A Behavioral Heterogeneous Agent New Keynesian Model

Oliver Pfäuti* Fabian Seyrich†

January 2026

Link to most recent version

Abstract

We develop a behavioral Heterogeneous Agent New Keynesian model embedding two empirically-disciplined features: households with high MPCs are more exposed to aggregate income fluctuations, and households cognitively discount aggregate news. Together, these features align the model with recent empirical evidence on the monetary transmission, and their interaction generates a demand-driven amplification mechanism making inflation stabilization particularly difficult for monetary policy after supply and government spending shocks. We calibrate the model to micro evidence and find that TFP shocks increase inflation twice as much as when abstracting from heterogeneity and cognitive discounting, with the largest amplification component coming from their complementarity.

Keywords: Monetary Policy, Heterogeneous Households, Behavioral Macroeconomics, Forward Guidance, Business Cycle Shocks, Inflation, Macroeconomic Stabilization

JEL Codes: E21, E52, E62, E71

*The University of Texas at Austin, Department of Economics, pfaeuti.oliver@gmail.com

†Frankfurt School of Finance & Management and DIW Berlin, fabian.seyrich@gmail.com.

We thank Klaus Adam, Marios Angeletos, Adrien Auclert (discussant), Neele Balke, Nicholas Barberis, Christian Bayer, Florin Bilbiie, Oli Coibion, Eduardo Dávila, Davide Debortoli, Andres Drenik, Stéphane Dupraz (discussant), Xavier Gabaix, Morteza Ghomi (discussant), Zhen Huo, Maximilian Jager, Diego Känzig, Alexander Kriwoluzky, Ralph Luetticke, Gernot Müller, Timo Reinelt, Ricardo Reis, Hannah Seidl, Alp Simsek, Alejandro Van der Ghote (discussant), Maximilian Weiß (discussant), Christian Wolf, and numerous seminar and conference participants for helpful comments and suggestions. Oliver Pfäuti gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project C02) and the Stiftung Geld & Währung. Fabian Seyrich gratefully acknowledges financial support by the Leibniz Association through the project "Distributional effects of macroeconomic policies in Europe". First version online: January 2022.

1 Introduction

Monetary policy stabilizes the business cycle and inflation by managing aggregate demand. When demand exceeds productive capacity, the central bank raises interest rates to curb private spending, thereby counteracting inflationary pressures. The effectiveness of monetary policy hinges on how the policy and the underlying shock affect households' incomes and expectations, and how these changes, in turn, shape consumption. A growing empirical literature challenges the workhorse model of monetary policy along these dimensions. Indirect effects operating through changes in household income are central to the monetary transmission. Additionally, households' incomes are affected heterogeneously, with income and consumption inequality moving countercyclically after monetary policy shocks. Furthermore, households' expectations underreact to macroeconomic news so that announced future policy changes are less powerful than what standard models predict. While each of these patterns has been studied in isolation, we show that their *joint* presence has first-order implications for stabilization policy as it severely limits monetary policy's effectiveness to manage demand.[1]

We show this by developing a behavioral Heterogeneous Agent New Keynesian (HANK) model that accounts for these empirical facts simultaneously. Two empirically disciplined features in our model are crucial. First, the earnings of households with high marginal propensities to consume (MPCs) are more sensitive to aggregate fluctuations induced by monetary policy. Second, households cognitively discount expected deviations from the stationary equilibrium, so that their expectations underreact to aggregate news.

While these two features reconcile the model with the empirical evidence on the monetary transmission, their interaction generates a shock-dependent, demand-driven amplification mechanism which is largely absent in rational-expectations or representative-agent models. This amplification makes stabilization difficult, in particular when demand exceeds productive capacity for reasons outside of households' consumption decisions such as a fall in TFP or an increase in government spending. In these cases, labor demand rises, shifting income toward high-MPC households, thereby putting upward pressure on consumption and inflation. At the same time, households' expectations underreact to the prospect of a sustained monetary tightening in response to the shock, causing their consumption to fall less than it would if they were rational. Because this muted expectations channel further amplifies labor-market overheating, it reinforces the redistribution toward high-MPC households and its upward pressure on consumption and inflation.

[1] See, e.g., Ampudia et al. (2018), Slacalek et al. (2020) and Holm et al. (2021) for the empirical relevance of indirect channels in the transmission of monetary policy, Auclert (2019), Patterson (2023), Slacalek et al. (2020), Alves et al. (2020), and Bilbiie et al. (2025) for evidence on households' income exposure, and Del Negro et al. (2023), D'Acunto et al. (2022), Miescu (2022), Roth et al. (2021), Coibion and Gorodnichenko (2015), Angeletos et al. (2021), Pfäuti (2024) for evidence on the (in-)effectiveness of announcements and households' inattention to the macroeconomy.

Calibrating the model to micro evidence on households' unequal exposure and their underreaction to aggregate news shows that a given TFP shock increases inflation about twice as much as in a model that abstracts from household heterogeneity and cognitive discounting. This amplification is mainly shaped by the complementarity between the two: removing either household heterogeneity or cognitive discounting reduces the amplification by roughly two thirds, as removing one feature also strongly dampens the impact of the remaining one. Thus, the joint presence of household heterogeneity and cognitive discounting can rationalize how seemingly modest shocks can generate outsized and persistent inflation responses.

Our model builds on the recent HANK literature which combines the typical Bewley-Huggett-Aiyagari-Imrohoroglu incomplete markets setup with nominal rigidities. Ex-ante identical households face uninsurable idiosyncratic productivity risk, incomplete markets and borrowing constraints. In contrast to that literature, households in our model do not necessarily hold rational expectations as we allow for cognitive discounting of aggregate variables: households anchor their expectations about future macroeconomic variables to the stationary equilibrium and cognitively discount expected future deviations, in the spirit of Gabaix (2020). Expectations then underreact to aggregate news, as we show to be the case empirically.[2] We capture heterogeneous earnings elasticities to the business cycle by assuming that labor hours are set by unions that potentially allocate fluctuations in aggregate hours unevenly across households, as in Auclert and Rognlie (2020).

We start by exploiting a specific calibration that allows for a closed-form solution of our model to establish three novel analytical results that cleanly demonstrate how households' unequal exposure and cognitive discounting interact in amplifying TFP shocks. First, heterogeneous exposure affects the propagation of TFP shocks in this analytical framework if and only if expectations are behavioral. Second, if expectations are behavioral, the more unequally exposed households are, the stronger the impact of TFP shocks on the output gap and on inflation. Likewise, the more behavioral households are, the stronger the impact of TFP shocks on the output gap and inflation. And third, the two features are complements: the marginal impact of unequal exposure is larger, the more behavioral households are and vice versa.

These results stem from a novel IS equation that is shaped by unequal exposure and cognitive discounting, whereas the rest of the model—the supply side and the monetary policy rule—is standard. Importantly, while this IS equation reconciles the model with the empirical evidence on the monetary transmission, it implies that monetary policy is substantially

[2] Angeletos and Lian (2023) show how other forms of bounded rationality or lack of common knowledge can be observationally equivalent. For further evidence on underreaction of expectations or general patterns of inattention, see, e.g., Coibion and Gorodnichenko (2015), Coibion et al. (2023), D'Acunto et al. (2022), Roth et al. (2021) or Angeletos et al. (2021). Kučinskas and Peters (2022) and Born et al. (2024) show that even when agents overreact to micro news, they underreact to macro news.

less effective in stabilizing the economy after TFP shocks. We can derive this IS equation in closed form because we can represent the household block as if there were two representative households and we refer to this special case as a behavioral TANK (Two Agent New Keynesian) model.[3] The first group of households is financially "unconstrained", and the second group consists of "hand-to-mouth" households. The latter exhibit high MPCs and their labor income is more exposed to monetary policy, in line with the data. Because unconstrained households face a risk of becoming hand-to-mouth, they exhibit a precautionary-savings motive. In this simplified version of the model, wages are sticky, whereas prices are fully flexible.

Relative to the rational-expectations, representative-agent New Keynesian model (RANK), the behavioral TANK IS equation differs in three respects. First, the output gap—the gap between actual output and output absent nominal rigidities—is more sensitive to contemporaneous deviations of the real rate from the natural rate. Second, the current output gap is less sensitive to expected future output gaps. Third, the natural rate, that is, the interest rate consistent with a zero output gap, responds more strongly to TFP shocks.

Contemporaneous (one-time) real rate changes are amplified through strong indirect, general-equilibrium effects, because high-MPC households are more exposed to output fluctuations induced by monetary policy. After an expansionary monetary policy shock, labor demand increases which redistributes toward hand-to-mouth households. As these households consume all their disposable income, aggregate consumption responds more strongly than if all households were exposed equally to monetary policy.

Under rational expectations, the same unequal exposure would instead make the current output gap more sensitive to expected future output gaps, counterfactually implying *stronger* effects of forward guidance (see also Bilbiie (2025), Werning (2015), Acharya and Dogra (2020)). After an announced future interest rate cut, households anticipate a boom in the future such that unconstrained households increase consumption already today to smooth it intertemporally. On top, unequal exposure leads unconstrained households to reduce their precautionary savings, because becoming hand-to-mouth is less costly in a future boom. However, since households cognitively discount future changes in aggregate variables, *both* of these channels are attenuated. As a result, the sensitivity of the current output gap to the future output gap is reduced and the model does not suffer from the forward guidance puzzle.

But precisely because future interest rate changes are less effective in moving demand today, the natural rate responds more strongly to TFP shocks than under rational expectations. Households expect interest rates to increase for as long as the productive capacity is

[3] This requires shrinking the idiosyncratic risk to two states, imposing zero-liquidity, and limiting ourselves to only one form of nominal rigidities. Models with a similar household structure are often referred to as TANK models with type switching or as THANK (Tractable HANK) models (Bilbiie, 2025).

reduced. Under rational expectations, this expectations channel is very effective in reducing consumption already today and, hence, relatively small interest rate hikes are sufficient to close the output gap. If households are rational, the real rate tracks the natural rate even without a change in the nominal rate due to an endogenous adjustment of expected inflation in our sticky-wage, flexible-price economy.[4] This implies that TFP shocks do *not* affect the output gap or wage inflation. This neutrality is independent of how unequally exposed households are because with a zero output gap, labor hours remain constant and, thus, incomes of both types of households fall one-for-one with aggregate income. Hence, household heterogeneity is irrelevant for the transmission of TFP shocks under rational expectations, even though heterogeneity affects the effectiveness of monetary policy.

In contrast, under behavioral expectations, households cognitively discount the higher future interest rates in response to the fall in TFP. Thus, stronger increases in interest rates would be required to sufficiently lower consumption to realign it with the lower productive capacity. This stronger movement in the natural rate implies a gap between the actual and the natural real rate for a given Taylor rule. Consequently, the output gap and wage inflation increase—the more so, the more behavioral households are.

Because of the positive output gap, heterogeneous exposure becomes relevant. As demand exceeds productive capacity, firms require more labor. This increase in hours redistributes income toward high-MPC households, whose incomes fall less than one-for-one with aggregate income. This redistribution, *ceteris paribus*, raises aggregate demand which further pushes actual output from its potential—the more so, the more unequally exposed households are. It also implies that consumption inequality falls after a negative TFP shock, consistent with novel empirical evidence on the distributional consequences of TFP changes that we present.

The amplification through unequal exposure is stronger, the more behavioral households are, because then the gap between the actual real rate and the natural rate is larger. As this increases the output gap, it strengthens the redistribution channel through households' unequal exposure. Thus, unequal exposure and bounded rationality are complements that reinforce each other in amplifying the impact of TFP shocks.

We show that the nature of the interaction between unequal exposure and cognitive discounting depends on the underlying shock. For example, after an increase in government spending, consumption needs to fall by the same amount as after a fall in TFP to close the output gap, triggering the same demand-driven amplification through the interaction between unequal exposure and cognitive discounting. In contrast, when the change in demand originates itself in households' response to monetary policy—due to persistent

[4] Prices adjust such that the real wage tracks TFP. With zero wage inflation this implies future deflation and, thus, a rise in the real rate strong enough to sufficiently lower consumption under rational expectations. While this endogenous price adjustment is necessary for our sharp irrelevance result, we show that the complementarity between unequal exposure and cognitive discounting also extends to economies with sticky prices.

monetary policy shocks or because a binding effective lower bound causes monetary policy to be contractionary—unequal exposure and cognitive discounting pull in opposite directions, because behavioral expectations weaken the expectations channel of monetary policy.

Equipped with our analytical insights, we turn to our full behavioral HANK model to quantify the importance of the demand-driven amplification after a negative TFP shock. We relax our calibration restrictions and instead allow for sticky wages and sticky prices, positive government debt supply, a non-degenerate wealth distribution, progressive labor taxes, as well as permanent and transitory shocks to households' income. We rely on estimates from Patterson (2023) and Bilbiie et al. (2025) to calibrate households' unequal exposure in labor earnings. We propose a new approach to estimate the degree of cognitive discounting from households' underreaction to aggregate news using survey data. Our empirical results clearly reject rational expectations, and we use our estimates to calibrate the degree of cognitive discounting.

A fall in TFP reduces output and raises the output gap, inflation, wage inflation, interest rates and the government debt level. Importantly, even though monetary policy raises interest rates more forcefully, consumption decreases less strongly than in models that abstract from household heterogeneity, cognitive discounting, or both. Consequently, the output gap and inflation increase substantially more in the behavioral HANK model. In our baseline calibration, inflation increases on impact 93% more than in the RANK version of our model. Decomposing the amplification reveals that the largest component arises from the interaction between household heterogeneity and cognitive discounting: While household heterogeneity in isolation accounts for about a third and cognitive discounting in isolation for about a fifth, their interaction accounts for 41% of the additional inflation response. Accounting for both features is even more important when monetary policy responds more gradually to the TFP shock, captured by interest-rate smoothing in the Taylor rule. In this case, the amplification becomes larger with the interaction explaining roughly half of the overall amplification.

To isolate the role of unequal exposure, we contrast our model with an acyclical, rational HANK model that shuts down unequal exposure in labor earnings. We find that inflation in our model rises 67% more on impact, showing that unequal exposure and its interaction with cognitive discounting are quantitatively powerful. This exercise also reveals that household heterogeneity generates additional amplification channels and that cognitive discounting positively interacts with those. Specifically, negative TFP shocks strongly increase the government debt level which is expansionary in the presence of non-Ricardian households (Auclert et al. (2024)). This channel is stronger under behavioral expectations because interest rates rise more, increasing the fiscal spillovers of the monetary tightening. As such, the monetary-fiscal interactions—already materially important in HANK models—become even more important when heterogeneous households are behavioral.

Related literature. The existing literature typically treats the empirical facts about the transmission of monetary policy in isolation. The heterogeneous-household literature emphasizes indirect, general-equilibrium channels of monetary policy (Kaplan et al. (2018), Auclert (2019), Auclert et al. (2020), Bilbiie (2020), Luetticke (2021)). Others show how these models can also help address the forward-guidance puzzle (McKay et al. (2016, 2017), Hagedorn et al. (2019a), Acharya and Dogra (2020), McKay and Wieland (2022)). Werning (2015) and Bilbiie (2025) combine these themes and show that, in HANK models, amplification of current monetary policy through redistribution toward high-MPC households comes at the cost of exacerbating the forward-guidance puzzle via dampened precautionary savings—what Bilbiie calls a "Catch-22".

Only few papers resolve this trade-off, and they do so through specific monetary- or fiscal-policy designs. Bilbiie (2025) shows how the forward guidance puzzle vanishes under price level targeting. Hagedorn et al. (2019b), building on Hagedorn (2016, 2018), show that introducing nominal government bonds together with a particular bond-supply rule can resolve the forward-guidance puzzle in a quantitative HANK model. In contrast to these approaches, our behavioral HANK model resolves the Catch-22 through households' consumption behavior, and hence, extends to standard monetary and fiscal rules.

A recent literature combines household heterogeneity with deviations from full-information rational expectations (FIRE). Farhi and Werning (2019) introduce level-k thinking in a HANK model and focus on resolving the forward-guidance puzzle in a setting that abstracts from unequal exposure. We instead consider cognitive discounting à la Gabaix (2020), disciplined using survey expectations, and we show how it interacts with unequal exposure to shape the transmission of monetary policy and business cycle shocks. Auclert et al. (2020) incorporate sticky information in a HANK model to generate hump-shaped responses after monetary policy shocks while matching intertemporal MPCs. Guerreiro (2023) focuses on the role of belief disagreement and how it correlates with households' exposure. In contrast to these papers, we establish new analytical and quantitative insights into how household heterogeneity and cognitive discounting shape the transmission of contemporaneous and announced future monetary policy shocks and what this, in turn, implies for how effective monetary policy is in stabilizing supply and other business cycle shocks.[5]

Our paper further complements existing work that highlights the inflationary effects of reductions in productive capacity, e.g., due to supply chain disruptions (Amiti et al., 2023), or the inflationary effects of unfunded fiscal shocks (Bianchi and Melosi, 2022; Bianchi et al.,

[5] There are also recent contributions to how household heterogeneity and deviations from FIRE interact in settings abstracting from nominal rigidities, see e.g. Broer et al. (2022) and Ilut and Valchev (2023). In addition, Gallegos (forthcoming) examines how the relaxation of the full-information assumption matters in a TANK model. Maxted et al. (2025) introduces *present bias* in a model of household heterogeneity set in partial equilibrium, but they maintain the FIRE assumption. Pfäuti et al. (2024) analyzes how heterogeneity in persistent optimism affects households' consumption-savings decision and fiscal policy.

2023). Others have highlighted the role of a non-linear Phillips Curve (Cerrato and Gitti, 2022; Gitti, 2024; Benigno and Eggertsson, 2023), state-dependent price adjusting frequencies (Cavallo et al., 2024; Blanco et al., 2024), or state-dependent inflation expectations (Weber et al., 2025; Pfäuti, 2025) for inflation. We complement that literature by focusing on the demand side and highlight how the interaction between cognitive discounting and household heterogeneity renders it particularly difficult for monetary policy to sufficiently lower aggregate demand to stabilize inflation after TFP or government spending shocks.

Outline. The rest of the paper is structured as follows. We present our behavioral HANK model in Section 2. In Section 3, we consider a special calibration that allows us to solve the model in closed form and we present our main theoretical insights. In Section 4, we then move to a standard calibration to quantify the role of household heterogeneity, cognitive discounting, and their interaction for the transmission of business cycle shocks. Section 5 concludes.

2 Model

This section presents our model that incorporates household heterogeneity in the standard incomplete markets setup (Bewley, 1976; Imrohoroğlu, 1989; Aiyagari, 1994), cognitive discounting à la Gabaix (2020), and standard New Keynesian nominal rigidities.

2.1 Households

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

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

where φ denotes the inverse Frisch elasticity, γ the relative risk aversion, and Ψ_N the relative weight on the disutility of labor.

Household i faces the budget constraint

$$C_{i,t} + \frac{B_{i,t+1}}{1 + r_t} = B_{i,t} + (1 - \tau_t^L) \left[W_t z(e_{i,t}) N_{i,t}\right]^{\tau_P} + D_t d(e_{i,t}) \tag{2}$$

and the borrowing constraint $B_{i,t+1} \geq \underline{B}$, where $\underline{B}$ denotes an exogenous borrowing limit, B denotes the household's bond holdings, r_t denotes the net real interest rate, W_t the real wage and $e_{i,t}$ the household's exogenous idiosyncratic state that follows a Markov chain with time-invariant transition matrix $\mathcal{P}$. The process for $e_{i,t}$ is the same for all households and the mass of households in state e at any point in time equals the probability of being in that state in the stationary equilibrium, $p(e)$. Conditional on their exogenous idiosyncratic state, households have the idiosyncratic productivity $z(e_{i,t})$ which we further specify in Sections 3.1 and 4.1, respectively, and they receive a share $d(e_{i,t})$ of total dividends D_t, specified in Section 4.1. Households' net labor income is subject to a progressive tax schedule, governed by τ_t^L and τ_P, as in Heathcote et al. (2017). We also allow households' time discount factor to be a function of e, $\beta(e_{i,t})$. We follow the recent HANK literature (e.g., Broer et al., 2020; Auclert et al., 2024) and assume that households' labor hours, N_{it}, are determined by labor unions discussed below.

