of the characteristic equation

$$
\mathbb{E}_t b_{t+2} - \phi_1 b_{t+1} - \phi_2 b_t = 0, \tag{67}
$$

where

$$\phi_1 \equiv \left[\frac{1}{s\bar{m}} + \beta^{-1}s + \frac{(1-s)^2\beta^{-1}(1-\lambda)}{s\lambda}\right] \tag{68}$$

and

$$\phi_2 \equiv -\frac{\beta^{-1}}{\bar{m}}. \tag{69}$$

Thus, the roots are given by

$$\mu_{1,2} = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}. \tag{70}$$

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

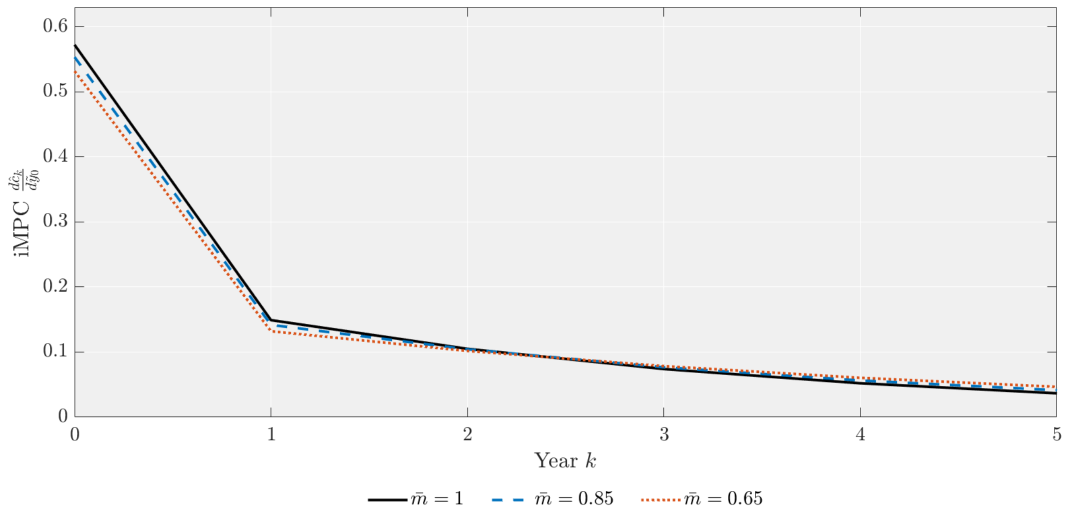
$$b_{t+1} = \mu_1 b_t + \frac{1-\lambda\chi}{s}\sum_{l=0}^{\infty}\mu_2^{-(l+1)}\mathbb{E}_t\left(\frac{1}{\bar{m}}\widehat{y}_{t+l} - \delta\widehat{y}_{t+1+l}\right). \tag{71}$$

Plugging this in equation (64) and 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 5.1, 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 5.1. The reason being that the behavioral, unconstrained households

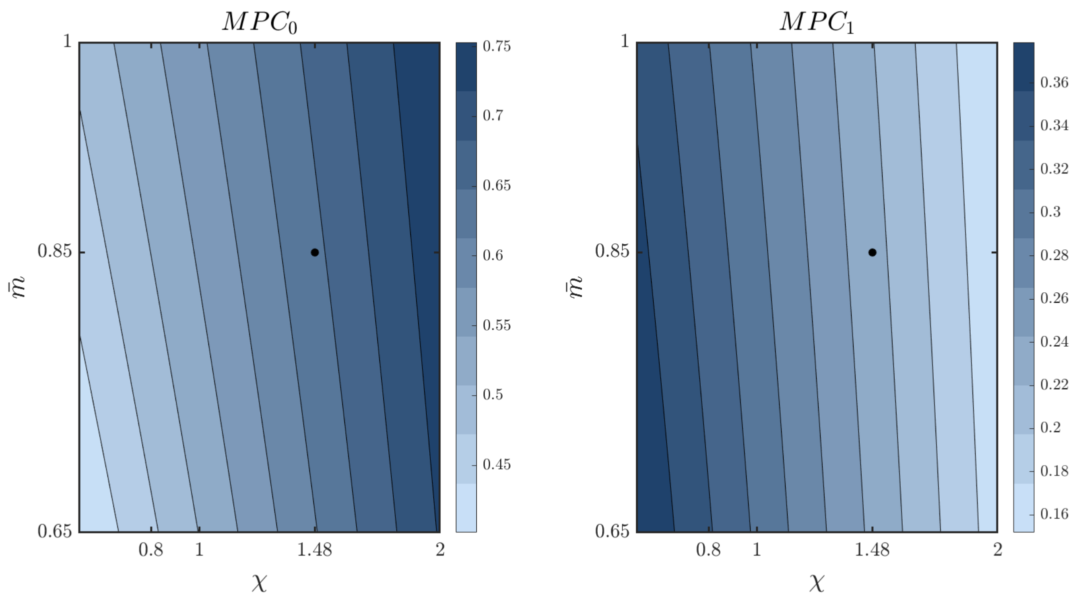
Figure 11: Intertemporal MPCs



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 $\bar{m}$.

save a relatively large amount of the received income windfall in period 0 as they cognitively 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 5.2 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).[28]

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

[28]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 Quantitative Behavioral HANK Model

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

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).

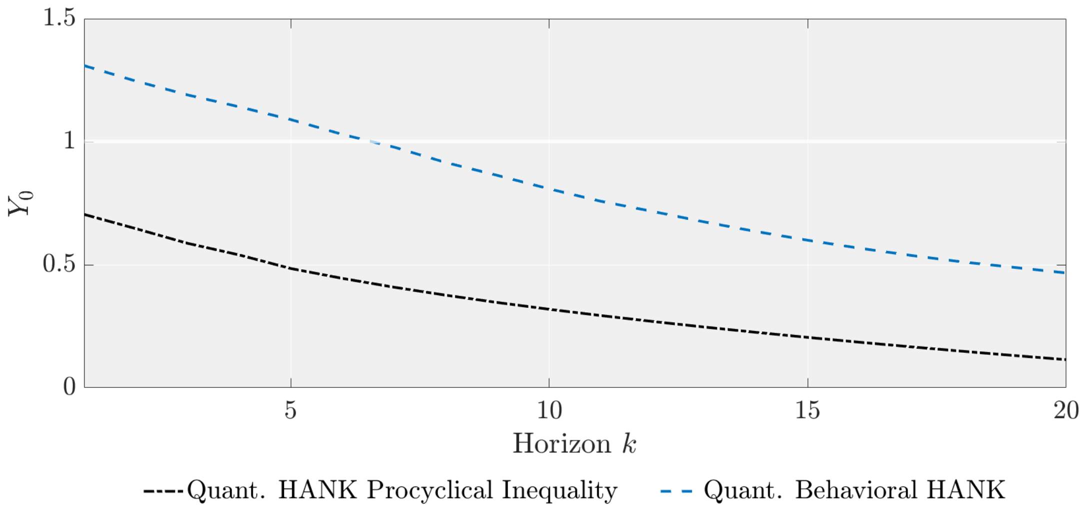
Parameter	Description	Value
R	Steady State Real Rate (annualized)	2%
γ	Risk aversion	2
φ	Inverse of Frisch elasticity	2
μ	Markup	1.2
θ	Calvo Price Stickiness	0
ρ_e	Autocorrelation of idiosyncratic risk	0.966
σ_e^2	Variance of idiosyncratic risk	0.0384
$\tau(e)$	Tax shares	[0, 0, 1]
$d(e)$	Dividend shares	$[0, \frac{0.2}{0.5}, \frac{0.8}{0.25}]$
$\frac{B^G}{4Y}$	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