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

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

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

Bounded rationality. We assume that households are fully rational with respect to their idiosyncratic risk, but they cognitively discount the effects of aggregate shocks.[6] To model cognitive discounting, we follow Gabaix (2020) but extend it to an economy with a non-degenerate distribution of households rather than focusing on a representative consumer.[7] Let X_t be a random variable (or vector of variables) and let us define $\bar{X}_t$ as a default value the agent has in mind and let $\widetilde{X}_{t+1} \equiv X_{t+1} - \bar{X}_t$ denote the deviation from this default value. The behavioral agent's expectation about X_{t+1} is then defined as

$$\mathbb{E}_t^{BR}\left[X_{t+1}\right] = \mathbb{E}_t^{BR}\left[\bar{X}_t + \widetilde{X}_{t+1}\right] \equiv \bar{X}_t + \bar{m}\mathbb{E}_t\left[\widetilde{X}_{t+1}\right], \tag{4}$$

where $\mathbb{E}_t\left[\cdot\right]$ is the rational expectations operator and $\bar{m} \in [0,1]$ is the cognitive discounting parameter. A higher $\bar{m}$ denotes a smaller deviation from rational expectations and rational expectations are captured by $\bar{m} = 1$.

When $\bar{m} < 1$, the behavioral agent anchors her expectations to the default value and cognitively discounts expected future deviations from this default value. Since households only cognitively discount the implications of aggregate shocks, we assume that the default value $\bar{X}_t$ is the stationary equilibrium value of the variable. Thus, in the absence of aggregate shocks, $\widetilde{X}_{t+1} = 0$, households are fully rational implying that cognitive discounting does not affect the stationary equilibrium of the economy.

To see how cognitive discounting matters for the dynamics in our model, note that the

[6]We later discuss in Footnote 19 that relaxing the assumption that households have rational expectations about their idiosyncratic risk does not materially affect our result.

[7]Following Gabaix (2020), we show how to microfound cognitive discounting as a noisy-signal extraction problem in Appendix A.11, but note that the exact microfoundation or underlying behavioral friction which leads to underreaction is not crucial for our analysis.

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

$$\begin{aligned} C_{i,t}^{-\gamma} &\geq \beta(e_{i,t})\mathbb{E}_t^{BR}\left[R_t C_{i,t+1}^{-\gamma}\right] \\ &= \beta(e_{i,t})\mathbb{E}_t^{BR}\left\{R_t\left[\bar{C}_{i,t+1}^{-\gamma} + \left(C_{i,t+1}^{-\gamma} - \bar{C}_{i,t+1}^{-\gamma}\right)\right]\right\} \\ &= \beta(e_{i,t})\mathbb{E}_t\left\{R_t\left[\bar{C}_{i,t+1}^{-\gamma} + \bar{m}\left(C_{i,t+1}^{-\gamma} - \bar{C}_{i,t+1}^{-\gamma}\right)\right]\right\}, \end{aligned} \tag{5}$$

where the rational expectations operator $\mathbb{E}_t[\cdot]$ denotes the expectations that a fully rational household would have in the behavioral economy. The last equality applies our definition of bounded rationality in equation (4) and, following Gabaix (2020), assumes that households are fully rational with respect to today's real interest rate. We show below that this assumption does not materially affect our results.

Equation (5) illustrates that when households form expectations of their marginal utility in the next period, their expectations of the marginal utilities in each individual state are anchored to the marginal utilities in these states in stationary equilibrium. Hence, if an aggregate shock perturbs the expected marginal utility for a given current idiosyncratic state, behavioral agents discount these deviations from stationary equilibrium.

Given $\bar{m} < 1$, expectations underreact to aggregate news about the future compared to rational expectations, that is, they do not fully incorporate aggregate news into their expectations. As we detail in Section 4.1, we propose an approach to estimate $\bar{m}$ using survey data. We find strong evidence for $\bar{m} < 1$, consistent with earlier findings that use different approaches to estimate $\bar{m}$.[8] In fact, our empirical evidence points to $\bar{m} < 0.81$ at quarterly frequency, indicating quite substantial underreaction.

2.2 Firms

We assume a standard New Keynesian firm side with monopolistic competition, (potentially) sticky prices and where firms have rational expectations. 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 $C_t(j) =$

[8]We also discuss a second approach to estimate $\bar{m}$ based on results in Gabaix (2020) and show that the results align with our proposed approach.

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

$$Y_t(j) = A_t N_t(j). \tag{6}$$

A_t represents aggregate productivity which evolves according to

$$A_t = (1 - \rho_a)\bar{A} + \rho_a A_{t-1} + \epsilon_t^A, \tag{7}$$

where ρ_a denotes the persistence and $\bar{A}$ the steady state of TFP, and ϵ_t^A are exogenous shocks to TFP. Firms can only update their prices infrequently with i.i.d. probability θ_p, as in Calvo (1983) and Yun (1996). The real marginal cost is given by $\frac{w_t}{A_t} = \frac{W_t}{P_t A_t}$. We assume that the government may pay a constant subsidy τ_S on revenues.[9] This subsidy is financed by a lump-sum tax on firms T_t^F. Hence, the profit function is:

$$D_t(j) = (1 + \tau_S)[P_t(j)/P_t]Y_t(j) - w_t N_t(j) - T_t^F. \tag{8}$$

2.3 Unions

We follow the recent HANK literature and assume that hours worked $N_{i,t}$ are determined by union labor demand. Our setup here follows Auclert et al. (2024), which is based on Erceg et al. (2000). Each worker provides $n_{i,k,t}$ hours of work to a continuum of unions indexed by $k \in [0, 1]$. Each union aggregates efficient units of work into a union-specific task

$$N_{k,t} = \int e_{i,t} n_{i,k,t} di.$$

A competitive labor packer then packages these tasks into aggregate employment services according to the CES technology

$$N_t = \left(\int_k N_{k,t}^{\frac{\epsilon_w - 1}{\epsilon_w}} \right)^{\frac{\epsilon_w}{\epsilon_w - 1}} \tag{9}$$

and sells these services to firms at price W_t.

We assume that unions receive the standard constant steady-state subsidy financed by lump-sum taxes paid by the union to undo distortions from monopolistic competition in steady state. To isolate the role of behavioral expectations of households, we assume that unions are rational.[10]

We model wage stickiness by imposing a quadratic utility cost $\frac{\psi}{2} \int_k \left(\frac{W_{k,t}}{W_{k,t-1}} - 1 \right)^2 dk$ that appears additively in the household's utility function. A union sets a common nominal wage $W_{k,t}$ per efficient unit for each of its members. In doing so, the union trades-off the marginal disutility of working given average hours against the marginal utility of consumption

[9] We assume a subsidy in the tractable model in Section 3 to undo the steady state distortion arising from monopolistic competition. We abstract from such a subsidy in our quantitative model in Section 4, allowing households to receive dividend income in steady state.

[10] We discuss the implications of behavioral unions in Footnote 26 showing that they would not affect our main qualitative insights.

given average consumption. The union then calls upon its members to supply hours. We assume the union ensures that each household supplies the same amount of hours in steady state. Outside of the steady state, we assume that the union follows specific labor allocation rules to allow for heterogeneous elasticities of earnings to the business cycle, as documented empirically. We specify these labor allocation rules in Section 3 and in Section 4.

Linearized around the zero-inflation steady state, the wage Phillips Curve is given by:

$$\pi_t^w = \beta\,\mathbb{E}_t[\pi_{t+1}^w] \;+\; \bar{\kappa}_w\Big(\frac{1+\varphi(1-\tau_P)}{\varphi}\,\widehat{n}_t + \gamma\,\widehat{c}_t - \tau_P\,\widehat{w}_t - \widehat{\tau}_t^L - \widehat{\xi}_t\Big), \tag{10}$$

with $\bar{\kappa}_w = \frac{\epsilon_w}{\psi}\frac{\Psi_N \bar{N}^{\varphi}}{\bar{C}^{-\gamma}}$ denoting the slope of the Phillips curve, and where hatted variables denote their log-deviations from their respective steady state value and bar variables denote their steady state value. Furthermore, $\xi_t \equiv \int e_{it}\left(\frac{e_{it}N_{it}}{N_t}\right)^{\tau_P} di$ is a labor decomposition term capturing how heterogeneous fluctuations in hours potentially affect the union's target wage.

2.4 Government

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

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

where B^G denotes the bonds issued by the government, T_t denotes tax income, and G_t denotes (exogenous) government spending. Taxes follow a simple debt feedback rule

$$T_t - \bar{T} = \vartheta\frac{B_{t+1}^G - \bar{B}^G}{\bar{Y}}, \tag{11}$$

where $\bar{T}$, $\bar{B}^G$, and $\bar{Y}$ denote the respective steady state values. The parameter ϑ governs how fast the government repays its debt.

The model is closed by a rule for monetary policy. As we will use different monetary policy rules we will specify them further below.

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

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

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

3 Analytical Results

To illustrate the implications of the joint presence of household heterogeneity and cognitive discounting, we start by focusing on a specific calibration strategy that admits a closed-form solution. It allows us to represent the model as if there were only two types of households, even though individual households face idiosyncratic risk and thus, switch between types. Hence, we refer to this specific calibration as behavioral TANK model with type switching or just *behavioral TANK* for convenience.[11] We will first use the behavioral TANK model to show how bounded rationality and household heterogeneity shape the monetary policy transmission to private consumption. We then turn to our main exercise, namely, how household heterogeneity and behavioral expectations interact in shaping the effectiveness of monetary policy to stabilize business cycle fluctuations. The key insight of this Section is that precisely because the behavioral TANK model is consistent with the empirical evidence on the transmission of monetary policy, it implies that monetary policy is substantially less effective in stabilizing the economy after TFP and government spending shocks. Appendix A.2 provides detailed derivations. We relax all specific calibration choices necessary for these sharp analytical results in Section 4 and show numerically that our analytical insights qualitatively carry over to a quantitative version of our model and that they are quantitatively important.

3.1 A Calibration towards a Closed-Form Solution

For now, we simplify the supply side to recover its textbook New Keynesian representation. To this end, we set the elasticity of intertemporal substitution (EIS) to $\frac{1}{\gamma} = 1$, impose only one source of nominal rigidities—following the recent HANK literature (Broer et al., 2020; Auclert et al., 2024), we assume sticky wages and flexible prices[12]—and adopt the optimal New Keynesian subsidy to intermediate firms, $\tau_S = \frac{1}{\epsilon - 1}$, financed by lump-sum taxes paid by firms. We otherwise abstract from fiscal policy and set $B_t^G = G_t = T_t = 0$. We also abstract from heterogeneity in idiosyncratic productivity, as discussed in more detail below. We focus on log deviations from the deterministic steady state, and we denote these deviations using hatted variables.

The supply side is then fully captured by two equations: a zero-profit condition for firms, which links price inflation, π_t, to wage inflation, π_t^w, and the growth rate of TFP, $\Delta\widehat{a}_t \equiv \widehat{a}_t - \widehat{a}_{t-1}$:

$$\pi_t = \pi_t^w - \Delta\widehat{a}_t, \tag{12}$$

[11] This type of model is sometimes also referred to as tractable HANK (THANK) model (Bilbiie, 2020, 2025).

[12] Our results, however, do not depend on this choice. We derive an equivalence result between our tractable model with sticky wages and a version of our tractable model with sticky prices and flexible wages in Appendix A.7.

and the textbook wage Phillips curve:

$$\pi_t^w = \beta \mathbb{E}_t \pi_{t+1}^w + \kappa_w \widehat{x}_t, \tag{13}$$

where $\widehat{x}_t = \widehat{y}_t - \widehat{y}_t^{pot}$ is the output gap, i.e., the log deviation of output $\widehat{y}_t$ from potential output, $\widehat{y}_t^{pot}$, and $\kappa_w = \frac{1+\varphi}{\varphi}\bar{\kappa}_w$ denotes the slope of the Phillips Curve.[13] Potential output is defined as output in the economy in which prices and wages are flexible. Given our assumptions, this is the same as in the standard RANK and behavioral RANK models (Gabaix, 2020) and is given by[14]

$$\widehat{y}_t^{pot} = \widehat{a}_t. \tag{14}$$

We then impose the following additional calibration restrictions to obtain an analytically-tractable representation of the demand side of our economy. First, we assume that there are only two exogenous idiosyncratic states, $e \in \{U, H\}$, and denote a household's probability to remain in her current state $p(e_{t+1} = U|e_t = U) = s$ and $p(e_{t+1} = H|e_t = H) = h$, respectively. As these transition probabilities are constant over time, the share of households in state H, $\lambda = \frac{1-s}{2-s-h}$, is also constant over time.

Second, we focus on the zero-liquidity equilibrium (see, e.g., Ravn and Sterk (2017); Auclert et al. (2024); Werning (2015); McKay et al. (2017); Bilbiie (2025)) and rule out household borrowing by setting $\underline{B} = 0$. Together with the assumption of zero government debt, this implies that households cannot save in equilibrium such that the idiosyncratic state is fully described by e.

Given the absence of taxes and profit income, the only source of heterogeneity across households is their labor income, $w_t n_{i,t} z(e_i)$. For simplicity, we also abstract from differences in idiosyncratic productivities across the two states, that is $z(e_H) = z(e_U) = 1$, which effectively leaves hours worked as the only source of heterogeneity in income. As discussed in Section 2, all households work the same number of hours in the stationary equilibrium implying that there is no inequality in the stationary equilibrium. We then assume that $\beta(H) < \beta(U)$ such that the Euler equation (5) always holds with equality for households in state U, while it always holds with inequality for households in state H. In other words, H households are *H*and-to-Mouth, while U households are *U*nconstrained.

If there is an aggregate shock and hours deviate from their steady state value, $\widehat{n}_t \neq 0$, the labor union allocates these hours according to the following allocation rule:

$$\widehat{n}_t^H = \chi \widehat{n}_t = \chi(\widehat{y}_t - \widehat{a}_t),$$

with χ capturing the H households' sensitivity of hours worked to changes in total hours

[13] We arrive at equation (13) by setting $\gamma = 1$, $\widehat{\tau}_t^L = 0$, and $\tau_P = 1$ in equation (10). Furthermore, $\widehat{\xi}_t^{pc} = 0$ because we for now abstract from heterogeneity in idiosyncratic productivities.

[14] Potential output moves one for one with TFP because in the benchmark economy with flexible prices and flexible wages, the income and substitution effects of a TFP shock on households' labor supply exactly offset each other such that aggregate labor hours remain constant. The fact that the labor supply of households is determined by a union does not affect this result as the union solves a "representative" labor-leisure equation.

worked, and where the second equality follows from the production function. Since this allocation rule is the only source of income heterogeneity, χ is a sufficient statistic of households' income exposure to monetary policy shocks (as $\widehat{a}_t = 0$ in this case). In the following, we will refer to $\chi \neq 1$ as "unequal exposure". As in Bilbiie (2020, 2025), χ is a sufficient statistic for household heterogeneity in our behavioral TANK model.[15]

From the H households' budget constraint, the production function, and market clearing, we obtain consumption of the two types:

$$\widehat{c}_t^H = \chi \widehat{y}_t + (1-\chi)\widehat{a}_t \tag{15}$$

$$\widehat{c}_t^U = \frac{1-\lambda\chi}{1-\lambda}\widehat{y}_t + \left(1 - \frac{1-\lambda\chi}{1-\lambda}\right)\widehat{a}_t. \tag{16}$$

Consumption inequality is given by:[16]

$$\widehat{c}_t^U - \widehat{c}_t^H = \frac{1-\chi}{1-\lambda}\left[\widehat{y}_t - \widehat{a}_t\right]. \tag{17}$$

Thus, if $\chi > 1$, inequality is countercyclical conditional on a monetary policy shock ($\widehat{a}_t = 0$), whereas it is procyclical if $\chi < 1$. The empirical evidence points towards $\chi > 1$: Coibion et al. (2017), Mumtaz and Theophilopoulou (2017), Ampudia et al. (2018) and Samarina and Nguyen (2024) all provide evidence of countercyclical inequality conditional on monetary policy shocks. Patterson (2023) documents a positive correlation between households' MPCs and their elasticities of earnings to GDP in the U.S. Bilbiie et al. (2025) report qualitatively similar findings for Norway albeit somewhat more muted. Alves et al. (2020) provide further evidence consistent with these patterns.[17] These findings also point toward $\chi > 1$ because then hand-to-mouth households that have higher MPCs are more strongly exposed to output fluctuations driven by demand shocks. In Appendix C, we provide novel empirical evidence on the effects of TFP changes on consumption inequality and show that our findings are consistent with $\chi > 1$ also when focusing on TFP shocks. In particular, our empirical findings show that consumption inequality increases after a positive TFP shock which is also what our model predicts with $\chi > 1$, as we discuss later. Given that all these different views on χ point towards $\chi > 1$, we focus on $\chi > 1$ unless explicitly stated otherwise.

Log-linearizing the Euler equation of U households (5) yields:

$$\widehat{c}_t^U = s\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^U\right] + (1-s)\mathbb{E}_t^{BR}\left[\widehat{c}_{t+1}^H\right] - (i_t - \mathbb{E}_t\pi_{t+1}). \tag{18}$$

[15]We use χ as this is reminiscent of the χ parameter in Bilbiie (2020, 2025) because, as it turns out, it captures the same sufficient statistic but in a model with sticky prices and flexible wages and with a different underlying mechanism of this redistribution channel. See also Appendix A.7 for a discussion how the equivalent sticky-price, flexible-wage version of our model nests Bilbiie (2020, 2025).

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

[17]Alves et al. (2020) show that exposure might be U-shaped along the income distribution when correctly accounting for top-income earners, especially the top 1%. This U shape is largely driven by performance-related bonus payments, which our model does not capture.

For the case without idiosyncratic risk, i.e., for $s = 1$, equation (18) boils down to a standard Euler equation under bounded rationality. For $s \in [0, 1)$, however, the household takes into account that she might experience an idiosyncratic shock and self-insures against becoming hand-to-mouth next period. How strongly this precautionary savings motive affects the household's consumption away from the stationary equilibrium will depend on the household's degree of bounded rationality.[18]

We can then derive an aggregate IS equation by combining the H households' consumption with the consumption of U households and their Euler equation (18).

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

$$\widehat{x}_t = \psi_f \mathbb{E}_t \widehat{x}_{t+1} - \psi_c \left[r_t + (1 - \bar{m}\rho_a)\widehat{a}_t \right] \tag{19}$$

where

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

and $r_t \equiv i_t - \mathbb{E}_t \pi_{t+1}$ is the (ex-ante) real interest rate. The natural rate, i.e. the real rate consistent with a zero output gap, is given by

$$r_t^n = -(1 - \bar{m}\rho_a)\widehat{a}_t. \tag{20}$$

There are three differences in the IS equation of the behavioral TANK model compared to the textbook New Keynesian model. The first is the coefficient in front of the gap between the actual real rate and the natural real rate, ψ_c. Despite behavioral expectations, ψ_c is the same coefficient as the one in front of the real rate in the IS equation in the rational TANK model (with or without type-switching) in Bilbiie (2025). There are, however, two differences: first, Bilbiie (2025) abstracts from fluctuations in the natural rate driven by TFP shocks and second, the underlying redistribution mechanism in our model is caused by the higher sensitivity of the labor income of high MPC households on aggregate labor hours rather than by countercyclical profits as in Bilbiie (2025). The second new coefficient, ψ_f, governs the sensitivity of today's output gap to next period's output gap. Proposition 1 shows that $\psi_f = \bar{m}\delta$, where $\bar{m}$ is the coefficient in front of future expected output in Gabaix (2020), whereas δ is isomorphic to the coefficient in front of future expected output in Bilbiie (2025), again with a different micro-foundation.[19] The third difference is that $\bar{m}$ shows up in the natural rate. Despite the presence of household heterogeneity, this natural rate is the

[18] As discussed in Section 2, we follow Gabaix (2020) in assuming that households are fully rational with respect to today's real interest rate. We show in Appendix A.2 that our results do not depend on this assumption.