Figure 13: Resolving the Forward Guidance Puzzle in HANK



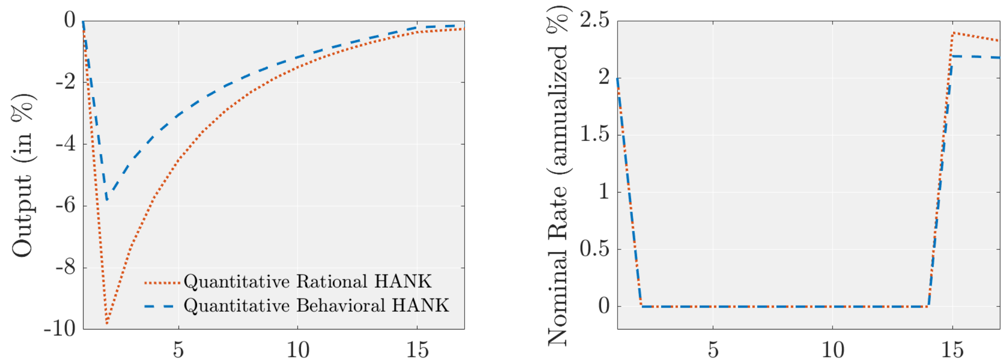
Note: This figure shows the response of total output in period 0 to anticipated i.i.d. monetary policy shocks occurring at different horizons k, relative to the response in the RANK model under rational expectations (equal to 1).

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

Stability at the ELB. Figure 14 shows the output and nominal interest rate response after a shock to the discount factor in the quantitative behavioral HANK model and in its rational counterpart. In particular, the discount factor jumps on impact by 0.8% for 12 quarters before it returns to steady state.

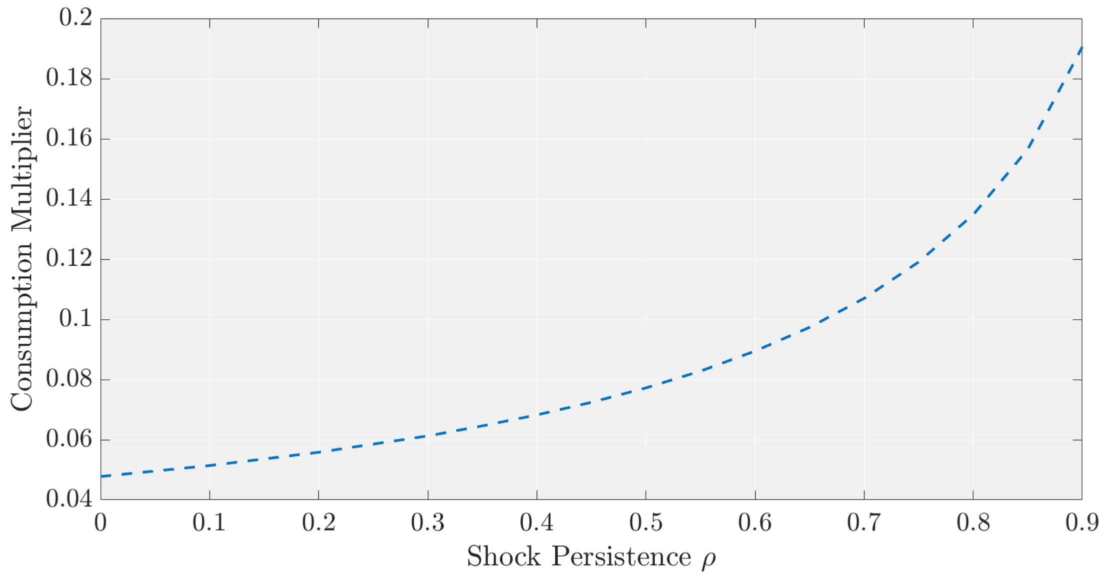
Figure 14: ELB recession in the quantitative behavioral HANK model



Note: This figure shows the impulse responses of total output and of the nominal interest rate after a discount factor shock that brings the economy to the ELB for 12 quarters.

Fiscal Multiplier. To verify that the quantitative behavioral HANK model generates positive consumption multiplier under a constant real rate, we redo the experiments in Section 3.3: the government exogeneously increases government consumption (which is assumed to be zero in steady state) which follows an AR(1)-process. The increase in government consumption is immediately financed by lump-sum taxes. Figure 15 shows the impact multiplier on consumption for various degrees of persistence, ρ_G. It shows that while the multiplier increases in persistence, it is bounded from below by zero. In other words, also the quantitative behavioral HANK model generates positive consumption multipliers, such that, the result from our tractable model carries over to the behavioral HANK model.

Figure 15: Consumption multiplier in the quantitative behavioral HANK



Note: This figure shows the impact consumption multiplier after an exogenous increase in government consumption which is financed by lump-sum taxes for various degrees of persistence.

F.1 Heterogeneous $\bar{m}$

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

$$x_{t+4} - \mathbb{E}_t^{e,BR} x_{t+4} = c^e + b^{e,CG} \left(\mathbb{E}_t^{e,BR} x_{t+4} - \mathbb{E}_{t-1}^{e,BR} x_{t+4} \right) + \epsilon_t^e, \tag{72}$$

to estimate $b^{e,CG}$ for different groups of households, indexed by e. As shown in Appendix B, $b^{e,CG} > 0$ is consistent with underreaction and the corresponding cognitive discounting parameter is approximately given by

$$\bar{m}^e = \left(\frac{1}{1 + b^{e,CG}} \right)^{1/4}. \tag{73}$$

Ideally, we would use actual data and expectations data about future marginal utilities of consumption which, however, are not available. Instead, we focus on expectations about future unemployment. The Survey of Consumers from the University of Michigan provides 1-year ahead unemployment expectations and we use the unemployment rate from the FRED database as our measure of actual unemployment. Consistent with the model, we split the households into three groups based on their income. The bottom and top income groups each contain the 25% households with the lowest and highest income, respectively, and the remaining 50% are assigned to the middle income group.

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

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

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

which after some rearranging yields

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

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

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

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

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

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

FRED as

$$u_{t+4} - \mu_t^e. \tag{74}$$

For the case of expected unemployment changes, we replace u_{t+4} with $(u_{t+4} - u_t)$ in equation (74).

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

$$u_{t+4} - \mu_t^e = c^e + b^{e,CG} \left(\mu_t^e - \mu_{t-1}^e \right) + \epsilon_t^e, \tag{75}$$

to estimate $b^{e,CG}$ for each income group e. Note, however, that the expectations in the forecast revisions are about unemployment at different points in time. To account for this, we instrument forecast revisions by the *main business cycle shock* obtained from Angeletos et al. (2020).

Table 5: Regression Results of Equation (72)

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

Note: This table provides the estimated $\widehat{b}^{e,CG}$ from regression (72) for different income groups. The first three columns show the results when the right-hand side in equation (72) is instrumented using the *main business cycle shock* from Angeletos et al. (2020) and the last three columns using OLS. Standard errors are robust with respect to heteroskedasticity and are reported in parentheses. The row "F-stat." reports the first-stage F-statistic for the IV regressions.

Table 5 shows the results. The first three columns report the estimated $b^{e,CG}$ from the IV regressions and the last three columns the same coefficients estimated via OLS. Standard errors are robust with respect to heteroskedasticity and are reported in parentheses. The row "F-stat." reports the first-stage F-statistic for the IV regressions. We see that in all cases $\widehat{b}^{e,CG}$ is positive, suggesting that households of all income groups tend to underreact, consistent with our assumption of $\bar{m} < 1$.

Using equation (73) we obtain $\bar{m}^e$ equal to 0.86, 0.87 and 0.88 for the bottom 25%, the middle 50% and the top 25%, respectively for the estimates from the IV regressions and 0.82, 0.83 and 0.85 for the OLS estimates. When estimating $\bar{m}^e$ using expected unemployment *changes* instead of the level, the estimated $\bar{m}^e$ equal 0.57, 0.59 and 0.64 for the IV regressions and 0.77, 0.80 and 0.86 for the OLS regressions.

There are two main take-aways from this empirical exercise: first, it further confirms that $\bar{m} = 0.85$ is a reasonable (but rather conservative) deviation from rational expectations.

Second, the data suggests that there is heterogeneity in the degree of rationality conditional on households income. In particular, households with higher income tend to exhibit higher degrees of rationality.[29]