[19] While we focus on the case in which households have rational expectations about their idiosyncratic risk, relaxing this assumption would affect ψ_f. If households overestimate the persistence of their idiosyncratic productivity (as in Rozsypal and Schlafmann (2023)) and, hence, they overreact to idiosyncratic news—as (Born et al., 2024) find for firms—we can replace the staying probability s in ψ_f with a perceived probability $\tilde{s} > s$, resulting in a lower ψ_f. As it does not qualitatively alter our following analysis, we abstract from non-rational expectations about idiosyncratic risk.

same as in Gabaix (2020)'s representative agent model.[20]

Just like in the RANK model, the full dynamics of our behavioral TANK model are then described by the IS equation (19), the wage Phillips curve (13) plus the zero-profit condition (12), a rule for monetary policy and processes for TFP and monetary policy shocks. In addition, all differences in aggregate dynamics following monetary policy shocks or fluctuations in TFP compared to RANK originate solely from the novel IS equation.

3.2 Monetary Policy in the Behavioral TANK Model

We start by isolating the impact of heterogeneous exposure of households and their behavioral expectations on the monetary transmission. To this end, we for now abstract from movements in the natural rate by holding TFP constant, $\widehat{a}_t = 0$. We further assume that monetary policy directly controls the real interest rate, $r_t = \epsilon_{t-k}$. We then analyze the transmission of monetary policy shocks at different horizons by considering: (i) a contemporaneous monetary policy shock with $k = 0$ and (ii) a forward guidance shock, i.e., a news shock today about a decrease in the real interest rate $k > 0$ periods in the future. Monetary policy keeps the real rate at its steady state level in all other periods. Since our framework reproduces several key insights from earlier work that studies either cognitive discounting or household heterogeneity in isolation (e.g., Gabaix, 2020; Bilbiie, 2025; Werning, 2015), we keep the discussion here brief.

Focusing on the equilibrium which returns to the stationary equilibrium in the long-run, the solution is then characterized by the real rate rule and the forward iterated IS equation:[21]

$$\widehat{y}_t = -\psi_c \mathbb{E}_t \sum_{j=0}^{\infty} \psi_f^j r_{t+j}. \tag{21}$$

Equation (21) highlights the impact of the two coefficients that are new relative to RANK: ψ_c acts as a level shifter of monetary policy, whereas ψ_f, governs the relative strength of real rate changes at different horizons on today's output. From that, the following proposition directly follows.

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

$$\psi_c > 1 \Leftrightarrow \chi > 1, \tag{22}$$

[20]The IS equation in Proposition 1 nests representative-agent models in two ways: first directly by letting $\lambda \to 0$. Second, with $\chi = 1$, we have $\psi_c = \delta = 1$ and, hence, we recover the dynamics of representative-agent models as the presence of the two types of households does not matter for aggregate dynamics (see also Bilbiie, 2025).

[21]In general, we obtain $\widehat{y}_t = -\psi_c \sum_{j=0}^{\infty} \psi_f^j r_{t+j} + \lim_{j\to\infty} \psi_f^j \mathbb{E}_t \widehat{y}_{t+j}$. By assuming that the economy returns to its stationary equilibrium in the long run, we impose that the last term is 0 in all model versions considered.

and the forward guidance puzzle is ruled out if and only if

$$\psi_f < 1 \Leftrightarrow \bar{m} < \frac{1}{\delta}. \tag{23}$$

Because ψ_c is not affected by behavioral expectations, unequal exposure $\chi > 1$ is a necessary and sufficient statistic for $\psi_c > 1$ captured in Condition (22). Hence, this is the same condition as in Bilbiie (2025), albeit with a different micro-foundation. The behavioral TANK model generates amplification of contemporaneous monetary policy relative to RANK whenever high-MPC households' incomes are relatively more sensitive to aggregate income fluctuations, conditional on monetary policy shocks. After a decrease in the interest rate, unconstrained households want to consume more and save less. To produce this increase in demand, firms increase labor hours. Given $\chi > 1$, unions allocate relatively more of these additional hours to H households. This implies that the incomes of households with higher MPCs move more than one-for-one with output after monetary policy shocks, in line with the empirical evidence. Because these households consume all their income, this shift in resources towards high-MPC households amplifies the effect on aggregate demand, captured by $\psi_c > 1$. More generally, because these households consume all their income, indirect effects through households' incomes play a major role in the monetary transmission. In Appendix A.5, we show that for a standard calibration, roughly 60% of the monetary transmission is through indirect effects which is quantitatively similar to larger models, as for example in Kaplan et al. (2018).

Condition (23) in Proposition 2 states that the forward guidance puzzle is ruled out if $\psi_f < 1$, because in that case, a forward guidance shock has a smaller effect on current output than a contemporaneous monetary policy shock of the same size. Hence, the effectiveness of forward guidance is governed by the interaction of behavioral expectations and household heterogeneity in the coefficient $\psi_f = \bar{m}\delta$. Starting with the second factor, which comes from the presence of household heterogeneity, $\chi > 1$ necessarily implies $\delta > 1$ as in Bilbiie (2025): Given that the incomes of H agents move more than one for one with aggregate income after monetary policy shocks, hand-to-mouth households are relatively better off in the period the real rate change occurs. Unconstrained households who self-insure against becoming hand-to-mouth in the future therefore cut back their precautionary savings when they expect a future decrease in the interest rate. The reduction in precautionary savings compounds the increase in output today after a forward guidance shock.

Yet, when households are boundedly rational, they cognitively discount these effects taking place in the future which is reflected in $\bar{m} < 1$. Importantly, unconstrained households cognitively discount both the usual consumption-smoothing response due to the future increase in consumption (as in Gabaix (2020)), as well as the general equilibrium implications for their precautionary savings as they do not fully incorporate the effects of this policy on their own income risk. Both of these effects decrease the effectiveness of the forward guidance shock on today's consumption. As a result, the forward guidance puzzle is resolved if $\bar{m} < \frac{1}{\delta}$ as stated in Condition (23). Given a standard calibration (discussed in Appendix A.1), this condition is met if $\bar{m} < 0.966$. Hence, already a small degree of cognitive discounting is sufficient to resolve the forward guidance puzzle in the behavioral TANK model. As we show in Appendix B, this is far above the upper bounds of our empirical estimates.

Resolving the Catch-22. Comparing the behavioral TANK to its rational counterpart illustrates that the behavioral TANK model overcomes a shortcoming inherent in rational TANK (and HANK) models—the *Catch-22* (Bilbiie (2025); see also Werning (2015)). The Catch-22 describes the tension that the rational HANK (and TANK) model can either generate amplification of contemporaneous monetary policy *or* solve the forward guidance puzzle. To see this, note that with $\bar{m} = 1$, we have $\psi_f = \delta$ and, thus, exactly the same Catch-22 conditions as in Bilbiie (2025): the forward guidance puzzle is resolved when $\delta < 1$ which requires $\chi < 1$. However, counterfactually assuming $\chi < 1$ leads to *dampening* of contemporaneous monetary policy instead of amplification. Note that also rational TANK models without type switching—nested in our model by setting $\bar{m} = s = 1$—or the behavioral RANK model—nested by setting $s = 1$ and $\lambda \to 0$—would not deliver amplification and resolve the forward guidance puzzle simultaneously. TANK models without type switching cannot solve the forward guidance puzzle while bounded rationality in RANK does not deliver initial amplification through indirect effects. We conclude that the joint presence of cognitive discounting and unequal exposure of households is not only supported by micro data but also warranted to solve the Catch-22.

Persistent monetary policy shocks. A corollary of Proposition 2 is that persistent monetary policy shocks can either be amplified or dampened relative to RANK depending on the persistence of the shock. On the one hand, persistent expansionary monetary policy shocks signal that the interest rate will remain below its steady state level for longer. Behavioral households, however, cognitively discount these lower future interest rates which dampens the effectiveness of this expectations channel. This is not the case in RANK, in which expectations about future lower interest rates are as effective as lower interest rates today—the forward guidance puzzle. On the other hand, unequal exposure, $\chi > 1$, amplifies persistent monetary policy shocks through the static redistribution channel as well as the reduction in households' precautionary savings, as discussed above.[22] When real rate changes are sufficiently persistent, the attenuation of the forward guidance channel dominates, making monetary policy in the behavioral TANK less effective than in RANK. Appendix A.3 derives the cutoff level of persistence.

[22] A corollary from the aggravated forward guidance puzzle in the rational TANK model with type switching is that the forward guidance channel is even more important than in RANK. Consequently, the amplification of rational TANK over RANK is increasing in the persistence of the real rate shock.

Determinacy and stability at the effective lower bound. In Appendix A.4, we show that the behavioral TANK model widens the determinacy region compared to RANK (and rational TANK models) by deriving the Taylor principle in our model. We also show that—just as after persistent monetary policy shocks—unequal exposure and cognitive discounting pull in opposite directions for the recessionary impact of extended periods at the effective lower bound.

3.3 Propagation of Business Cycle Shocks

Having characterized their role in the monetary transmission, we now analyze how the joint presence of household heterogeneity and cognitive discounting shapes the effectiveness of monetary policy in stabilizing the economy after business-cycle shocks. We start by considering TFP shocks with a persistence of $\rho_a \in (0, \frac{1}{\delta})$.[23] We define the endogenous response of monetary policy to these business cycle shocks by a standard Taylor rule:

$$i_t = \phi_w \pi_t^w, \tag{24}$$

where wage inflation, π_t^w, is determined by the Phillips curve (13) and $\phi_w > 1$ governs the hawkishness of the central bank. Writing the Taylor rule in terms of wage inflation instead of price inflation has three advantages in our sticky-wage, flexible-price setup: First, it allows us to derive a closed-form solution without any additional state variables which, second, allows us to establish an equivalence result with a sticky price model.[24] Third—and most importantly—it aligns with the fact that wage inflation is the primary source of inefficiency in this framework, whereas flexible prices are an important adjustment margin. In contrast, a Taylor rule responding to price inflation would create large output losses via "Keynesian supply shocks". After a negative TFP shock, the only way to stabilize price inflation is to induce wage *deflation* (see equation (12)), which forces output to fall by more than potential (see Appendix A.6).

The solution for the aggregate dynamics of the economy after a TFP shock is then characterized by the following proposition:

Proposition 3. *The dynamics of the output gap and wage inflation after a TFP shock are given by*

$$\widehat{x}_t = -\alpha \widehat{a}_t, \text{ and } \pi_t^w = \tilde{\kappa}_w \widehat{x}_t \tag{25}$$

[23] Assuming $\rho_a < \frac{1}{\delta}$ ensures that the rational TANK version of our model is always non-explosive. Given our baseline calibration, $\frac{1}{\delta} = 0.97$. In the behavioral TANK model, $\rho_a < \frac{1}{\bar{m}\delta}$ would be sufficient which is less restrictive than $\rho_a < 1$, given that $\bar{m}\delta < 1$ in the behavioral TANK model.

[24] If the nominal rate would respond to price inflation, we would have to keep track of $\widehat{a}_{t-1}$ due to the non-profit condition which pins down price inflation. If instead the nominal rate responds to wage inflation, price inflation only enters the model through the expected inflation term in the IS equation, and hence, $\widehat{a}_t$ remains the only state variable. See Appendix A.7 for the equivalence result with a sticky price, flexible wage version of the model.

where

$$\alpha \equiv \frac{\rho_a(1-\bar{m})\psi_c}{1-\psi_f\rho_a+\psi_c(\phi_w-\rho_a)\tilde{\kappa}_w} \geq 0, \quad \tilde{\kappa}_w \equiv \frac{\kappa_w}{1-\beta\rho_a}.$$

Proposition 3 shows that a negative TFP shock affects the output gap if and only if expectations are behavioral, that is, if $\bar{m} < 1$.[25] In that case, a negative TFP shock increases the output gap. Proposition 3 further shows that the path of the output gap is a sufficient statistic for wage inflation dynamics, since wage inflation is the sum of today's and (discounted) future output gaps, which can be summarized by $\tilde{\kappa}_w$, $\pi_t^w = \tilde{\kappa}_w \widehat{x}_t$. This implies that, for a given $\tilde{\kappa}_w$, the output gap dynamic after TFP shocks, captured by $-\alpha$, is a sufficient statistic for how successful a given monetary policy rule is in stabilizing business cycle fluctuations driven by TFP shocks.[26]

The following Proposition summarizes how bounded rationality and heterogeneous exposure affect the impact of TFP shocks on the output gap:

Proposition 4. *Bounded rationality, $\bar{m}$, affects the solution α in the following way:*

$$\frac{\partial\alpha}{\partial\bar{m}} = \frac{-\rho_a\psi_c\left[1-\delta\rho_a+\psi_c(\phi-\rho_a)\tilde{\kappa}_w\right]}{(1-\psi_f\rho_a+\psi_c(\phi_w-\rho_a)\tilde{\kappa}_w)^2} < 0. \tag{26}$$

Hence, the lower the degree of rationality, the stronger the effects of TFP shocks on the output gap.

Heterogeneous exposure, χ, affects the solution α in the following way:

$$\frac{\partial\alpha}{\partial\chi} = \frac{\rho_a(1-\bar{m})}{(1-\psi_f\rho_a+\psi_c(\phi_w-\rho_a)\tilde{\kappa}_w)^2}\left[\frac{\partial\psi_c}{\partial\chi}(1-\psi_f\rho_a)+\frac{\partial\delta}{\partial\chi}\psi_c\bar{m}\rho_a\right] \geq 0, \tag{27}$$

with

$$\frac{\partial\alpha}{\partial\chi} = 0 \text{ if } \bar{m} = 1$$
$$\frac{\partial\alpha}{\partial\chi} > 0 \text{ if } \bar{m} < 1,$$

where $\frac{\partial\psi_c}{\partial\chi} = \frac{\lambda(1-\lambda)}{(1-\lambda\chi)^2} > 0$. Therefore, unequal exposure affects the transmission of TFP shocks if and only if expectations are behavioral. Under behavioral expectations, more unequal exposure amplifies the effects of TFP shocks on the output gap.

The first part of Proposition 4 shows that the more behavioral households are—that is

[25] It is well known that TFP shocks affect the output gap in the textbook, rational-expectations New Keynesian model with sticky prices. As we discuss below Proposition 4, the key distinction is that with flexible prices and sticky wages, the endogenous price adjustment is stabilizing. In Appendix A.7, we show that the qualitative nature of the interaction between cognitive discounting and household heterogeneity carries over to the sticky-price, flexible-wage model. We further show how the irrelevance result of TFP shocks under rational expectations can arise in the sticky price, flexible wage model.

[26] Behavioral unions would attenuate the mapping from expected future output gaps to wage inflation—reflected in a lower $\tilde{\kappa}_w$—and, thus, also affect the monetary policy response, but they would not directly affect the amplification mechanism operating through the IS equation. In particular, when unions are behavioral, the Phillips curve (13) remains of the same form except that the discount factor β would be multiplied by $\bar{m}$, see Appendix A.2. This implies that our qualitative results in this section would be unaffected by behavioral unions.

the lower $\bar{m}$—the larger the impact of a TFP shock on the output gap and wage inflation, reflected in a larger α. The second part of Proposition 4 describes how heterogeneous exposure χ impacts the output gap response to a TFP shock. It shows that household heterogeneity is irrelevant for the transmission of TFP shocks under rational expectations as unequal exposure does not affect the solution α with $\bar{m} = 1$. Under rational expectations, TFP shocks do not affect the output gap at all—irrespective of the presence of heterogeneous households or a representative agent, as shown in Proposition 3. In contrast, under behavioral expectations, the presence of household heterogeneity amplifies the business cycle implications of TFP shocks: the more unequally exposed households are, that is the larger χ, the higher the impact of TFP shocks on the output gap and on wage inflation.

These results can be understood through the IS equation, in particular, from how household heterogeneity and bounded rationality affect the gap between the actual real rate and the natural real rate and how this gap, in turn, propagates through the economy. As stated in Proposition 1, the impact of TFP fluctuations on the natural rate is unaffected by household heterogeneity. When output is at potential, aggregate labor is constant and the incomes of both types move one-for-one with aggregate output. Hence, there is no redistribution, making household heterogeneity irrelevant for aggregate dynamics and, consequently, for the natural rate.

In contrast, bounded rationality affects the natural rate for the same reason that it weakens forward guidance. After a negative TFP shock, the real rate in the flexible-price, flexible-wage economy must rise as long as TFP is below its steady-state level and expectations of these higher future interest rates reduce demand already today. The more behavioral households are, however, the less effective this expectations channel becomes in managing current demand. Consequently, larger real-rate adjustments at all horizons are required to bring actual output back to potential.

Fluctuations in TFP also directly affect the actual real rate, because TFP shocks drive a wedge between price and wage inflation to realign the real wage with TFP. Substituting the zero-profit condition (12) into the real rate, we obtain:

$$r_t \equiv i_t - \mathbb{E}_t \pi_{t+1} = i_t - \mathbb{E}_t(\pi^w_{t+1} - \Delta \widehat{a}_{t+1}) = i_t - \mathbb{E}_t \pi^w_{t+1} - (1 - \rho_a)\widehat{a}_t, \tag{28}$$

with $-(1-\rho_a)\widehat{a}_t$ being the direct effect of TFP on the real rate. This direct effect is unaffected by behavioral expectations or household heterogeneity and it coincides with the natural rate under rational expectations.[27] Hence, the real rate tracks the natural rate without monetary policy responding implying that TFP has no impact on the output gap when agents are

[27] We follow Gabaix (2020) and assume that households are rational with respect to today's real interest rate. If, instead we assume that households are also behavioral with respect to the ex ante real rate, such that $r_t = i_t - \bar{m}\mathbb{E}_t \pi_{t+1}$, our results are qualitatively unchanged. We discuss this case in Appendix A.2.

rational.[28]

Moreover, as established above, when output is at potential and therefore, aggregate hours are constant, household heterogeneity is irrelevant for aggregate dynamics. Hence, rational TANK (with or without type switching) and RANK models deliver identical aggregate dynamics after a TFP shock even though they yield different aggregate outcomes following monetary policy shocks.

By contrast, under behavioral expectations, TFP no longer affects the actual and the natural real rate identically: While the direct effect on the actual real rate remains unchanged, TFP moves the natural rate more strongly under behavioral expectations. Given this gap between actual and natural real rate, output falls by less than potential output after a negative TFP shock, generating a positive output gap and wage inflation, reflected in $\alpha > 0$. Importantly, for any given Taylor coefficient ϕ_w, the absolute size of the responses of the output gap, wage inflation, and nominal interest rates increases with the degree of bounded rationality.

Precisely because behavioral expectations break the neutrality of TFP shocks on the output gap, household heterogeneity is no longer irrelevant for aggregate dynamics after TFP shocks but instead amplifies their business cycle impact under behavioral expectations. The reason is that movements in the output gap redistribute across households in a similar way as after a monetary policy shock. The positive output gap after a negative TFP shock implies an increase in labor hours. Since the labor hours of H households are more sensitive to aggregate hours, their income falls less than one-for-one with aggregate income. This redistribution stabilizes demand, pushing output further away from its potential for two reasons. First, because it redistributes income to households with higher MPCs. Second, because unconstrained households reduce their precautionary savings, anticipating that they will be relatively better off if they become hand-to-mouth. The greater the heterogeneity in exposure, the more pronounced this redistribution channel and hence the stronger the impact of TFP shocks on the output gap and inflation.

A corollary of this finding is that unequal exposure adds a novel channel through which TFP shocks affect *actual* output. This becomes evident when writing the IS equation in terms of actual output instead of the output gap: $\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c r_t - [(1-\bar{m}\rho_a)\psi_c - (1-\rho_a\psi_f)]\,\widehat{a}_t$. If monetary policy keeps the real rate constant, a negative TFP shock increases output because more labor is required to produce the demanded output, which in turn, redistributes to high MPC households. In contrast, in representative agent models or more generally with equal exposure, $\chi = 1$, the last term of the IS equation is zero and output only moves due

[28] The exact irrelevance under rational-expectations in our sticky-wage, flexible-price economy hinges on the assumption of an EIS of 1. If, for example, the EIS is lower than 1, the direct impact of TFP shocks on the actual real rate remains the same, but TFP shocks have a stronger impact on the natural rate. As we show in Appendix A.10, however, the intuition and the other results in this section do not hinge on the unit EIS assumption.

to changes in the real rate, regardless of whether expectations are behavioral or rational.

Proposition 4 cleanly highlights how household heterogeneity and behavioral expectations interact in shaping the business cycle. It shows that the two are qualitatively complements as household heterogeneity only becomes relevant under behavioral expectations and that, in this case, the additional presence of household heterogeneity further lowers monetary policy's ability in stabilizing the business cycle. In fact, the amplification effects of heterogeneous exposure for the transmission of TFP shocks become even stronger, the higher the degree of bounded rationality—and vice versa. We first show this in the limiting case without idiosyncratic risk, $s = 1$, as this case allows for clean analytical results. Given $s = 1$, heterogeneous exposure χ only matters as it affects the impact on the static redistribution among households, captured by ψ_c.

Lemma 1. *In the limiting case of no idiosyncratic risk, $s = 1$, we have:*

$$\frac{\partial^2 \alpha}{\partial \bar{m} \partial \chi} = -\rho_a \frac{\partial \psi_c}{\partial \chi} \frac{(\phi_w - \rho_a)\tilde{\kappa}_w \psi_c[2(1 - \bar{m}\rho_a) - (1 - \rho_a)] + (1 - \rho_a)(1 - \bar{m}\rho_a)}{(1 - \psi_f \rho_a + \psi_c(\phi_w - \rho_a)\tilde{\kappa}_w)^3} < 0. \tag{29}$$

Thus, the impact of household heterogeneity on aggregates becomes stronger, the higher the degree of bounded rationality.

The intuition for Lemma 1 becomes apparent from Proposition 4: The higher the degree of bounded rationality, the more the output gap opens up after a TFP shock. A larger output gap and, thus, a stronger response in aggregate hours, in turn, amplifies the impact of an increase in the unequal exposure of households towards changes in aggregate hours and, thus, also on the aggregate outcomes—hence, the negative cross-derivative. Consequently, Lemma 1 implies that bounded rationality has an even greater impact on the transmission of TFP shocks when households' exposure is more uneven.

In the presence of idiosyncratic risk ($s < 1$), two additional channels influence the cross-derivative with opposing signs: A lower $\bar{m}$ increases the output gap which magnifies the reduction in precautionary-savings incentives for unconstrained households induced by an increase in χ. At the same time, however, the lower $\bar{m}$, the less unconstrained households respond to the reduction in their precautionary-savings incentives induced by an increase in χ. Hence, the overall sign of the cross derivative is no longer unambiguous, but rather depends on the parameter values, particularly those of the type-switching parameter, $(1-s)$. Yet, Appendix A.9 shows that for the range of reasonable values of this idiosyncratic risk parameter, the sign of the cross-derivative remains negative. In Section 4, we rely on an empirically-disciplined idiosyncratic risk process for the quantification of the interaction in our quantitative behavioral HANK model and we find that this interaction is quantitatively substantive.

3.4 Lower TFP or higher G: An equivalence result

We now show that the insights on how the joint presence of cognitive discounting and heterogeneous exposure amplifies the business cycle effects of TFP shocks carry over to government spending shocks. Importantly, while demand now exceeds productive capacity because of an expansion in aggregate demand, it again originates outside households' consumption. Consequently, it leaves the central bank with a similar stabilization task. In fact, as we detail in Appendix A.8, an increase in government spending financed by a labor income tax moves *potential consumption*—consumption under flexible prices and wages—and hence the natural rate, identically as a negative TFP shock. The reason is that in the flexible-price, flexible-wage economy, hours remain constant after both shocks such that potential consumption falls one-for-one with the size of the respective shock.

With sticky wages and flexible prices, however, the impact of both shocks on the real rate diverge. The reason is that government spending shocks do not replicate the stabilizing, direct effect of TFP shocks on the real rate (see equation (28)). If we assume that monetary policy mimics this direct effect by directly responding to government spending shocks, g_t, as follows:

$$i_t = \phi_w \pi_t^w + \phi_g g_t, \tag{30}$$

with $\phi_g = (1 - \rho_g)$, where ρ_g is the persistence of the government spending shock, the two shocks have exactly the same impact on the output gap and, thus, on wage inflation, as the following Lemma shows.[29]

Lemma 2. *If $\phi_g = (1 - \rho_g)$, the dynamics of the output gap in response to government spending shocks, g_t, that are financed via labor-income taxes, is given by $\widehat{x}_t = \alpha_g g_t$, where*

$$\alpha_g \equiv \frac{(1 - \bar{m}\rho_g - \phi_g)\psi_c}{1 - \psi_f \rho_g + \psi_c \tilde{\kappa}_w (\phi_w - \rho_g)} = \frac{\rho_g (1 - \bar{m})\psi_c}{1 - \psi_f \rho_g + \psi_c \tilde{\kappa}_w (\phi_w - \rho_g)} \tag{31}$$

which is identical to the output gap response to negative TFP shocks, α, if $\rho_a = \rho_g$.

From Lemma 2, it then directly follows that unequal exposure, cognitive discounting, and their interaction shape the output gap and the inflation response to a positive government spending shock in exactly the same way as after a negative TFP shock.

If, instead, monetary policy does not directly respond to government spending, i.e., $\phi_g = 0$, the impact on the output gap and wage inflation differ across the two shocks, as the effects of government spending shocks are more pronounced. Yet, the impact of cognitive

[29] The correction term in the Taylor rule, ϕ_g, exactly mirrors the correction term necessary to achieve equivalence between the sticky price and sticky wage model after a TFP shock as we show in Appendix A.7. In other words, the sticky wage model behaves equivalently after a positive government spending shock as the sticky price model after a negative TFP shock. A corollary is that negative TFP and positive government spending shocks have identical effects in the sticky price, flexible wage model with or without any correction term in the Taylor rule.

discounting, unequal exposure and their interaction on inflation and the output gap remain qualitatively unchanged, as we show in Appendix A.8.[30]

3.5 Limitations of the behavioral TANK model

While it allows for clean analytical insights, the behavioral TANK model naturally has some limitations when it comes to the quantification of the impact of the joint presence of household heterogeneity and cognitive discounting. First, the limited household heterogeneity structure prevents us from matching empirical evidence on households idiosyncratic income risk as well as on their unequal exposure. Second, because we focus on the zero-liquidity equilibrium, higher interest rates neither impact households' asset income nor the government budget and, thus, taxes. Third, fiscal policy is kept very stylized implying that the behavioral TANK model abstracts from meaningful interactions between monetary and fiscal policy, which has been shown to matter quite strongly in quantitative HANK models (Kaplan et al., 2018; Alves et al., 2020). Finally, the behavioral TANK model features only one nominal rigidity. Consequently, the central bank could achieve divine coincidence in the behavioral TANK after a TFP shock by aggressively increasing interest rates, which is not feasible anymore if both prices and wages are rigid (Erceg et al., 2000). Hence, we now turn to the full quantitative HANK model which overcomes these limitations.

4 Quantitative Behavioral HANK Model

We now relax the specific calibration assumptions imposed for the analytical results and analyze a quantitative version of the model with non-degenerate wealth and income distributions, as well as sticky wages and sticky prices. The key insights are, first, that the qualitative nature of the interaction carries over to the quantitative version of our model: the complementarity between behavioral expectations and unequal exposure of households amplifies inflation caused by demand exceeding productive capacity for reasons outside of households' consumption decisions. Second, this amplification is quantitatively powerful: in our baseline calibration, inflation increases 93% more strongly in response to a negative TFP shock than in the representative-agent, rational-expectations benchmark, with the interaction term between household heterogeneity and behavioral expectations accounting for 41% of this amplification. We then conduct a set of robustness checks showing that this amplification mechanism is robust and applies quite generally.

[30] While we abstract from debt-financed government spending shocks in the behavioral TANK model, we show in Appendix Figure 7 by means of our quantitative behavioral HANK model, that the intuition extends to deficit-financed government spending as well.

Table 1: Calibration

Parameter	Description	Value
$\bar{r}$	Steady-state real rate (annualized)	2%
$1/\gamma$	Intertemporal elasticity of substitution	0.5
φ	Inverse of Frisch elasticity	2
β	Discount factor	0.9517
$\bar{m}$	Cognitive discounting	0.81
ζ	Unequal-exposure parameter	-1.00
e_{it}	33-state earnings process	as in Kaplan et al. (2018)
μ_{D_1}	Dividend-share parameter	3
B^G/Y	Government debt (share of annual GDP)	80%
κ_p	Slope of price Phillips curve	0.096
$\bar{\kappa}_w$	Slope of wage Phillips curve	0.015
τ_P	Tax progressivity	0.9
τ_L	Tax level	0.01
ϑ	Debt feedback	0.45
ϕ	Taylor-rule coefficient	2.0

4.1 Calibration

We build on a standard one-asset calibration in the HANK literature extended by unequal exposure of households and behavioral expectations. Table 1 summarizes our baseline calibration. We follow McKay et al. (2016) and set the IES to $1/\gamma = 0.5$, the inverse of the Frisch elasticity to $\varphi = 2$, and the discount factor β to match a steady state real rate of 2% (annualized). In contrast to Section 3, we now abstract from differences in time discounting, $\beta(e_{i,t}) = \beta$ for all $e_{i,t}$, such that borrowing constraints only bind for endogenous reasons. We calibrate the idiosyncratic skill process with a discretized version of the idiosyncratic process in Kaplan et al. (2018). This process captures both persistent and transitory income risk of households in a 33-state Markov chain. We set the government debt level to 80% of annual GDP which results in an aggregate MPC of 0.21 which lies in the middle range of the empirical estimates between 0.15 and 0.25. As in Wolf (2025, 2023), we assume the following functional form for dividend share endowments:

$$d_{it} = \begin{cases} 0, & \text{if } e^p_{it} \leq e_p, \\ \mu_{D_0} \left(e^p_{it} - e_p\right)^{\mu_{D_1}} d_t, & \text{otherwise,} \end{cases}$$

with e^p_{it} denoting the persistent component of households' idiosyncratic productivity. The cutoff e_p, as well as μ_{D_0}, and μ_{D_1}, are set to ensure that (i) the bottom half of households receives no dividends consistent with the illiquid wealth distribution in the 2016 SCF, (ii) that the top 10 percent of households' share of total dividends is consistent with total illiquid wealth in Kaplan et al. (2018), and (iii) that $\int_0^1 d_{it} di = d_t$.

To capture the fact that higher MPC households tend to be on average more exposed to aggregate income changes induced by monetary policy, we assume that the union allocates labor hours according to the following allocation rule proposed in Auclert and Rognlie (2020):

$$N_{i,t} = N_t \frac{(e_{i,t})^{\zeta \log \frac{N_t}{N}}}{\mathbb{E}[e^{1+\zeta \log \frac{N_t}{N}}]}. \tag{32}$$

We set $\zeta = -1.00$ to match a covariance between the gross earnings elasticities of workers after output fluctuations induced by demand shocks and their transitory MPCs of 0.06 as estimated for the U.S. in Patterson (2023). Because the estimates in Patterson (2023) are not conditional on demand shocks, but on output fluctuations in general, we consider as robustness checks also a higher and a lower degree of unequal exposure in Section 4.4. Specifically, we set $\zeta^{high} = -2.00$ matching the headline covariance result of 0.1 in Patterson (2023) using MPC estimates for a more persistent income shock, and we also consider $\zeta^{low} = -0.43$ to target a covariance of 0.02 as estimated in Bilbiie et al. (2025) using Norwegian data.

We show how we can estimate households' underreaction to aggregate news and how we can infer the cognitive discounting parameter $\bar{m}$ from these estimates in Appendix B. In short, we regress survey expectations about inflation and unemployment in time $t+1$ and their realizations in time $t+1$ on realizations in t and use the ratio of the two regression coefficients to obtain an estimate for $\bar{m}$. Intuitively, when expectations about time $t+1$ co-move less strongly with information arriving in time t than realizations in time $t+1$, households underreact to aggregate news governed by $\bar{m}$ below one. While this approach relies on the definition of cognitive discounting, it does not require any specific assumptions about the data generating process or households' perceived law of motion. Our point estimates for the average $\bar{m}$ range from 0.51 to 0.81. As we find at most modest heterogeneity in rationality, we maintain the assumption in our baseline that all households have the same degree of rationality. For our baseline calibration, we use the upper bound of our empirical estimates, $\bar{m} = 0.81$. We also consider $\bar{m}^{low} = 0.6$ which is closer to the lower bound of our point estimates for $\bar{m}$ as well as $\bar{m}^{high} = 0.9$ in Section 4.4. Overall, these choices are conservative deviations from FIRE compared to our estimates. Moreover, the evidence in Roth et al. (2021)—which provides an upper bound for households' rationality with respect to monetary policy announcements, as it relies on hypothetical scenarios in which households receive perfect information about future interest rate changes and are hence by definition fully informed—does also not imply cognitive discounting degrees larger than 0.9. We also consider the case of heterogeneous degrees of rationality in Section 4.5, motivated by a small positive correlation between rationality and income in our estimates.

We assume a standard price adjustment frequency of 0.25 such that the average price duration is one year. This implies a slope of the Phillips curve of $\kappa_p = 0.096$. We set the wage rigidity such that the slope of the wage Phillips curve is equal to $\bar{\kappa}_w = 0.015$, consistent with the conventional view in the literature that wages are more rigid than prices (Auclert et al., 2024). This ensures that dividends are mildly procyclical, implying a correlation of 0.1 between dividends and output following a monetary policy shock with persistence $\rho_{MP} = 0.61$. In robustness checks, we also show results for higher and lower degrees of nominal rigidities.

We set the tax progressivity parameter to $\tau_P = 0.9$ as in Bayer et al. (2024) capturing the average progressivity in U.S. taxes post World War II. Instead, setting $\tau_P = 0.82$ as in Auclert et al. (2024) or assuming only linear taxes by setting $\tau_P = 1$ does not materially affect our results, as we show in Section 4.6. We then set τ_L to ensure that the government budget constraint holds. We set $\vartheta = 0.45$, implying that the government runs deficits in the short run when the government budget situation worsens, with a somewhat delayed repayment of government debt. In robustness exercises, we also consider a more delayed tax repayment by setting $\vartheta = 0.3$, as well as a faster repayment by setting $\vartheta = 0.9$.

Regarding monetary policy, we assume the following Taylor rule:

$$i_t = \rho_i i_{t-1} + (1 - \rho_i)\phi\pi_t, \tag{33}$$

setting $\phi = 2.0$ as in Bayer et al. (2024). In our baseline, we set $\rho_i = 0$ to abstract from interest rate smoothing in the Taylor rule, but we also consider the case in which monetary policy reacts more gradually by assuming $\rho_i = 0.8$ as in Bayer et al. (2024). We also show results with a more hawkish central bank by setting $\phi = 3.0$ and a more dovish central bank by setting $\phi = 1.5$. We show in Section 4.6 that a Taylor rule that responds to wage inflation instead of price inflation does not materially affect our results.

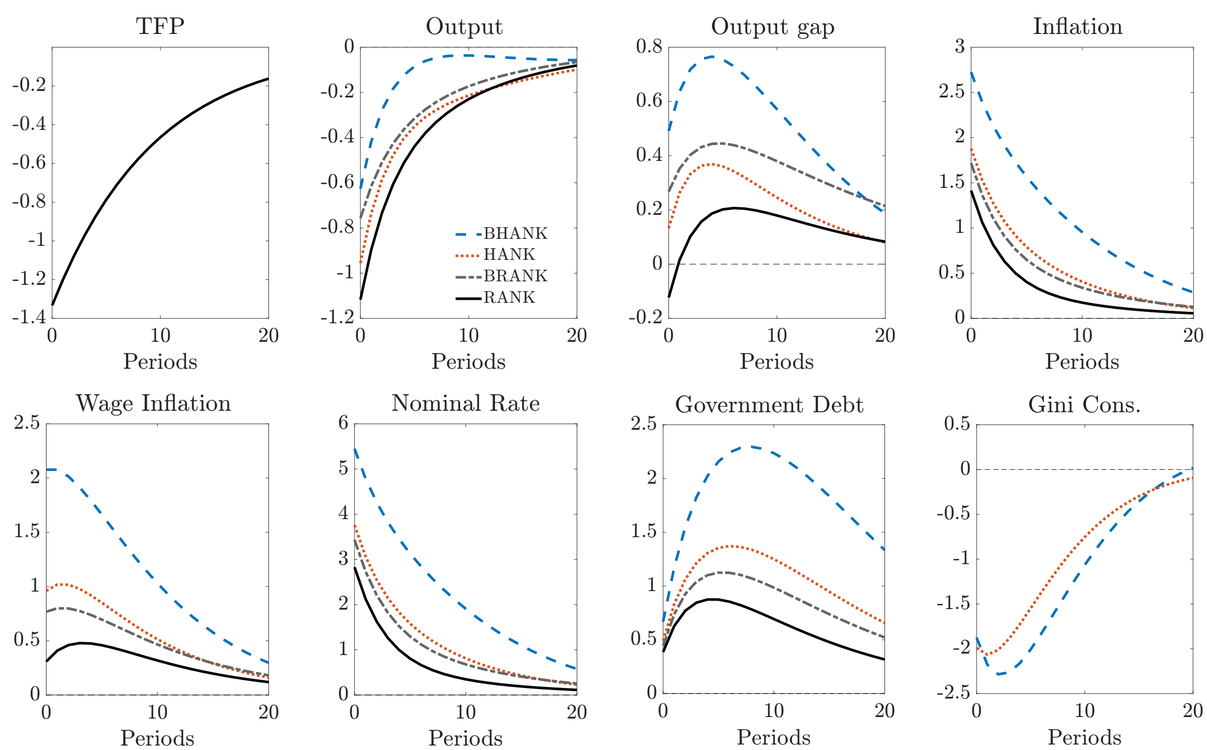
For the TFP shock process, we assume a persistence of $\rho_a = 0.9$ in our baseline, but also consider TFP shocks with lower ($\rho_a = 0.7$) and higher ($\rho_a = 0.95$) persistence. We then compute the linearized perfect-foresight transition dynamics around the deterministic stationary equilibrium of the model using the sequence-space method of Auclert et al. (2021).

Transmission of monetary policy. Appendix D.1 confirms that our analytical insights from Section 3 regarding the transmission of monetary policy carry over to our quantitative behavioral HANK model. Specifically, monetary policy is largely transmitted via indirect effects. These indirect effects are particularly pronounced because expansionary monetary policy redistributes towards high-MPC households. In addition, our behavioral HANK model does not suffer from the forward guidance puzzle, but instead the effectiveness of announced real rate changes strongly decays with their horizon. In contrast, the rational version of our model exacerbates the forward guidance puzzle as the effectiveness of real rate changes strongly increases in their horizon. Our quantitative behavioral HANK model therefore provides a suitable framework to understand the effectiveness of monetary policy in stabilizing business cycle shocks.

4.2 Quantifying the impact of household heterogeneity and cognitive discounting

We now use our model to quantify the impact of the joint presence of household heterogeneity and cognitive discounting on the effectiveness of monetary policy to stabilize the economy after business cycle shocks. Given our analytical insights in Section 3, we mainly focus on TFP shocks that, *ceteris paribus*, cause an excess in demand by temporarily reducing the productive capacity. Figure 1 compares the impulse responses in our behavioral HANK model (labeled *BHANK*, shown by the blue dashed lines) to three different models. First, to its rational-expectations counterpart, which assumes that households are fully rational by setting $\bar{m} = 1$ (*HANK*, orange dotted lines). Second, to the representative-agent version of the behavioral model which retains $\bar{m} = 0.81$ but replaces household heterogeneity by a representative agent (*BRANK*, black dashed-dotted lines). And third, to the representative-agent, rational-expectations version which shuts down household heterogeneity and sets $\bar{m} =$ 1, (*RANK*, black solid lines). Everything else is held constant across the four models.

Figure 1: Inflationary supply shock: Baseline



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

Qualitatively, the impulse responses are aligned across the four models: a negative TFP shock reduces potential output (not shown),[31] thereby increasing the output gap requiring firms to produce above capacity which induces inflationary pressure to prices and wages.[32] The central bank leans against the inflationary pressure by increasing interest rates to bring down consumption and, thus, actual output. Due to the higher interest rate payments and lower tax income, government debt increases. Consistent with our empirical evidence in Appendix C, consumption inequality, measured by the Gini index, decreases both in the rational as well as in the behavioral HANK model. While higher interest rates redistribute to relatively consumption-rich households, this effect on consumption inequality is dominated by the increase in labor hours which redistributes to relatively consumption-poor households via their labor earnings.