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

F.1.1 Heterogeneous $\bar{m}$: Alternative Scenarios

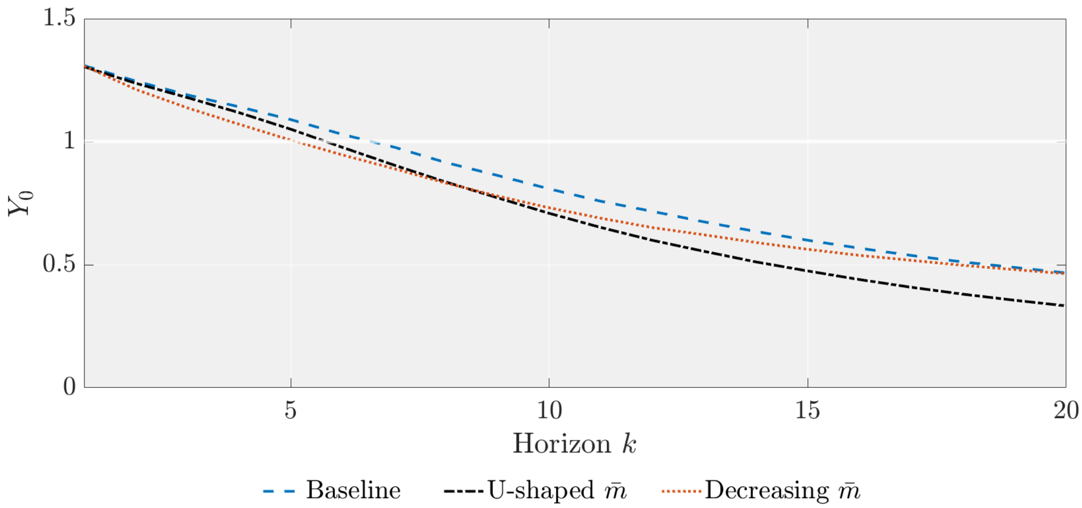
Empirically, we document that richer households tend to deviate somehwat less from rational expectations than poorer households. Broer et al. (2021a) find that the relation between income and forecast accuracy is non-monotonic. In particular, they find that relatively rich and poor households tend to make smaller forecast errors than households with medium level income. To mirror this, we set $\bar{m} = 0.9$ for the high- and low-productivity households and $\bar{m} = 0.8$ for the medium-productivity households. Given that 50% of households fall into the medium category, this calibration again features an average $\bar{m}$ of 0.85. The black-dashed-dotted line in Figure 16 shows the results when re-running the monetary policy experiments outlined in Section 4.

Overall, the results are quite similar to the baseline calibration. Forward guidance is somewhat weaker, which is driven by the lower $\bar{m}$ of the medium-productivity households. These households are usually unconstrained and thus, respond to forward guidance directly. Since they account for half of all households their lower $\bar{m}$ outweighs the higher $\bar{m}$ of high-productivity households, even though these are even less likely to be constrained. But only 25% of all households are high-productivity households whereas 50% are medium-productivity households.

The orange-dotted line shows the result for the case in which low-productivity households are closest to rational expectations, i.e., when their $\bar{m}$ is set to 0.9 and the high-productivity households have a $\bar{m}$ of 0.8. We see that compared to the baseline calibration, the effectiveness of monetary policy drops faster with the horizon as we increase the horizon.

[29] This is consistent with other empirical findings on heterogeneous deviations from FIRE. Broer et al. (2021a), for example, document that wealthier households tend to have more accurate beliefs, as measured by forecast errors.