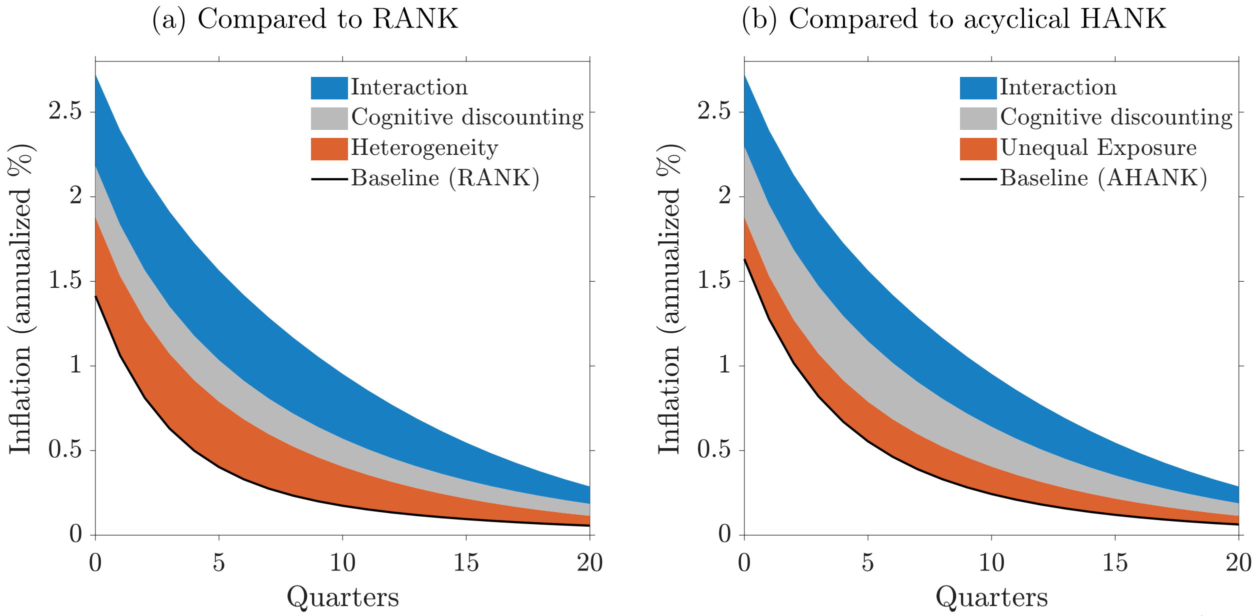
Quantitatively, however, there are stark differences across the four models. In particular, the effects of the TFP shock are most pronounced in the behavioral HANK model in line with our analytical results in Section 3. Because higher expected future interest rates are less effective in reducing current demand, actual output falls less. This pushes up the output gap and increases labor hours more and, thus, redistributes toward households with high MPCs. While the behavioral RANK and the rational HANK model each feature one of these two mechanisms in isolation, only the behavioral HANK model embeds both. Consequently, the output gap, wage and price inflation react most strongly in the behavioral HANK model—despite interest rates rising the most. The strong increase in interest rates enlarges the fiscal footprint of the monetary tightening as higher interest rates increase the cost of government debt which is financed by issuing more debt in the short-run.

Figure 1 shows that in the behavioral HANK model, inflation increases most strongly. Compared to the RANK model which abstracts from household heterogeneity and cognitive discounting, inflation increases 93% more on impact. The left panel of Figure 2 decomposes this additional inflation increase in our behavioral HANK model relative to the RANK benchmark (black solid line) into three components. First, the orange area isolates the effects of household heterogeneity, measured as the inflation response in HANK minus the inflation response in RANK. Second, the gray area captures the contribution of cognitive discounting alone, measured as the inflation response in the behavioral RANK minus the inflation response in RANK. Third, the blue area reflects the interaction term, that is, the part of the inflation difference between the behavioral HANK and RANK that cannot be explained by heterogeneity or cognitive discounting in isolation, but which arises in their joint presence.

[31]As in the analytical framework, we define potential output as the output in the respective economy absent nominal rigidities.

[32]Note that the quantitative version of the RANK model used in this Section differs in four ways from the analytical RANK model analyzed in Section 3 which breaks the TFP irrelevance: first, the Taylor rule targets price inflation, second, the EIS is 0.5 instead of 1, third it has time-varying distortionary taxes, and fourth, it incorporates both sticky wages and sticky prices. Under our calibration, monetary policy depresses output briefly below potential output in RANK, which is reflected in a negative output gap on impact.

Figure 2: Decomposition of the additional inflation increase



Note: This figure shows the decomposition of the additional inflation increase in the behavioral HANK compared to the RANK model in panel (a) and compared to the rational, acyclical HANK model in panel (b). See main text for the computation of the decomposition.

The left panel of Figure 2 shows that all three components increase inflation, with the blue area being the largest, indicating that the largest part of the inflation increase is due to the complementarity between household heterogeneity and cognitive discounting. On impact, this complementarity increases inflation by 38% of the RANK inflation response (which corresponds to 41% of the total amplification relative to RANK). By comparison, household heterogeneity alone increases inflation by 33% on impact and cognitive discounting increases it by 21%. Hence, abstracting from either household heterogeneity or cognitive discounting reduces the inflation response by 71% and 59% of the RANK impact, respectively.

Increase in government spending. Figure 7 in the Appendix shows impulse responses after an increase in government spending. In this case, demand exceeds productive capacity due to an expansion in demand rather than a reduction in productive capacity. While the exact equivalence between a positive government spending shock and a negative TFP shock breaks down in the quantitative model, Figure 7 demonstrates that the qualitative insights are the same: the output gap, inflation, interest rates and the government debt level all rise much more strongly in our behavioral HANK model than they do in RANK or in models that consider either household heterogeneity or cognitive discounting in isolation. Hence, our analytical insights from the behavioral TANK model are also quantitatively important for government spending shocks: monetary policy is less effective in stabilizing inflation after higher government spending.

4.3 Isolating the interaction between unequal exposure and cognitive discounting

In contrast to the behavioral TANK model discussed in Section 3, the quantitative model gives rise to additional channels besides unequal exposure in labor earnings through which household heterogeneity affects aggregate outcomes. We now isolate the impact of unequal exposure by comparing our baseline BHANK model to an *acyclical* HANK model in which we switch off the unequal exposure of households' labor earnings. In particular, we set $\zeta = 0$ in the union's allocation rule (32) such that hours are allocated uniformly across households also outside the stationary equilibrium.

The right panel of Figure 2 benchmarks the inflation response in our baseline model relative to the rational version of the acyclical HANK model after the same negative TFP shock as in Figure 1. The black line shows that inflation in the rational, acyclical HANK increases by 1.6 percentage points on impact. Relative to this benchmark, inflation increases 67% more strongly in our baseline model. We decompose this amplification again into three components. First, the orange area isolates the effects of unequal exposure, measured as the inflation response in HANK minus the inflation response in the acyclical HANK. Second, the gray area captures the contribution of cognitive discounting alone, measured as the inflation response in the behavioral acyclical HANK minus the inflation response in the acyclical HANK. Third, the blue area captures the interaction between unequal exposure and cognitive discounting. On impact, this interaction increases inflation by 26% of the acyclical HANK inflation response. In comparison, unequal exposure in isolation increases inflation by 15% on impact and cognitive discounting increases it by 26%. Again, abstracting from either one of the two model features reduces the inflation response by 41% and 52% of the acyclical HANK inflation response, respectively.

There are three main takeaways from comparing the right panel of Figure 2 to its left panel. First, the interaction between unequal exposure in isolation and cognitive discounting is itself powerful, consistent with the analytical insights in Section 3. Second, in the quantitative model, unequal exposure in labor earnings is not the only reason why household heterogeneity amplifies the inflation response to a negative TFP shock. Inflation is also higher because profits fall somewhat more than labor income—redistributing toward higher-MPC households—and because the rise in government debt boosts demand given non-Ricardian households (Auclert et al., 2024; Angeletos et al., 2024). And third, cognitive discounting also interacts positively with these additional amplification channels generated by household heterogeneity: its incremental effect is larger in the acyclical HANK (26%) than it is in the RANK model (21%). The reason is that under behavioral expectations, real rates increase more, which amplifies the debt response and the behavioral expectations make households even more non-Ricardian.

4.4 Varying unequal exposure and cognitive discounting

Unequal exposure and cognitive discounting as well as their interaction can severely limit the effectiveness of monetary policy in stabilizing the business cycle. To obtain a range of plausible values regarding their influence on the inflationary response to TFP shocks, we now consider alternative calibrations for $\bar{m}$ and ζ discussed in Section 4.1. Table 2 reports the on-impact increase in inflation across these alternative calibrations of the behavioral HANK model, measured relative to the RANK benchmark. It also decomposes the amplification into the contribution of household heterogeneity alone, the contribution of cognitive discounting alone, and their interaction, analogous to Figure 2. We here focus on benchmarking our results against the RANK inflation response. Table 7 in the Appendix shows that the insights carry over if we instead benchmark the results against the inflation response in the acyclical HANK.

Table 2: Varying unequal exposure and cognitive discounting

	Comparison with RANK			
Scenario	Amplification of π	Heterogeneity	Cognitive discounting	Interaction
Baseline	93%	33%	21%	38%
Less unequal exposure	65%	23%	21%	20%
More unequal exposure	139%	46%	21%	71%
Higher $\bar{m}$	67%	33%	13%	21%
Lower $\bar{m}$	141%	33%	35%	73%
Het. $\bar{m}$, specification 1	96%	33%	25%	38%
Het. $\bar{m}$, specification 2	95%	33%	32%	30%

Note: The table shows the on-impact amplification of inflation in the behavioral HANK model compared to the RANK model, as well as the respective on-impact contributions due to household heterogeneity, cognitive discounting, and their interaction. See the main text for the computation of this decomposition. *Less unequal* refers to $\zeta = -0.425$ (instead of $\zeta = -1.0$); *more unequal* to $\zeta = -2.0$. Higher $\bar{m}$ refers to $\bar{m} = 0.90$ (instead of $\bar{m} = 0.81$), lower to $\bar{m} = 0.60$. *Het. $\bar{m}$, specification 1* refers to the case in which we target our empirical results based on inflation expectations and *Het. $\bar{m}$, specification 2* refers to the case in which we target our empirical results based on unemployment expectations.

We begin with the two alternative calibrations for households' unequal exposure. With lower unequal exposure, the redistribution towards high-MPC households following a decline in TFP is muted. Consequently, the fall in actual output is less dampened by the presence of household heterogeneity. This mutes both the direct effect of household heterogeneity and the interaction between household heterogeneity and cognitive discounting on inflation. Consequently, the amplification of inflation is somewhat reduced compared to our baseline, but still sizable with 65% (instead of 93%).

By contrast, with more unequal exposure, there is more redistribution towards high-MPC households amplifying the impact of TFP shocks on the business cycle: inflation rises 139% more than in RANK, driven largely by a strong complementarity effect, which now accounts for more than half of the total amplification. These results quantify the analytical insights from our behavioral TANK model in Proposition 4 and Lemma 1: More unequal exposure makes inflation stabilization substantially more difficult for monetary policy after a TFP shock, and the more unequally exposed households are, the larger is the impact of a given degree of cognitive discounting, as captured by the growing significance of the interaction term.

Turning to households' underreaction to aggregate news, we begin by considering a higher degree of rationality, i.e., a higher $\bar{m}$. As households are more rational, expected higher interest rates in the future are more effective in reducing demand today, muting the impact on the output gap. Compared to our baseline calibration, this weakens both the contribution of cognitive discounting alone but also the interaction with household heterogeneity as there is less redistribution towards high-MPC households in general equilibrium. Still, inflation rises 67% more than in RANK. By contrast, with a lower degree of rationality, the effects of cognitive discounting—and especially the interaction—are more pronounced: inflation increases 141% more than in RANK, with the interaction accounting for more than half of it. Again, the insights from our behavioral TANK are quantitatively important: the less rational households are, the more difficult it is for monetary policy to stabilize inflation after a TFP shock. Moreover, the less rational households are, the larger is the impact of a given degree of unequal exposure as captured by the growing significance of the interaction term.

4.5 Heterogeneous cognitive discounting

So far, we have assumed that all households exhibit the same degree of rationality. Yet, as we show in Appendix B, while underreaction is found across all income groups, the data suggests that higher income households deviate somewhat less from rational expectations. To capture this, we now assume that households' degree of cognitive discounting is a decreasing function of the persistent component of their productivity level e^p_{it}. We then calibrate this function to match our estimated degrees of underreaction by income quartile—once using our estimates based on inflation expectations and once using those for unemployment expectations.[33]

Figure 8 in the Appendix compares the models with heterogeneous degrees of bounded rationality to our baseline behavioral HANK model for the same TFP shock. The results are virtually identical across all three specifications. Consequently, their overall amplification relative to RANK is also essentially the same, as shown in Table 2.

Two opposing effects across the three different specifications almost exactly offset each other: first, the estimates with heterogeneous cognitive discounting imply a lower average $\bar{m}$ than our conservative baseline calibration. Second, the positive correlation between $\bar{m}$ and income somewhat mutes the aggregate role of cognitive discounting, *ceteris paribus*, because

[33]Specifically, we assign the estimated quartile-specific $\bar{m}$ to the idiosyncratic productivity grid point containing the respective income-quartile cutoff and linearly interpolate $\bar{m}$ across the remaining grid points.

stronger cognitive discounting is concentrated among lower-income households who are more likely not on their Euler equation.

4.6 Robustness

Finally, we conduct robustness checks for our specifications of monetary and fiscal policy, the wage Phillips curve, the degree of nominal rigidities, and the persistence of the TFP shock. Table 3 summarizes the results by comparing inflation in the behavioral HANK model to the rational RANK benchmark and decomposing the additional amplification as described in Section 4.2.

Table 3: Robustness

	Comparison with RANK			
Scenario	Amplification of π	Heterogeneity	Cognitive discounting	Interaction
Baseline	93%	33%	21%	38%
Persistence in Taylor	128%	42%	29%	57%
More dovish MP	126%	60%	17%	49%
More hawkish MP	67%	16%	24%	27%
Slower debt rep.	122%	40%	21%	61%
Faster debt rep.	59%	20%	21%	18%
No progressivity	102%	35%	20%	47%
More progressivity	85%	31%	22%	31%
Transfer repayment	60%	23%	15%	22%
No ξ_t in PC	74%	28%	21%	25%
No ξ_t and no taxes in PC	62%	30%	15%	17%
π^w-targeting	115%	50%	30%	35%
More persistent shock	129%	30%	47%	51%
Less persistent shock	61%	38%	1%	21%
More nom. rigidities	95%	48%	10%	37%
Less nominal rigidities	81%	22%	28%	31%
More sticky wages	81%	32%	16%	33%

Note: The table shows the on-impact amplification of inflation in the behavioral HANK model compared to the RANK model, as well as the respective on-impact contributions due to household heterogeneity, cognitive discounting, and their interaction. See the main text for the exact parameter changes analyzed and see Section 4 for the computation of the decomposition.

We start by examining the role of the specification of the monetary policy rule. First, we add persistence to the Taylor rule, setting $\rho_i = 0.8$ as in Bayer et al. (2024). In this case, inflation rises 128% more than in RANK. With persistence in the Taylor rule, monetary policy relies more on expectations of higher future interest rates to restrain current demand. But this strategy is markedly less effective when households are behavioral and particularly so, when simultaneously accounting for household heterogeneity, as reflected in the growing contribution of the interaction term. The muted expectations channel implies even larger current output gaps, reinforcing the redistribution towards high-MPC households and, thus, the importance of household heterogeneity.

Regarding the central bank's aggressiveness, amplification is larger under a more dovish rule ($\phi = 1.5$) and smaller under a more hawkish one ($\phi = 3.0$). Notably, the degree of hawkishness shifts the balance between the heterogeneity-only and cognitive-discounting-only components, while the interaction term remains important in both cases. Intuitively, a dovish stance implies that monetary policy leans less strongly against the overheating of the economy, leading to larger output gaps which elevates the role of unequal exposure. In contrast, a hawkish monetary policy raises real rates at all horizons more forcefully, effects that are discounted when households are behavioral.

If the central bank targets wage instead of price inflation as in our behavioral TANK model, the amplification becomes somewhat more pronounced. Intuitively, as wages are more sticky than prices, monetary policy leans less against the overheating of the economy by reducing output and the output gap which especially increases the role of household heterogeneity.

We then turn to the specification of fiscal policy. Given the strong response of government debt, assumptions about its repayment should impact aggregate outcomes (Kaplan et al., 2018; Alves et al., 2020; Angeletos et al., 2024; Bayer et al., 2025). We therefore run robustness with respect to our fiscal specification, starting with varying the speed of repayment. If we assume that debt is repaid more slowly by setting $\vartheta = 0.3$ instead of $\vartheta = 0.45$, inflation reacts even more strongly: on impact it rises 122% more than in RANK, with roughly half of the extra inflation coming from the interaction term. Intuitively, because households are non-Ricardian, delaying repayment is expansionary, and the effect is especially pronounced in the behavioral HANK model, where government debt responds most strongly. By contrast, with faster repayment, $\vartheta = 0.9$, inflation rises 59% more than in RANK and, thus while the amplification remains sizable, it becomes weaker than in our baseline. Still, the interaction between household heterogeneity and cognitive discounting accounts for about a third of the overall amplification. The stark differences across different repayment speeds indicate that the role of fiscal policy and its interaction with monetary policy—which has been shown to be materially important in HANK models—becomes even more important when heterogeneous households are also behavioral.

We also analyze robustness with respect to the debt repayment instrument, holding the repayment speed fixed. First, we vary the degree of tax progressivity. Increasing progressivity ($\tau_P = 0.82$ as in Auclert et al. (2024)) or abstracting from progressivity altogether ($\tau_P = 1$), has only a small impact. With higher progressivity, amplification is 85%, whereas with linear taxes it is 102%. Intuitively, progressive taxes act as an automatic stabilizer, somewhat reducing the role of unequal exposure and muting the inflationary response. Second, we assume that repayment occurs through a reduction in lump-sum transfers. In this case, the overall amplification is 60%. Hence, monetary policy becomes more effective if the government adjusts its budget through transfers rather than through taxes, in line with Kaplan et al. (2018).

In the quantitative behavioral HANK model, TFP shocks have two additional indirect effects on inflation through the wage Phillips curve (10), which we now isolate. First, heterogeneous responses in labor hours affect the union's target wage. Shutting down this channel by setting $\widehat{\xi}_t = 0$, somewhat mutes the amplification to 74% (instead of 93%), indicating that the composition of the additional labor input is inflationary. Second, via their impact on the government budget, TFP shocks induce an adjustment in distortionary taxes. If, in addition to setting $\widehat{\xi}_t = 0$, we treat these tax changes as non-distortionary in the wage Phillips curve, the amplification is further reduced but remains sizable at 62% relative to RANK inflation.

Varying the persistence of the shock shows that the amplification grows in the persistence. For example, with $\rho_a = 0.95$ instead of 0.9, the inflation response is 126% higher than in RANK. The roles of cognitive discounting and the interaction term grow particularly strongly. In contrast, with lower persistence ($\rho_a = 0.7$), inflation is 61% higher relative to RANK. With lower persistence, the role of cognitive discounting is weaker. Still, the interaction term accounts for more than a third of the total amplification in this case.

Lastly, we analyze the role of nominal rigidities for our results. We first consider two specifications varying the overall degree of rigidity, one in which we halve the slope of the price Phillips curve and one in which we double it. In both cases, we recalibrate wage stickiness so that the correlation between dividends and output following a monetary policy shock is the same as in our baseline. While the inflation dynamics across all models differ noticeably from the baseline, the amplification compared to RANK is quite similar across the different degrees of nominal rigidity. What changes, however, is the relative importance of household heterogeneity and cognitive discounting taken in isolation, whereas their interaction remains sizable in all cases. With higher nominal rigidities, inflation responds less, such that monetary policy leans less aggressively against inflation by reducing output and the output gap. This raises the importance of household heterogeneity and reduces the role of cognitive discounting. By contrast, with more flexible prices and wages, inflation reacts more strongly and, hence, monetary policy reduces output and the output gap more strongly, diminishing the role of household heterogeneity and increasing the role of cognitive discounting. Finally, we consider a different ratio of price vs. wage stickiness by reducing the slope of the wage Phillips curve from 0.015 to 0.012 but keeping the slope of the price Phillips curve fixed. This increases the procyclicality of profits conditional on monetary policy shocks. We find that, in this case, the overall amplification slightly decreases to 81% with the interaction still accounting for the largest component.

5 Conclusion

We highlight the importance to consider household heterogeneity and behavioral expectations jointly to analyze the conduct of monetary policy. We first show that extending the standard New Keynesian model to include both household heterogeneity and cognitive discounting allows the model to match recent empirical evidence on the transmission of monetary policy: monetary policy is largely transmitted through indirect, general-equilibrium channels; the effectiveness of monetary policy announcements declines with their horizon; and expansionary monetary policy redistributes toward high-MPC households. Second, precisely because the model captures these facts, it implies that monetary policy is substantially less effective in stabilizing inflation after shocks that do not originate in households' consumption and which, therefore, require a fall in consumption below its steady state level. When households are behavioral and underreact to aggregate news—as we document using survey data—the natural rate responds more strongly, widening the gap between the actual and the natural real rate. Precisely in those circumstances when the actual real rate does not track the natural rate, household heterogeneity matters as overheating labor markets and positive output gaps redistribute toward high-MPC households. Moreover, the required large movements in the real rate amplify the fiscal footprint of monetary stabilization, increasing the importance of monetary–fiscal interactions. Quantitatively, a negative TFP shock raises inflation by roughly twice as much as in a model without these features. Because of the strong complementarity between household heterogeneity and cognitive discounting, removing either feature reduces the predicted amplification by more than two thirds.

Taken together, simultaneously accounting for household heterogeneity and cognitive discounting offers a theory that can at least partly explain how seemingly modest shocks can trigger pronounced inflation surges. Our explanation—through reduced efficacy in monetary policy's demand management in the face of lowered productive capacities or higher government spending—thus complements existing work focusing on supply side explanations for inflation surges. While we show how household heterogeneity and cognitive discounting affect monetary and fiscal policy, we leave it to future research to study how policy optimally responds to different shocks in this framework. We also stop short of considering state-dependence of people's behavioral biases and how systematic changes in people's biases may interact with their exposure to economic fluctuations.

References

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

Afsar, A., J.-E. Gallegos, R. Jaimes, and E. Silgado-Gómez (2024): "A behavioral hybrid New Keynesian model: Quantifying the importance of belief formation frictions," *Economic Modelling*, 132, 106626.

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

Alves, F., G. Kaplan, B. Moll, and G. L. Violante (2020): "A further look at the propagation of monetary policy shocks in HANK," *Journal of Money, Credit and Banking*, 52, 521–559.

Amiti, M., S. Heise, F. Karahan, and A. Şahin (2023): "Inflation Strikes Back: The Role of Import Competition and the Labor Market," *NBER Macroeconomics Annual, Volume 38*.

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

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

Angeletos, G.-M., C. Lian, and C. K. Wolf (2024): "Can Deficits Finance Themselves?" *Econometrica*, 92, 1351–1390.

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

Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2021): "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models," *Econometrica*, 89, 2375–2408.

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

Auclert, A., M. Rognlie, and L. Straub (2020): "Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model," *Working paper*.

——— (2024): "The intertemporal Keynesian cross," *Journal of Political Economy*, 132, 4068–4121.

Bayer, C., B. Born, and R. Luetticke (2024): "Shocks, frictions, and inequality in US business cycles," *American Economic Review*, 114, 1211–1247.

Bayer, C., A. Kriwoluzky, G. J. Müller, and F. G. L. Seyrich (2025): "Being and consciousness: Fiscal attitudes according to HANK," Tech. rep., ECONtribute Discussion Paper.

Benchimol, J., L. Bounader, and M. Dotta (2025): "Estimating Behavioral Inattention," *Journal of Economic Behavior & Organization*, 236, 107068.

Benigno, P. and G. B. Eggertsson (2023): "It's baaack: The surge in inflation in the 2020s and the return of the non-linear phillips curve," *Working paper*.

Bewley, T. (1976): "The Permanent Income Hypothesis: A Theoretical Formulation." .

Bhandari, A., J. Borovička, and P. Ho (2024): "Survey data and subjective beliefs in business cycle models," *Review of Economic Studies*.

Bianchi, F., R. Faccini, and L. Melosi (2023): "A Fiscal Theory of Persistent Inflation," *The Quarterly Journal of Economics*.

Bianchi, F. and L. Melosi (2022): "Inflation as a fiscal limit," *2022 Jackson Hole Symposium Volume*.

Bilbiie, F. and M. Trabandt (2025): "Sticky prices or sticky wages? An equivalence result," *Review of Economics and Statistics*, 1–27.

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

——— (2025): "Monetary policy and heterogeneity: An analytical framework," *Review of Economic Studies*, 92, 2398–2436.

Bilbiie, F. O., S. M. Galaasen, R. Gürkaynak, M. Mæhlum, and K. Molnar (2025): "Hanksson," *Working paper*.

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

Blanco, A., C. Boar, C. J. Jones, and V. Midrigan (2024): "The Inflation Accelerator," *Working paper*.

Born, B., Z. Enders, M. Menkhoff, G. J. Müller, and K. Niemann (2024): "Firm Expectations and News: Micro v Macro," *Working paper*.

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

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

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

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

Cavallo, A., F. Lippi, and K. Miyahara (2024): "Large shocks travel fast," *American Economic Review: Insights*, 6, 558–574.

Cerrato, A. and G. Gitti (2022): "Inflation since covid: Demand or supply," *Working paper*.

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

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

Coibion, O., Y. Gorodnichenko, E. S. Knotek, and R. Schoenle (2023): "Average inflation targeting and household expectations," *Journal of Political Economy Macroeconomics*, 1, 403–446.

Coibion, O., Y. Gorodnichenko, and D. Koustas (2021): "Consumption inequality and the frequency of purchases," *American Economic Journal: Macroeconomics*, 13, 449–482.

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

Debortoli, D., 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. P. Giannoni, and C. Patterson (2023): "The forward guidance puzzle," *Journal of Political Economy Macroeconomics*, 1, 43–79.

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

Eggertsson, G. B. and M. Woodford (2003): "Optimal monetary policy in a liquidity trap," *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.

Farhi, E. and I. Werning (2019): "Monetary policy, bounded rationality, and incomplete

markets," *American Economic Review*, 109, 3887–3928.
Fernald, J. (2014): "A quarterly, utilization-adjusted series on total factor productivity," Federal Reserve Bank of San Francisco.
Fieldhouse, A. and K. Mertens (2023): "The returns to government R&D: Evidence from US appropriations shocks," .
Fuhrer, J. C. and G. D. Rudebusch (2004): "Estimating the Euler equation for output," *Journal of Monetary Economics*, 51, 1133–1153.
Gabaix, X. (2020): "A behavioral New Keynesian model," *American Economic Review*, 110, 2271–2327.
Gallegos, J.-E. (forthcoming): "HANK beyond FIRE: Amplification, forward guidance, and belief shocks," *International Economic Review*.
Gitti, G. (2024): "Nonlinearities in the Regional Phillips Curve with Labor Market Tightness," *Working paper*.
Guerreiro, J. (2023): "Belief disagreement and business cycles," *Working Paper*.
Hagedorn, M. (2016): *A demand theory of the price level*.
——— (2018): "Prices and inflation when government bonds are net wealth," .
Hagedorn, M., J. Luo, I. Manovskii, and K. Mitman (2019a): "Forward guidance," *Journal of Monetary Economics*, 102, 1–23.
Hagedorn, M., I. Manovskii, and K. Mitman (2019b): "The fiscal multiplier," *National Bureau of Economic Research*.
Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal tax progressivity: An analytical framework," *The Quarterly Journal of Economics*, 132, 1693–1754.
Holm, M. B., P. Paul, and A. Tischbirek (2021): "The transmission of monetary policy under the microscope," *Journal of Political Economy*, 129, 2861–2904.
Huber, P. J. (1964): "Robust estimation of a location parameter," *Annals of Mathematical Statistics*, 35: 73–101.
Ilut, C. and R. Valchev (2023): "Economic agents as imperfect problem solvers," *The Quarterly Journal of Economics*, 138, 313–362.
Imrohoroğlu, A. (1989): "Cost of business cycles with indivisibilities and liquidity constraints," *Journal of Political Economy*, 97, 1364–1383.
Jordà, Ò. (2005): "Estimation and inference of impulse responses by local projections," *American Economic Review*, 95, 161–182.
Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.
Kučinskas, S. and F. S. Peters (2022): "Measuring under-and overreaction in expectation formation," *The Review of Economics and Statistics*, 1–45.
Luetticke, R. (2021): "Transmission of monetary policy with heterogeneity in household portfolios," *American Economic Journal: Macroeconomics*, 13, 1–25.
Mankiw, N. G. (2000): "The savers-spenders theory of fiscal policy," *American Economic Review*, 90, 120–125.
Maxted, P., D. Laibson, and B. Moll (2025): "Present bias amplifies the household balance-sheet channels of macroeconomic policy," *The Quarterly Journal of Economics*, 140, 691–743.
McKay, A., E. Nakamura, and J. Steinsson (2016): "The power of forward guidance revisited," *The American Economic Review*, 106, 3133–3158.
——— (2017): "The discounted Euler equation: A note," *Economica*, 84, 820–831.
McKay, A. and J. F. Wieland (2022): "Forward guidance and durable goods demand,"

American Economic Review: Insights, 4, 106–22.

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

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

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

Pfäuti, O. (2024): "Inflation—Who Cares? Monetary Policy in Times of Low Attention," *Journal of Money, Credit and Banking*.

——— (2025): "The inflation attention threshold and inflation surges," *Working paper*.

Pfäuti, O., F. Seyrich, and J. Zinman (2024): *Beyond Bad Luck: Macroeconomic Implications of Persistent Heterogeneity In Optimism*, National Bureau of Economic Research.

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

——— (2021): "Macroeconomic fluctuations with HANK & SAM: An analytical approach," *Journal of the European Economic Association*, 19, 1162–1202.

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

Rozsypal, F. and K. Schlafmann (2023): "Overpersistence bias in individual income expectations and its aggregate implications," *American Economic Journal: Macroeconomics*, 15, 331–371.

Samarina, A. and A. D. Nguyen (2024): "Does monetary policy affect income inequality in the euro area?" *Journal of Money, Credit and Banking*, 56, 35–80.

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.

Weber, M., B. Candia, H. Afrouzi, T. Ropele, R. Lluberas, S. Frache, B. Meyer, S. Kumar, Y. Gorodnichenko, D. Georgarakos, et al. (2025): "Tell Me Something I Don't Already Know: Learning in Low-and High-Inflation Settings," *Econometrica*, 93, 229–264.

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

Wolf, C. (2023): "The missing intercept: A demand equivalence approach," *American Economic Review*, 113, 2232–2269.

——— (2025): "Interest Rate Cuts vs. Stimulus Payments: An Equivalence Result," *Journal of Political Economy*, 133, 1235–1275.

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

Appendix

A Appendix to Section 3

A.1 Calibration of the behavioral TANK model

For most of our analytical insights in Section 3, the exact calibration is immaterial. When we need a parameterization, we rely on a calibration strategy that uses standard parameters in the literature on TANK models with type switching (see, e.g., Bilbiie (2020, 2025)). We set the share of H agents to one third, $\lambda = 0.33$, and unequal exposure to $\chi = 1.35$ which implies $\psi_c = 1.2$. In case we vary these parameters, we always keep $\lambda < \chi^{-1}$. We set the probability of a U household to become hand-to-mouth next period to 5.4%, i.e., $s = 0.946$ (this corresponds to $s = 0.8$ in annual terms). This calibration results in $\delta = 1.034$, such that $\bar{m} < 0.966$ is sufficient to rule out the forward guidance puzzle (reflected in $\psi_f = \bar{m}\delta < 1$). We set $\beta(U) = 0.99$, and the slope of the Phillips Curve to $\kappa_w = 0.02$, which is the same as in Bilbiie et al. (2022) for the price inflation Phillips curve. The cognitive discounting parameter, $\bar{m}$ is set to 0.81.

A.2 Derivations

Derivation of Proposition 1. We start from the linearized Euler equation (18):

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

and we set $r_t = i_t - \mathbb{E}_t^{BR}\pi_{t+1}$ to consider the more general case of (potentially) non-rational expectations about today's real rate. We then plug in the expressions for the two types' consumption:

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

to obtain

$$\begin{aligned}
\frac{1-\lambda\chi}{1-\lambda}\widehat{y}_t + \left(1 - \frac{1-\lambda\chi}{1-\lambda}\right)\widehat{a}_t = s\mathbb{E}_t^{BR}&\left[\frac{1-\lambda\chi}{1-\lambda}\widehat{y}_{t+1} + \left(1 - \frac{1-\lambda\chi}{1-\lambda}\right)\widehat{a}_{t+1}\right] \\
&+ (1-s)\mathbb{E}_t^{BR}\left[\chi\widehat{y}_{t+1} + (1-\chi)\widehat{a}_{t+1}\right] - r_t.
\end{aligned} \tag{34}$$

Potential output is given by $\widehat{y}_t^{pot} = \widehat{a}_t$, and the output gap is defined as $\widehat{x}_t = \widehat{y}_t - \widehat{y}_t^{pot}$, so that we can express the Euler equation as:

$$\begin{aligned}
\frac{1-\lambda\chi}{1-\lambda}\widehat{x}_t + \widehat{a}_t &= s\mathbb{E}_t^{BR}\left[\frac{1-\lambda\chi}{1-\lambda}\widehat{x}_{t+1} + \widehat{a}_{t+1}\right] + (1-s)\mathbb{E}_t^{BR}\left[\chi\widehat{x}_{t+1} + \widehat{a}_{t+1}\right] - r_t \\
\Leftrightarrow \frac{1-\lambda\chi}{1-\lambda}\widehat{x}_t &= s\mathbb{E}_t^{BR}\left[\frac{1-\lambda\chi}{1-\lambda}\widehat{x}_{t+1}\right] + (1-s)\mathbb{E}_t^{BR}\left[\chi\widehat{x}_{t+1}\right] - \left[r_t + \widehat{a}_t\left(1-\bar{m}\rho_a\right)\right] \\
\Leftrightarrow \frac{1-\lambda\chi}{1-\lambda}\widehat{x}_t &= s\mathbb{E}_t^{BR}\left[\frac{1-\lambda\chi}{1-\lambda}\widehat{x}_{t+1}\right] + (1-s)\mathbb{E}_t^{BR}\left[\chi\widehat{x}_{t+1}\right] - \left[r_t - r_t^n\right]
\end{aligned}$$

where we use $\mathbb{E}_t^{BR}[\widehat{a}_{t+1}] = \bar{m}\rho_a\widehat{a}_t$, and the definition of the natural rate $r_t^n = -(1-\bar{m}\rho_a)\widehat{a}_t$. Note that $\psi_c^{-1} \equiv \frac{1-\lambda\chi}{1-\lambda}$, so that if we divide by $\frac{1-\lambda\chi}{1-\lambda}$, we obtain

$$\widehat{x}_t = s\mathbb{E}_t^{BR}\left[\widehat{x}_{t+1}\right] + (1-s)\chi\psi_c\mathbb{E}_t^{BR}\left[\widehat{x}_{t+1}\right] - \psi_c\left[r_t - r_t^n\right].$$

Collecting and rearranging the terms in front of $\mathbb{E}_t^{BR}\left[\widehat{x}_{t+1}\right]$ yields Proposition 1, with

$$\begin{aligned}
\delta &= s + (1-s)\chi\psi_c \\
&= s + (1-s)\chi\frac{1-\lambda}{1-\chi\lambda} \\
&= \frac{s(1-\chi\lambda) + (1-s)\chi(1-\lambda)}{1-\lambda} \\
&= 1 - \frac{1-\lambda\chi}{1-\lambda\chi} + \frac{s(1-\chi\lambda) + (1-s)\chi(1-\lambda)}{1-\lambda} \\
&= 1 + (\chi-1)\frac{1-s}{1-\lambda\chi},
\end{aligned}$$

as stated in the proposition.

Derivation of Proposition 3. We start by using the derived IS equation in Proposition 1 and replace the real rate with $r_t = \phi\pi_t^w - \mathbb{E}_t^{BR}[\pi_{t+1}]$, which follows from the definition of the real rate and the Taylor rule. We consider here the general case with (potentially) behavioral expectations about today's real rate, defining $\mathbb{E}_t^{BR}[\pi_{t+1}] = \bar{m}_\pi\mathbb{E}_t\pi_{t+1}$. We use $\bar{m}_\pi$ instead of $\bar{m}$ here, because in most cases we will assume that agents have rational real rate expectations (as in Gabaix (2020)), captured by $\bar{m}_\pi = 1$, while assuming that agents are behavioral with respect to their future marginal utility (in deviations from the stationary equilibrium). We also write the natural rate in terms of TFP shocks: $r_t^n = -(1-\bar{m}\rho_a)\widehat{a}_t$. We then arrive at the IS equation:

$$\widehat{x}_t = \delta\mathbb{E}_t^{BR}\left[\widehat{x}_{t+1}\right] - \psi_c\left[\phi\pi_t^w - \mathbb{E}_t^{BR}[\pi_{t+1}] + (1-\bar{m}\rho_a)\widehat{a}_t\right].$$

Note that the only state variable is $\widehat{a}_t$.[34] From the wage Phillips Curve it follows that we can express wage inflation as follows:

$$\begin{aligned}
\pi_t^w &= \kappa_w \widehat{x}_t + \beta \mathbb{E}_t \pi_{t+1}^w &(35)\\
&= \kappa_w \widehat{x}_t + \beta \rho_a \pi_t^w &(36)\\
&= \frac{\kappa_w}{1-\beta\rho_a}\widehat{x}_t &(37)\\
&= \tilde{\kappa}_w \widehat{x}_t, &(38)
\end{aligned}$$

where the second line follows from the fact that $\widehat{a}_t$ is the only state variable and, thus, in this linearized economy all endogenous variables inherit the persistence of the exogenous $\widehat{a}_t$. The last equality captures our definition of $\tilde{\kappa}_w$ in Proposition 3. Note that we abstract from behavioral unions. If unions were also behavioral, the wage Phillips curve would be given by

$$\pi_t^w = \kappa_w \widehat{x}_t + \beta \bar{m}_U \mathbb{E}_t \pi_{t+1}, \tag{39}$$

such that $\tilde{\kappa}_w = \frac{\kappa_w}{1-\beta\bar{m}_U\rho_a}$, where $\bar{m}_U$ captures the union's degree of cognitive discounting. Note that this is slightly different from the formulation in the price Phillips Curve in Gabaix (2020) as we consider Rotemberg adjustment costs instead of a Calvo friction.

The zero-profit condition is given by:

$$\Delta \widehat{a}_t = \pi_t^w - \pi_t, \tag{40}$$

which also holds in expectations (behavioral or not), such that

$$\begin{aligned}
\mathbb{E}_t^{BR}\left[\Delta \widehat{a}_{t+1}\right] &= \mathbb{E}_t^{BR}\left[\widehat{\pi}_{t+1}^w\right] - \mathbb{E}_t^{BR}\left[\widehat{\pi}_{t+1}\right] &(41)\\
\Leftrightarrow \mathbb{E}_t^{BR}\left[\widehat{\pi}_{t+1}\right] &= \mathbb{E}_t^{BR}\left[\widehat{\pi}_{t+1}^w\right] - \mathbb{E}_t^{BR}\left[\Delta \widehat{a}_{t+1}\right] &(42)\\
&= \rho_a \bar{m}_\pi \pi_t^w - \bar{m}_\pi(\rho_a - 1)\widehat{a}_t &(43)\\
&= \rho_a \bar{m}_\pi \tilde{\kappa}_w \widehat{x}_t - \bar{m}_\pi(\rho_a - 1)\widehat{a}_t. &(44)
\end{aligned}$$

where the penultimate line follows from the definition of the TFP process.