Figure 16: Heterogeneous $\bar{m}$ and Monetary Policy



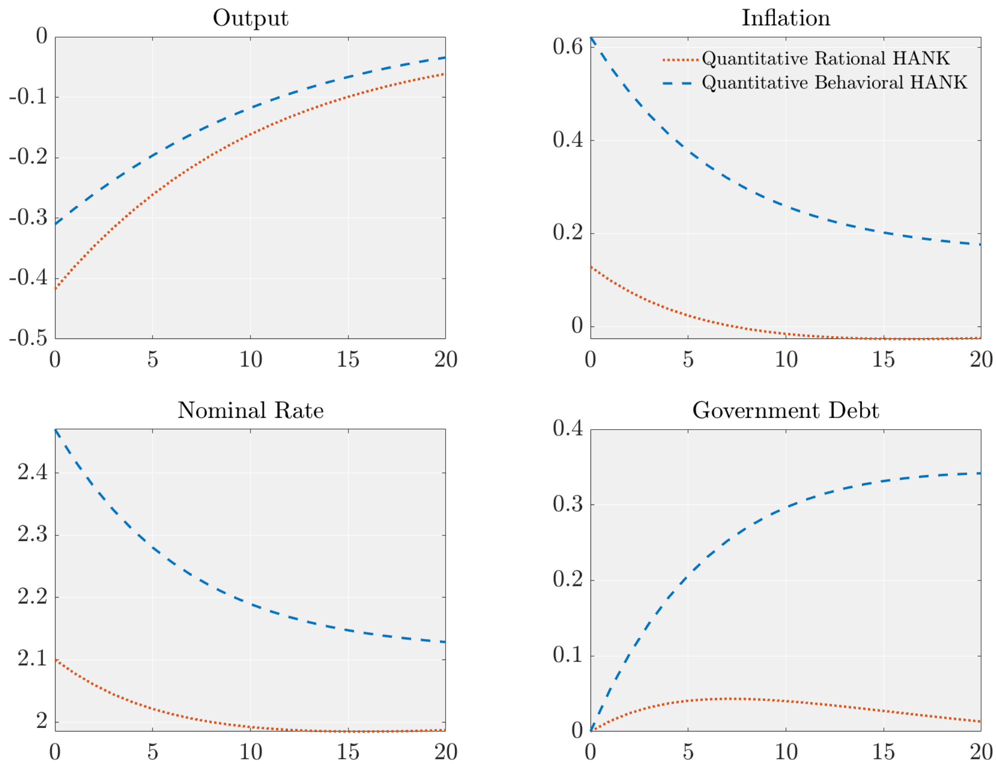
Note: This figure shows the response of total output in period 0 to anticipated i.i.d. monetary policy shocks occurring at different horizons k for the baseline calibration with $\bar{m} = 0.85$ for all households (blue-dashed line), for the model in which low- and high productivity households have $\bar{m} = 0.9$ and medium-level productivity households have $\bar{m} = 0.9$ (black-dashed-dotted line), and the model with $\bar{m} = 0.9$, $\bar{m} = 0.85$, and $\bar{m} = 0.8$ for low- medium- and high-productivity households, respectively (orange-dotted line).

F.2 Additional Figures to Section 4.3

Figure 17 shows the impulse-response functions of output, inflation, nominal interest rates and government debt (as a share of annual GDP) for the same experiment as considered in Section 4.3 but for the case in which monetary policy follows a simple Taylor rule with a response coefficient of 1.5.

As in the case where monetary policy fully stabilizes inflation, the nominal interest rate increases more strongly in the behavioral HANK model than in its rational version. The difference across the two models, however, is somewhat smaller compared to the case in which inflation is completely stable. Inflation, however, increases more strongly in the behavioral model and also government debt increases more substantially.

Figure 17: Cost-push Shock: Taylor Rule

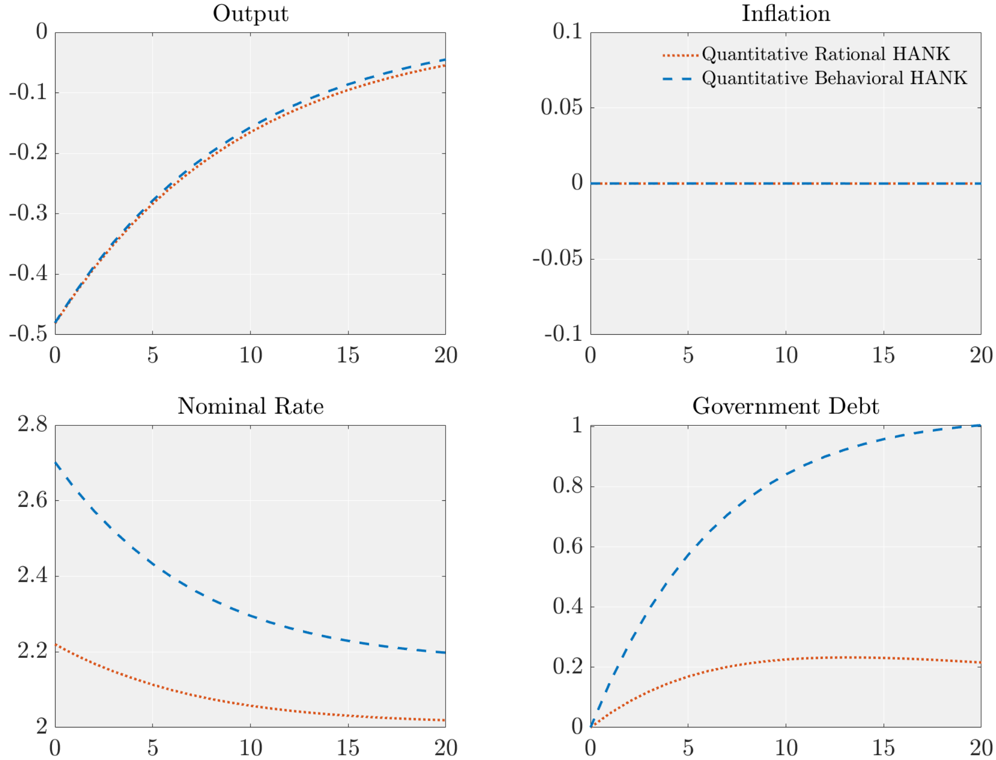


Note: This figure shows the impulse responses after a cost-push shock for the case that monetary policy follows a simple Taylor rule. Output is shown as percentage deviations from steady state, inflation and nominal interest rate as annualized percentage points and government debt level as percentage point deviations in debt-per annual GDP level.

Figure 18 shows the results for the case in which monetary policy fully stabilizes inflation but the economy features a relatively high initial debt level of 90% of annual GDP, similar to the situation many advanced economies face post-Covid.

Compared to the case with a debt level of 60% (as shown in Figure 6), the government debt and the nominal interest rate increases even more in the behavioral HANK model after an inflationary cost-push shock. The increase in government debt (as a share of annual GDP) is about twice as strong as in the baseline calibration, even though we increased the initial debt level by less than 100%.

Figure 18: Cost-push shock post-Covid



Note: This figure shows the impulse responses after a cost-push shock for the case that monetary policy fully stabilizes inflation and and the initial government debt level is 90% per annual GDP.Output is shown as percentage deviations from steady state, inflation and nominal interest rate as annualized percentage points and government debt level as percentage point deviations in debt-per annual GDP level.

G Details on the Backward-Looking Behavioral HANK Model

Here, we discuss how we can calibrate the backward-looking behavioral HANK model from Section 5.3 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 (13).

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.

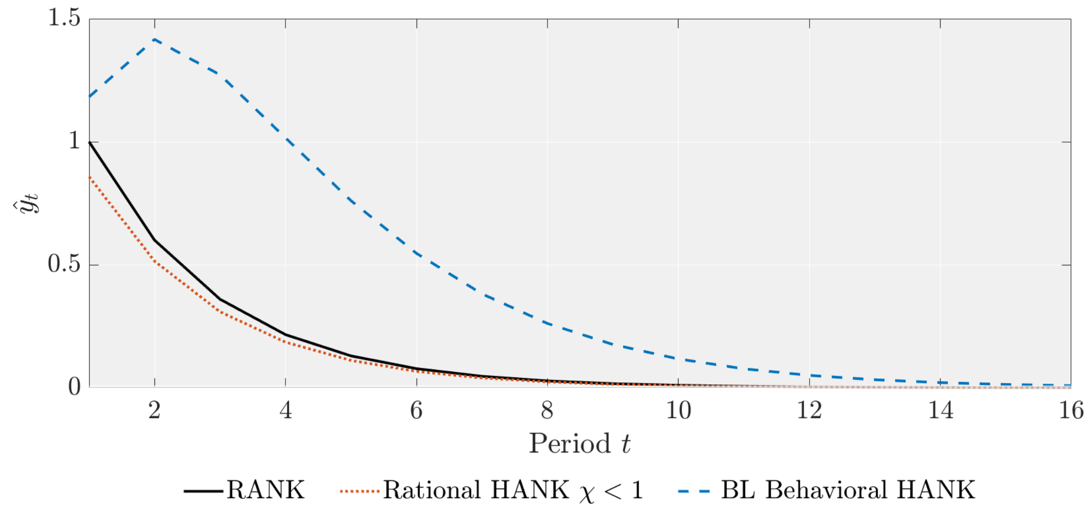
Impulse-Response Functions. We now show how the backward-looking behavioral HANK model generates hump-shaped impulse responses and a novel behavioral amplification channel. To this end, we examine how output in the backward-looking behavioral HANK model responds to an expansionary monetary policy shock 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 19 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 and the black-solid line of RANK.