Furthermore, note that $\mathbb{E}_t^{BR}[\widehat{x}_{t+1}] = \bar{m}\rho_a\widehat{x}_t$. If we plug all these components in the IS equation above, we obtain

$$\begin{aligned}
\widehat{x}_t &= \delta\bar{m}\rho_a\widehat{x}_t - \psi_c\left[\phi\tilde{\kappa}_w\widehat{x}_t - \rho_a\bar{m}_\pi\tilde{\kappa}_w\widehat{x}_t + \bar{m}_\pi(\rho_a-1)\widehat{a}_t + (1-\bar{m}\rho_a)\widehat{a}_t\right] &(45)\\
\widehat{x}_t &= -\psi_c\frac{(1-\bar{m}\rho_a) - \bar{m}_\pi(1-\rho_a)}{1-\rho_a\psi_f + \psi_c\tilde{\kappa}_w(\phi - \rho_a\bar{m}_\pi)}\widehat{a}_t, &(46)
\end{aligned}$$

[34] This is true because we do not need to solve for actual inflation, but only expected inflation given our Taylor rule in terms of wage inflation. If we would instead write the Taylor rule in terms of price inflation, $\widehat{a}_{t-1}$ would be an additional state variable due the zero-profit condition.

which becomes

$$\widehat{x}_t = -\alpha \widehat{a}_t \tag{47}$$

where

$$\alpha \equiv \psi_c \frac{(1-\bar{m}_\pi) + \rho_a(\bar{m}_\pi - \bar{m})}{1 - \rho_a \psi_f + \psi_c \tilde{\kappa}\left(\phi - \rho_a \bar{m}_\pi\right)}. \tag{48}$$

If $\bar{m}_\pi = 1$, this completes Proposition 3.

Proof of Proposition 4. We have

$$\alpha \equiv \psi_c \frac{(1-\bar{m}_\pi) + \rho_a(\bar{m}_\pi - \bar{m})}{1 - \rho_a \psi_f + \psi_c \tilde{\kappa}\left(\phi - \rho_a \bar{m}_\pi\right)}. \tag{49}$$

Let $\mathcal{D} \equiv 1 - \rho_a \psi_f + \psi_c \tilde{\kappa}_w \big(\phi - \rho_a \, \bar{m}_\pi\big)$ denote the denominator in the expression for α. We obtain

$$\begin{aligned}
\frac{\partial \alpha}{\partial \chi} &= \frac{\left[(1-\bar{m}_\pi) + \rho_a(\bar{m}_\pi - \bar{m})\right] \frac{\partial \psi_c}{\partial \chi} \mathcal{D} - \left[(1-\bar{m}_\pi) + \rho_a(\bar{m}_\pi - \bar{m})\right] \psi_c \left(\frac{\partial \psi_c}{\partial \chi} \tilde{\kappa}_w \big(\phi - \rho_a \, \bar{m}_\pi\big) - \rho_a \bar{m} \frac{\partial \delta}{\partial \chi}\right)}{\mathcal{D}^2} \\
&= \frac{\left[(1-\bar{m}_\pi) + \rho_a(\bar{m}_\pi - \bar{m})\right]}{\mathcal{D}^2} \left[\frac{\partial \psi_c}{\partial \chi} \mathcal{D} + \rho_a \bar{m} \frac{\partial \delta}{\partial \chi} \psi_c - \psi_c \frac{\partial \psi_c}{\partial \chi} \tilde{\kappa}_w \big(\phi - \rho_a \, \bar{m}_\pi\big)\right] \\
&= \frac{\left[(1-\bar{m}_\pi) + \rho_a(\bar{m}_\pi - \bar{m})\right]}{\mathcal{D}^2} \left[\frac{\partial \psi_c}{\partial \chi} (1 - \rho_a \psi_f) + \rho_a \bar{m} \frac{\partial \delta}{\partial \chi} \psi_c\right],
\end{aligned}$$

which is exactly the expression in Proposition 4 for $\bar{m}_\pi = 1$. If $\bar{m} = \bar{m}_\pi = 1$, the term in front of the brackets is zero. Since ψ_c and δ are both increasing in χ and $\psi_f \rho_a < 1$[35], it follows that $\frac{\partial \alpha}{\partial \chi}$ is strictly positive for $\bar{m} < 1$. This holds independently of whether $\bar{m}_\pi$ is 1 or not, i.e., whether households hold rational real rate expectations or not.

For the derivative with respect to $\bar{m}$, we differentiate two cases: $\bar{m}_\pi = 1$ or $\bar{m}_\pi = \bar{m} \leq 1$. In the first case, we obtain

$$\begin{aligned}
\frac{\partial \alpha}{\partial \bar{m}} &= \frac{-\rho_a \psi_c \mathcal{D} + \rho_a \delta \rho_a (1-\bar{m}) \psi_c}{\mathcal{D}^2} && (50) \\
&= \frac{-\rho_a \psi_c \left[1 - \rho_a \psi_f + \psi_c \tilde{\kappa}_w \big(\phi - \rho_a\big) - \rho_a \delta + \rho_a \psi_f\right]}{\mathcal{D}^2} && (51) \\
&= \frac{-\rho_a \psi_c \left[1 + \psi_c \tilde{\kappa}_w \big(\phi - \rho_a\big) - \rho_a \delta\right]}{\mathcal{D}^2}, && (52)
\end{aligned}$$

which is strictly negative, as $\delta \rho_a < 1$. This concludes the proof of Proposition 4.

For completeness and to show that our results do not depend on the assumption of

[35]This follows in the behavioral TANK model because of $\psi_f < 1$ and the definition of mean-reverting shocks, $\rho_a < 1$, and is assumed in the rational TANK model by assumption to ensure non-explosiveness of this model as discussed in the main text.

rational expectations about today's real rate, we now also consider the case $\bar{m}_\pi = \bar{m} \leq 1$:

$$\frac{\partial \alpha}{\partial \bar{m}} = \frac{-\psi_c \mathcal{D} + (1-\bar{m})\psi_c(\rho_a \delta + \psi_c \tilde{\kappa}_w \rho_a)}{\mathcal{D}^2} \tag{53}$$

$$= \frac{-\psi_c(1-\rho_a\psi_f + \psi_c\tilde{\kappa}(\phi - \rho_a\,\bar{m}_\pi)) + (1-\bar{m})\psi_c(\rho_a\delta + \psi_c\tilde{\kappa}_w\rho_a)}{\mathcal{D}^2} \tag{54}$$

$$= \frac{-\psi_c\left[1 + \psi_c\tilde{\kappa}_w\big(\phi - \rho_a\big) - \rho_a\delta\right]}{\mathcal{D}^2}, \tag{55}$$

which is identical to the expression for the case $\bar{m}_\pi = 1$, except that $\bar{m}_\pi$ is now also in $\mathcal{D}$ and the numerator is not multiplied by ρ_a. Hence, the sign of the derivative is unaffected and our conclusions remain the same.

Proof of Lemma 1. When we assume $s = 1$, the sufficient statistic for the business cycle impact of TFP shocks simplifies to:

$$\alpha \equiv \frac{\rho_a(1-\bar{m})\psi_c}{1-\rho_a\bar{m} + \psi_c\tilde{\kappa}\left(\phi - \rho_a\right)}. \tag{56}$$

The derivative with respect to χ is then given by

$$\frac{\partial \alpha}{\partial \chi} = \rho_a(1-\bar{m})\,\frac{\partial \psi_c}{\partial \chi}\,\frac{1-\rho_a\bar{m}}{\left(1-\rho_a\bar{m} + \psi_c\tilde{\kappa}(\phi-\rho_a)\right)^2}. \tag{57}$$

Now, taking the cross-derivative with respect to $\bar{m}$ yields:

$$\frac{\partial^2 \alpha}{\partial \bar{m}\partial \chi} = -\rho_a\frac{\partial \psi_c}{\partial \chi}\frac{(\phi_w-\rho_a)\tilde{\kappa}_w\psi_c[1+\rho_a-2\bar{m}\rho_a] + (1-\rho_a)(1-\bar{m}\rho_a)}{(1-\psi_f\rho_a + \psi_c(\phi_w-\rho_a)\tilde{\kappa}_w)^3}, \tag{58}$$

which is negative. To see this, note that the denominator is always positive and the last term in the numerator is always positive. The derivative of ψ_c with respect to χ is also positive, and ρ_a is also positive. Furthermore, we have that $\phi_w > 1 > \rho_a$, $\tilde{\kappa}_w > 0$ and $\psi_c > 0$, such that it is sufficient to show that the term in brackets, given by: $1 + \rho_a - 2\bar{m}\rho_a$, is positive. This follows directly from $\bar{m} \leq 1$ and $\rho_a < 1$.

A.3 Persistent monetary policy shocks

In Section 3.2, we derived the following expression for output:

$$\widehat{y}_t = -\frac{\psi_c}{\gamma}\mathbb{E}_t\sum_{j=0}^{\infty}\psi_f^j r_{t+j}. \tag{59}$$

Instead of assuming that real rate shocks are i.i.d. we now consider the case in which they follow an AR(1) process with persistence ρ_r:

$$r_t = \rho_r r_{t-1} + \epsilon_t^r, \text{ with } \epsilon_t^r \sim \mathcal{N}(0, \sigma_r^2),$$

with $\rho_r \in [0, \psi_f^{-1})$. In that case, we can write equation (59) as follows:

$$\begin{aligned}\widehat{y}_t &= -\frac{\psi_c}{\gamma}\mathbb{E}_t \sum_{j=0}^{\infty}(\psi_f \rho_r)^j r_t \\ &= -\frac{\psi_c}{\gamma}\frac{1}{1-\psi_f\rho_r} r_t.\end{aligned}$$

It then directly follows that persistent monetary policy shocks have larger effects (in absolute terms) in the behavioral TANK model than in RANK whenever

$$\frac{\psi_c}{\gamma}\frac{1}{1-\psi_f\rho_r} > \frac{1}{\gamma}\frac{1}{1-\rho_r} \tag{60}$$

$$\Leftrightarrow \frac{\psi_c}{1-\psi_f\rho_r} > \frac{1}{1-\rho_r} \tag{61}$$

$$\Leftrightarrow \frac{1-\psi_c}{\psi_f-\psi_c} > \rho_r, \tag{62}$$

i.e., when the shocks are not *too persistent*. If households are more behavioral, reflected in a lower ψ_f, the cutoff value for ρ_r decreases, as explained in Section 3.2. It is also immediate from the last expression that the rational TANK model with $\psi_c > 1$ always amplifies monetary policy shocks relative to RANK, as $\rho_r < 1$.

A.4 Determinacy and Effective Lower Bound

In this section, we revisit the determinacy conditions in the behavioral TANK model and discuss the implications for the stability at the effective lower bound constraint on nominal interest rates. We therefore focus on the case where monetary policy follows a standard Taylor rule described in equation (24), and we abstract from TFP shocks. With constant TFP, the zero-profit condition (12) implies $\pi_t^w = \pi_t$ and, hence, it is irrelevant whether monetary policy targets price or wage inflation. In addition, it follows that the wage Phillips Curve (equation (13)) fully describes price inflation. We can, thus, express the real interest rate in the IS equation as $r_t = i_t - \mathbb{E}_t\pi_{t+1}^w = \phi_w\pi_t^w - \mathbb{E}_t\pi_{t+1}^w$ and summarize the model by two endogenous variables $\widehat{x}_t$ and π_t^w, whose dynamics are described by the following two equations:

$$\begin{pmatrix}\mathbb{E}_t\pi_{t+1}^w \\ \mathbb{E}_t\widehat{x}_{t+1}\end{pmatrix} = \begin{pmatrix}\frac{1}{\beta} & -\frac{\kappa_w}{\beta} \\ \frac{\psi_c}{\psi_f}\left(\phi_w - \frac{1}{\beta}\right) & \frac{1}{\psi_f}\left(1+\frac{\psi_c\kappa_w}{\beta}\right)\end{pmatrix}\begin{pmatrix}\pi_t^w \\ \widehat{x}_t\end{pmatrix}. \tag{63}$$

For determinacy, we need

$$det(A) > 1; \quad det(A) - tr(A) > -1; \quad det(A) + tr(A) > -1. \tag{64}$$

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

$$\frac{1}{\beta\psi_f}\left(1 + \frac{\psi_c\kappa_w}{\beta}\right) + \frac{\kappa_w\psi_c}{\beta\psi_f}\left(\phi_w - \frac{1}{\beta}\right) > 1 \tag{65}$$

and

$$\frac{1}{\beta\psi_f}\left(1 + \frac{\psi_c\kappa_w}{\beta}\right) + \frac{\kappa_w\psi_c}{\beta\psi_f}\left(\phi_w - \frac{1}{\beta}\right) - \frac{1}{\beta} - \frac{1}{\psi_f}\left(1 + \frac{\psi_c\kappa_w}{\beta}\right) > -1. \tag{66}$$

From condition (65), it follows that:

$$\phi_w > \frac{\beta\psi_f - 1}{\kappa_w\psi_c}. \tag{67}$$

Note that in the behavioral TANK model $\psi_f < 1$, and hence, the right hand side is negative. Therefore, this condition is usually not binding. Condition (66) implies:

$$\phi_w > 1 + \frac{(\psi_f - 1)(1 - \beta)}{\kappa_w\psi_c}. \tag{68}$$

Thus, if $\psi_f = 1$, we would recover the standard Taylor principle saying that $\phi_w > 1$ is required for determinacy. Yet, given that $\psi_f < 1$ in the behavioral TANK model, the standard Taylor principle is relaxed. Hence, behavioral expectations widen the determinacy region, similar as in Gabaix (2020). Under some calibrations, the model is determinate even under an interest rate peg, i.e., even if $\phi_w = 0$. Equation (68) also shows that rational TANK models narrow the determinacy region as in this case $\psi_f > 1$ (see also Acharya and Dogra, 2020; Ravn and Sterk, 2021; Bilbiie, 2025).

Stability at the effective lower bound. Related to the indeterminacy issues that may arise under rational expectations—and that become even more pronounced in TANK models with type switching—the standard 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. Similar to the forward guidance puzzle, this is even more severe in rational HANK models.

We now show that the behavioral TANK model resolves these issues. To this end, we isolate the recessionary impact of a binding effective lower bound in the following way: First, we introduce a natural rate shock, $r_t^{n,\epsilon}$. However, in contrast to our analysis in Section 3, we abstract from how fundamental shocks map into natural rate fluctuations and instead compare the same $r_t^{n,\epsilon}$ shock across models. We further assume that the natural rate movements do not arise from the supply side, i.e., that TFP stays constant, such that we can replace $\mathbb{E}_t\pi_{t+1}$ with $\mathbb{E}_t\pi^w_{t+1}$. Hence, the IS equation is given by:

$$x_t = \psi_f \mathbb{E}_t x_{t+1} - \psi_c \left(i_t - \mathbb{E}_t \pi^w_{t+1} - r_t^{n,\epsilon} \right).$$

Second, for simplicity and because it does not qualitatively affect our result, we consider a static Phillips curve as in Bilbiie (2025), hence, $\mathbb{E}_t\pi^w_{t+1} = \kappa_w \mathbb{E}_t x_{t+1}$. And third, we assume that, absent the effective lower bound, monetary policy tracks the natural rate perfectly, that is, $i_t = max(r_t^{n,\epsilon}, -\bar{i})$. Thus, the output gap is only affected by the shock because of the binding lower bound.

We then consider the following experiment following Gabaix (2020) and Eggertsson and Woodford (2003). In period t the natural rate decreases to a value $r_t^{n,\epsilon} = \widetilde{r}^n < -\bar{i}$ 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 the output gap in period t is given by

$$x_t = -\psi_c \underbrace{\left(-\bar{i} - \widetilde{r}^n\right)}_{>0} \sum_{j=0}^{k} \left(\psi_f + \kappa_w \psi_c\right)^j, \tag{69}$$

where the term $(-\bar{i} - \widetilde{r}^n) > 0$ captures the shortfall of the policy response due to the binding ELB. Under rational expectations, we have that $\psi_f > 1$ (and $\kappa_w\psi_c > 0$), meaning that the output gap 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 TANK model, however, this is not the case. As long as $\psi_f + \kappa_w\psi_c < 1$ the output response in t is bounded even as $k \to \infty$. We find that $\bar{m} < 0.94$ is enough to rule out unboundedly-severe recessions at the ELB even if the ELB is expected to persist forever.

Equation (69) also shows that, in contrast to TFP or government spending shocks, unequal exposure and cognitive discounting have opposite effects on the recessionary impact of a binding ELB. While more unequal exposure increases ψ_c and ψ_f, it increases the recessionary impact of a binding lower bound, whereas more cognitive discounting lowers ψ_f and, thus, dampens the recessionary impact of a binding lower bound, similarly as in Gabaix (2020). Because a sustained binding lower bound triggers anticipation of a too contractionary monetary policy in the future, the intuition is similar to the one discussed for persistent monetary policy shocks in Section 3.

A.5 Indirect effects

To decompose the total effects into direct and indirect effects from real rate changes, this Section now derives the aggregate consumption function. 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

$$\frac{1}{\gamma}\mathbb{E}_t^{BR}\widehat{q}_{t,t+i}^U + \bar{m}^i\mathbb{E}_t\widehat{c}_{t+i}^U = \widehat{c}_t^U + (1-s)\mathbb{E}_t\sum_{k=1}^{i}\bar{m}^k\left(\widehat{c}_{t+k}^U - \widehat{c}_{t+k}^H\right). \tag{70}$$

The (linearized) budget constraint can be written as

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

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 (70) when summing them up. The left-hand side of the budget constraint can thus be written as

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

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

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

and similarly,

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

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

Combining these results, we see that

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

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

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

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 aggregate consumption function

$$\widehat{c}_t = [1 - \beta(1-\lambda\chi)]\,\widehat{y}_t - \frac{(1-\lambda)\beta}{\gamma}\widehat{r}_t + \beta\bar{m}\delta(1-\lambda\chi)\mathbb{E}_t\widehat{c}_{t+1}. \tag{71}$$

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 aggregate consumption function (71), we get

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

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

Given our calibration and assuming an AR(1) monetary policy shock with a persistence of 0.6, indirect effects account for about 62%, consistent with larger quantitative models as for example in Kaplan et al. (2018).[36] Holm et al. (2021) state that the overall importance of indirect effects they find in the data is comparable to those in Kaplan et al. (2018), with the difference that these effects unfold after some time, whereas direct effects are more important on impact. Because in our stylized model the response to a monetary policy shock peaks on impact, indirect effects are important right away. Slacalek et al. (2020) provide further evidence that indirect effects are strong drivers of aggregate consumption in response to monetary policy shocks. For comparison, the representative agent model generates an indirect share of $\frac{1-\beta}{1-\beta\bar{m}\rho}$, which, given our calibration, amounts to about 2% whether or not households are behavioral.

A.6 Price inflation targeting and Keynesian supply shocks

In Section 3, we assume a Taylor rule that responds to wage inflation and not price inflation. Aside from making the model more tractable, we also make this assumption to rule out *Keynesian supply shocks*, i.e., monetary policy responses that depress output more than potential output by increasing real rates more than the natural rate in response to negative TFP shocks. We now show that this issue arises in our model with sticky wages and flexible prices when we assume a Taylor rule that responds to price inflation.