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 19 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

Figure 19: Output Response to a Monetary Policy Shock



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

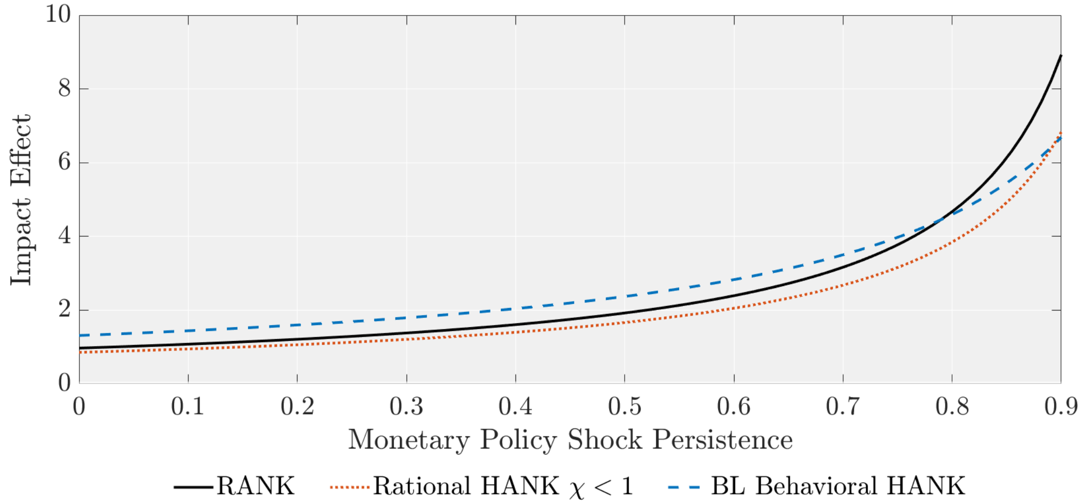
it becomes tomorrow's default value for the household's expectations. The behavioral anchor induces *endogenous* persistence which further increases today's output response through more optimistic expectations. Yet, there is an opposing channel at work: an *exogenously* persistent shock not only decreases 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.[30] Given our calibration, the first channel dominates, thereby generating amplification as depicted in Figure 19.

Given the two opposing forces at work, the degree of initial amplification depends on the persistence of the shock. Figure 20 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.

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 $\chi < 1$.

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

Figure 20: 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.

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

$$\frac{1}{\gamma}\frac{1-\lambda}{1-\lambda\chi} > 0. \tag{76}$$

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 (76). 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.

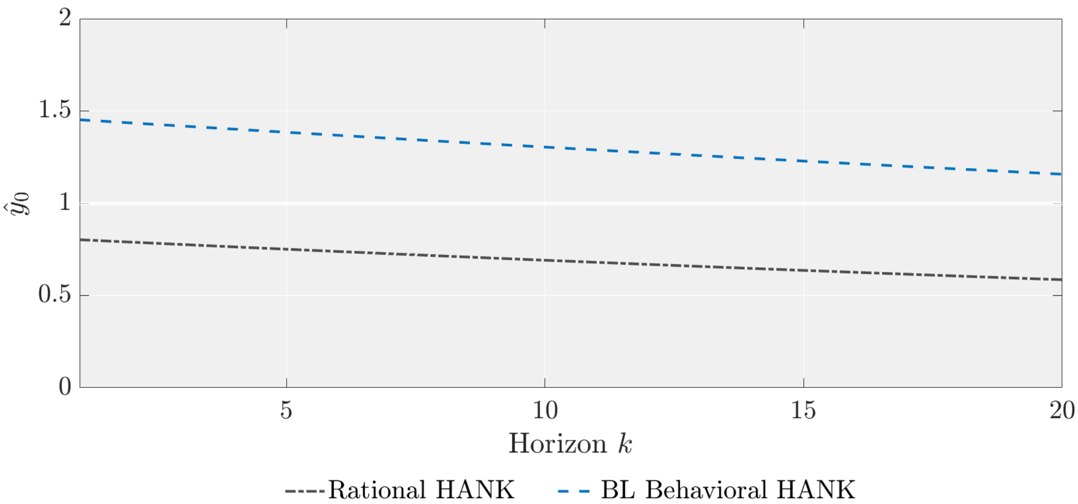
Turning to forward guidance, an expected change in the nominal interest rate in period $t+1$, affects output in t by

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

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 21 shows these patterns graphically. First, the behavioral amplification chan-

Figure 21: 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.

nel 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.