Recall that the natural rate is given by $r_t^n = -(1-\bar{m}\rho_a)\widehat{a}_t$ (see Proposition 1). The actual real rate under the Taylor rule $i_t = \phi\pi_t$ is given by:

$$\begin{aligned} & r_t = i_t - \mathbb{E}_t\pi_{t+1} \\ \Leftrightarrow \quad & r_t = \phi\pi_t - \mathbb{E}_t\pi_{t+1}. \end{aligned}$$

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

Inserting the zero-profit condition yields:

$$r_t = \phi(\pi_t^w - \Delta\widehat{a}_t) - \mathbb{E}_t(\pi_{t+1}^w - \Delta\widehat{a}_{t+1}).$$

For simplicity, we assume a static Phillips Curve: $\pi_t^w = \kappa_w \widehat{x}_t$. The only state variables are $\widehat{a}_t$ and $\widehat{a}_{t-1}$ and we therefore guess that all variables can be expressed as linear functions of those two state variables. In particular, we guess:

$$\widehat{x}_t = \alpha_1 \widehat{a}_t + \alpha_2 \widehat{a}_{t-1}, \tag{72}$$

where α_1 and α_2 are two yet unknown parameters. Hence, we can express the real rate as:

$$r_t = \phi\left(\kappa_w \alpha_1 \widehat{a}_t + \kappa_w \alpha_2 \widehat{a}_{t-1} - \widehat{a}_t + \rho_a \widehat{a}_{t-1}\right) - \kappa_w \left[\rho\alpha_1 + \alpha_2\right] \widehat{a}_t + (\rho_a - 1)\widehat{a}_t. \tag{73}$$

If we plug this expression into the IS equation, we obtain:

$$\alpha_1 \widehat{a}_t + \alpha_2 \widehat{a}_{t-1} = \psi_f \left[\alpha_1 \rho_a \widehat{a}_t + \alpha_2 a_t\right] - \psi_c \Big\{ \phi\Big(\kappa_w \alpha_1 \widehat{a}_t + \kappa_w \alpha_2 \widehat{a}_{t-1} - \widehat{a}_t + \rho_a \widehat{a}_{t-1}\Big) \tag{74}$$

$$- \kappa_w \left[\rho\alpha_1 + \alpha_2\right] \widehat{a}_t + (\rho_a - 1)\widehat{a}_t + (1 - \bar{m}\rho_a)\widehat{a}_t \Big\}. \tag{75}$$

Grouping all the $\widehat{a}_{t-1}$ terms yields

$$\alpha_2 = -\frac{\rho_a \psi_c \phi}{1 + \psi_c \kappa_w \phi}. \tag{76}$$

Grouping the $\widehat{a}_t$ terms yields

$$\begin{aligned}
\alpha_1 &= \psi_f[\alpha_1 \rho_a + \alpha_2] - \psi_c\left[\phi(\kappa_w \alpha_1 - 1) - \kappa_w(\rho_a \alpha_1 + \alpha_2) + \rho_a(1 - \bar{m})\right] \\
\alpha_1\left[1 - \psi_f \rho_a + \psi_c \kappa_w(\phi - \rho_a)\right] &= \psi_f \alpha_2 + \psi_c \phi + \psi_c \kappa_w \alpha_2 - \psi_c \rho_a (1 - \bar{m}) \\
\alpha_1 &= \frac{\psi_f \alpha_2 + \psi_c \phi + \psi_c \kappa_w \alpha_2 - \psi_c \rho_a (1 - \bar{m})}{1 - \psi_f \rho_a + \psi_c \kappa_w(\phi - \rho_a)} \\
&= \frac{-\psi_f \frac{\rho_a \psi_c \phi}{1 + \psi_c \kappa_w \phi} + \psi_c \phi - \psi_c \kappa_w \frac{\rho_a \psi_c \phi}{1 + \psi_c \kappa_w \phi} - \psi_c \rho_a (1 - \bar{m})}{1 - \psi_f \rho_a + \psi_c \kappa_w(\phi - \rho_a)}.
\end{aligned}$$

For Keynesian supply shocks to materialize, it is sufficient to show that $\alpha_1 > 0$, in which case a negative TFP shock leads to a decrease in the output gap, i.e., when a negative TFP shock hits the economy, actual output decreases more than potential output because the real rate increases more than the natural rate.

The denominator in the expression for α_1 is positive, hence, we can focus on the numerator:

$$0 < -\psi_f \frac{\rho_a \psi_c \phi}{1 + \psi_c \kappa_w \phi} + \psi_c \phi - \psi_c \kappa_w \frac{\rho_a \psi_c \phi}{1 + \psi_c \kappa_w \phi} - \psi_c \rho_a (1 - \bar{m}) \tag{77}$$

$$= -\frac{\psi_f \rho_a \phi}{1 + \psi_f \kappa_w \phi} + \phi - \frac{\kappa_w \rho_a \psi_c \phi}{1 + \psi_f \kappa_w \phi} - \rho_a (1 - \bar{m}) \tag{78}$$

$$\Leftrightarrow \quad \psi_f \rho_a \phi + \kappa_w \rho_a \psi_c \phi + \rho_a (1 - \bar{m})(1 + \psi_c \kappa_w \phi) < \phi(1 + \psi_c \kappa_w \phi). \tag{79}$$

Since we want to rule out Keynesian supply shocks for any $\bar{m} \in [0, 1]$, we focus for simplicity here on the rational-expectations model, $\bar{m} = 1$. In this case, it is straightforward to see that Keynesian supply shocks arise:

$$\psi_f \rho_a \phi + \kappa_w \rho_a \psi_c \phi < \phi(1 + \psi_c \kappa_w \phi) \tag{80}$$

$$\psi_f \rho_a + \kappa_w \rho_a \psi_c < 1 + \psi_c \kappa_w \phi \tag{81}$$

$$0 < \underbrace{1 - \psi_f \rho_a}_{>0} + \underbrace{\psi_c \kappa_w (\phi - \rho_a)}_{>0}. \tag{82}$$

To rule out these Keynesian supply shocks for all different model versions we consider, we focus on a Taylor rule that targets wage inflation, not price inflation, in Section 3.

A.7 Equivalence with sticky prices and flexible wages

We focus here on the linearized economy. While we keep the optimal subsidy to absorb all profits in steady state, with sticky prices and flexible wages, profits are not necessarily equal to zero outside steady state. Thus, we need to define the profit shares, $d(e_{i,t}$, in households budget constraint (2). We set $d(e_H) = \frac{\mu_D}{\lambda}$, with μ_D being the total share of profits that goes to all H households and λ the share of H households. The budget constraint of the H households is given by:

$$\widehat{c}_t^H = \widehat{n}_t^H + \frac{\mu_D}{\lambda} d_t + \widehat{w}_t. \tag{83}$$

Aggregate profits are equal to:

$$d_t = \widehat{a}_t - \widehat{w}_t. \tag{84}$$

The linearized labor-leisure equations are given by:

$$\varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H = w_t \tag{85}$$

$$\varphi \widehat{n}_t^U + \gamma \widehat{c}_t^U = w_t \tag{86}$$

$$\Leftrightarrow \varphi \widehat{n}_t^H + \gamma \widehat{c}_t^H = \varphi \widehat{n}_t^U + \gamma \widehat{c}_t^U. \tag{87}$$

Market clearing in the labor and the goods market yield:

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

Replacing $\widehat{n}_t^U$ and $\widehat{c}_t^U$ in the equalized labor leisure equations, and plugging the expression for the wage in the H's budget constraint, yields:

$$\widehat{c}_t^H = \chi \widehat{y}_t - \zeta(\chi - 1)\widehat{a}_t, \tag{88}$$

with the exposure of H households being a function of the inverse Frisch elasticity, φ, the share of H households, λ, and their share in profits, μ_D, $\chi = 1 + \varphi(1 - \frac{\mu_D}{\lambda})$, and $\zeta = \frac{1+\varphi}{\gamma+\varphi}$. Note that potential output is given by $\widehat{y}_t^{pot} = \zeta \widehat{a}_t$. Hence, we can also express consumption of the H households as:

$$\widehat{c}_t^H = \chi \widehat{x}_t + \zeta \widehat{a}_t. \tag{89}$$

Using this expression in the market clearing condition for consumption yields consumption of the U households:

$$\widehat{c}_t^U = \frac{1}{\psi_c}\widehat{x}_t + \zeta \widehat{a}_t, \tag{90}$$

where $\psi_c = \frac{1-\lambda}{1-\lambda\chi}$, which is exactly the same as in our baseline model.

Using these expressions in the IS equation, setting $\gamma = 1$, and using the natural rate $r_t^n = -(1 - \bar{m}\rho_a)\widehat{a}_t$,[37] we obtain:

$$\widehat{x}_t = \psi_f \mathbb{E}_t \widehat{x}_{t+1} - \psi_c \left(r_t - r_t^n\right), \tag{91}$$

where the real rate is given by $r_t = i_t - \mathbb{E}_t \pi_{t+1}$.

Hence, we fully recover the same IS equation as in our sticky wage, flexible price baseline model in Section 3. This implies that all our results regarding the transmission of monetary policy in Section 3.2 readily carry over to the sticky price model. The difference is that now, the redistribution towards high-MPC households after monetary policy shocks works via countercyclical profits which are mainly paid to unconstrained households exactly as in Bilbiie (2025) instead of via the unequal exposure of households' labor earnings. In fact, this sticky price model now fully nests the model in Bilbiie (2025) if we set $\bar{m} = 1$ and abstract from fluctuations in the natural rate, implying that it is now also equivalent in terms of wage

[37] Note that the natural rate is independent of the nominal rigidity as it is defined as the real rate in the flexible-wage, flexible-price economy.

and profit dynamics.

The supply side of this sticky price model, however, cannot be equivalent to the sticky wage model in the presence of TFP shocks (see also Bilbiie and Trabandt (2025)). We nevertheless show under what conditions our sticky price model behaves equivalently in response to TFP shocks, in the sense, that the output gap responds identically as in the sticky wage model as well as that price inflation replicates the path of wage inflation in our baseline model. To this end, we assume that the monetary authority sets the nominal interest rate according to the following rule:

$$i_t = \phi_\pi \pi_t + \phi_a \widehat{a}_t, \tag{92}$$

allowing monetary policy to directly respond to TFP shocks. The price Phillips curve is standard:

$$\pi_t = \kappa \widehat{x}_t + \beta \mathbb{E}_t \pi_{t+1}. \tag{93}$$

The solution to the model is given by:

$$\widehat{x}_t = -\alpha \widehat{a}_t, \tag{94}$$

where

$$\alpha = \frac{\psi_c (1 - \bar{m}\rho_a + \phi_a)}{1 - \rho_a \psi_f + (\phi_\pi - \rho_a)\tilde{\kappa}}, \tag{95}$$

with $\tilde{\kappa} = \frac{\kappa}{1-\beta\rho_a}$. Hence, if $\kappa = \kappa_w$ and $\phi_\pi = \phi_w$, we obtain equivalence between the two models if and only if:

$$\phi_a = -(1 - \rho_a). \tag{96}$$

Hence, if monetary policy "mimics" the direct effect of TFP on the real rate in the sticky wage model through flexible prices (see (28)), the two models coincide. Since in this case, the solution is exactly the same, our results in Section 3 regarding the effects of changes in χ or $\bar{m}$, as well as their interaction, remain completely unchanged.

If we instead focus on the case with $\phi_a = 0$, first note that in this case, we recover the typical textbook New Keynesian result that TFP shocks affect the output gap even with rational expectations: for any given Taylor rule, monetary policy does not reduce consumption sufficiently to realign output with its potential. Yet, the qualitative nature of the impact of bounded rationality, unequal exposure, and their interaction remain the same. Specifically, we obtain the following results:

$$\frac{\partial \alpha}{\partial \chi} = \frac{\frac{\partial \psi_c}{\partial \chi}(1 - \bar{m}\rho_a)\left(1 - \rho_a \psi_f + (\phi_\pi - \rho_a)\tilde{\kappa}\right) + \rho_a \frac{\partial \psi_f}{\partial \chi}\psi_c (1 - \bar{m}\rho_a)}{\left(1 - \rho_a \psi_f + (\phi_\pi - \rho_a)\tilde{\kappa}\right)^2} > 0 \tag{97}$$

$$\frac{\partial \alpha}{\partial \bar{m}} = -\frac{\psi_c \rho_a \left(1 - \rho_a \psi_f + (\phi_\pi - \rho_a)\tilde{\kappa}\right) - \rho_a \delta \psi_c (1 - \bar{m}\rho_a)}{\left(1 - \rho_a \psi_f + (\phi_\pi - \rho_a)\tilde{\kappa}\right)^2} \tag{98}$$

$$= -\frac{\psi_c \rho_a (1 - \delta) + \psi_c \rho_a (\phi_\pi - \rho_a)\tilde{\kappa}}{\left(1 - \rho_a \psi_f + (\phi_\pi - \rho_a)\tilde{\kappa}\right)^2}. \tag{99}$$

$\frac{\partial \alpha}{\partial \bar{m}}$ is negative if

$$1 + (\phi_\pi - \rho_a)\tilde{\kappa} > \delta,$$

which holds as long as the type-switching probability $1 - s$ is not too high. For our baseline calibration, the cutoff value is $s < 0.68$ which would imply an almost 80% annual chance of becoming hand-to-mouth which is far more extreme than typical calibrations for s.

Let us now consider $s = 1$, in which case, ψ_c is a sufficient statistic for heterogeneity. We then get:

$$\frac{\partial^2 \alpha}{\partial \psi_c \partial \bar{m}} = -\frac{(\phi_\pi - \rho_a)\tilde{\kappa}\rho_a}{\left(1 - \rho_a \psi_f + (\phi_\pi - \rho_a)\tilde{\kappa}\right)^2} < 0. \tag{100}$$

Hence, our results in the sticky wage model regarding how $\bar{m}$, χ and the interaction of the two shape the amplification of TFP shocks extend to the sticky price model.

A.8 Details on the equivalence result between higher government spending and lower TFP

In this section, we consider government spending shocks G_t that follow an AR(1)-process and assume for now constant TFP $A_t = \bar{A}$. We assume that government spending is financed by raising labor-income taxes τ_t^L and abstract from progressivity, that is $\tau_P = 1$, such that the household's budget constraint in equilibrium reads:

$$C_{i,t} = (1 - \tau_t^L)(W_t z(e_{i,t}) N_{i,t}), \tag{101}$$

as bond holdings are 0 in equilibrium. The government budget constraint is balanced in all periods t:

$$G_t = \tau_t^L W_t N_t. \tag{102}$$

Given constant TFP, the zero-profit condition implies that the real wage W_t is constant.

Linearization around $G = 0$ steady state. We start by linearizing the government budget constraint around the $G = \tau^L = 0$ steady state, and we use lower-case letters to express absolute deviations of G and τ^L from this steady state. Hence, the government budget constraint in linearized form is given by

$$g_t = \tau_t^L. \tag{103}$$

The production function and the resource constraint yield

$$\widehat{n}_t = \widehat{y}_t = \widehat{c}_t + g_t. \tag{104}$$

In the economy with flexible wages and prices, the aggregate labor-leisure condition coming from the labor unions is given by

$$N_t^{\varphi} = (1 - \tau_t^L)(W_t)C_t^{-1}, \tag{105}$$

which, in linearized terms, reads as:

$$\varphi \widehat{n}_t = -\tau_t^l - \widehat{c}_t \tag{106}$$

$$\Leftrightarrow \varphi \widehat{n}_t = -g_t - (\widehat{n}_t - g_t) \tag{107}$$

$$\Leftrightarrow (\varphi + 1)\widehat{n}_t = -g_t(1 - 1) \tag{108}$$

$$\Leftrightarrow \widehat{n}_t = 0. \tag{109}$$

Hence, hours in the flexible-wage economy do not respond to government spending shocks. Hence, potential output does not respond either, and potential consumption is given by $\widehat{c}_t = -g_t$. Note that hours and consumption in the flexible-wage economy behaves exactly the same as after TFP shocks when $\widehat{a}_t = -g_t$.

Now, turning to the economy with wage stickiness, we start by linearizing the H household's budget constraint using our allocation rule $\widehat{n}_t^H = \chi \widehat{n}_t$:

$$\widehat{c}_t^H = -\tau_t^l + \chi \widehat{n}_t. \tag{110}$$

Labor market clearing implies

$$\widehat{n}_t^U = \frac{1 - \lambda\chi}{1 - \lambda} \widehat{n}_t, \tag{111}$$

so that the U household's budget constraint is given by

$$\widehat{c}_t^U = -\tau_t^l + \frac{1 - \lambda\chi}{1 - \lambda} \widehat{n}_t. \tag{112}$$

Plugging these expressions for the two type's consumption into the Euler equation and replacing $\widehat{n}_t$ with $\widehat{y}_t$, we obtain

$$-\tau_t^l + \frac{1 - \lambda\chi}{1 - \lambda} \widehat{y}_t = s\bar{m}\mathbb{E}_t \left[-\tau_{t+1}^l + \frac{1 - \lambda\chi}{1 - \lambda} \widehat{y}_{t+1} \right] + (1 - s)\bar{m}\mathbb{E}_t \left[-\tau_{t+1}^l + \chi \widehat{y}_{t+1} \right] - r_t \tag{113}$$

$$\frac{1}{\psi_c} \widehat{y}_t = s\bar{m} \frac{1}{\psi_c} \mathbb{E}_t \widehat{y}_{t+1} + (1 - s)\bar{m}\chi \mathbb{E}_t \widehat{y}_{t+1} + \tau_t^l (1 - \bar{m}\rho_g) - r_t \tag{114}$$

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \left[r_t - \tau_t^l (1 - \bar{m}\rho_g) \right] \tag{115}$$

$$\widehat{y}_t = \psi_f \mathbb{E}_t \widehat{y}_{t+1} - \psi_c \left[r_t - g_t(1 - \bar{m}\rho_g) \right] \tag{116}$$

$$\widehat{x}_t = \psi_f \mathbb{E}_t \widehat{x}_{t+1} - \psi_c \left[r_t - g_t(1 - \bar{m}\rho_g) \right], \tag{117}$$

with the last equality coming from the fact that potential output does not move after a government spending shock.

Now, the wage Phillips Curve is given by

$$\frac{\kappa_w}{\varphi} \left(\varphi \widehat{n}_t + \widehat{c}_t + g_t \right) = \pi_t^w - \beta \mathbb{E}_t \pi_{t+1}^w \tag{118}$$

$$\frac{\kappa_w}{\varphi} \varphi \widehat{x}_t = \pi_t^w - \beta \mathbb{E}_t \pi_{t+1}^w \tag{119}$$

$$\pi_t^w = \frac{\kappa_w}{\varphi} \varphi \widehat{x}_t + \beta \mathbb{E}_t \pi_{t+1}^w \tag{120}$$

$$= \tilde{\kappa}_w \widehat{x}_t, \tag{121}$$

with $\tilde{\kappa}_w \equiv \frac{\kappa_w}{1-\beta\rho_g}$.

We allow monetary policy to potentially directly react to the government spending shock, such that the Taylor rule is given by

$$i_t = \phi_w \pi_t^w + \phi_g g_t = \phi_w \tilde{\kappa}_w \widehat{x}_t + \phi_g g_t. \tag{122}$$

With constant TFP, $\pi_t = \pi_t^w$ and, thus, expected inflation is given by

$$\mathbb{E}_t \pi_{t+1} = \tilde{\kappa}_w \mathbb{E}_t \widehat{x}_{t+1} \tag{123}$$

so that the real rate is given by

$$r_t = \tilde{\kappa}_w (\phi_w - \rho_g) \widehat{x}_t + \phi_g g_t. \tag{124}$$

Plugging this into the IS equation yields:

$$\widehat{x}_t = \psi_f \mathbb{E}_t \widehat{x}_{t+1} - \psi_c \left[\tilde{\kappa}_w (\phi_w - \rho_g) \widehat{x}_t + \phi_g g_t - g_t(1 - \bar{m}\rho_g) \right]. \tag{125}$$

Hence, we obtain the solution:

$$\widehat{x}_t = \alpha_g g_t \tag{126}$$

with

$$\alpha_g \equiv \frac{(1 - \bar{m}\rho_g - \phi_g)\psi_c}{1 - \psi_f \rho_g + \psi_c \tilde{\kappa}_w (\phi_w - \rho_g)}. \tag{127}$$

It follows that for $\phi_g = 1 - \rho_g$, α_g is identical to the solution for TFP shocks α (see Proposition 3), when $\widehat{a}_t = -g_t$ and $\rho_a = \rho_g$, which proves Lemma 2.

If $\phi_g = 0$, the output gap, consumption, and price inflation behave after a positive government spending shock exactly like the sticky price model after negative TFP shocks which we derived in Appendix A.7. This implies that the derivatives with respect to $\bar{m}$, χ, and the cross derivative are isomorphic to the ones in equations (97)-(100). This also shows that in the sticky price model, a negative TFP and a positive government spending shock move the output gap, consumption, and price inflation equivalently.

A.9 Idiosyncratic risk $s < 1$

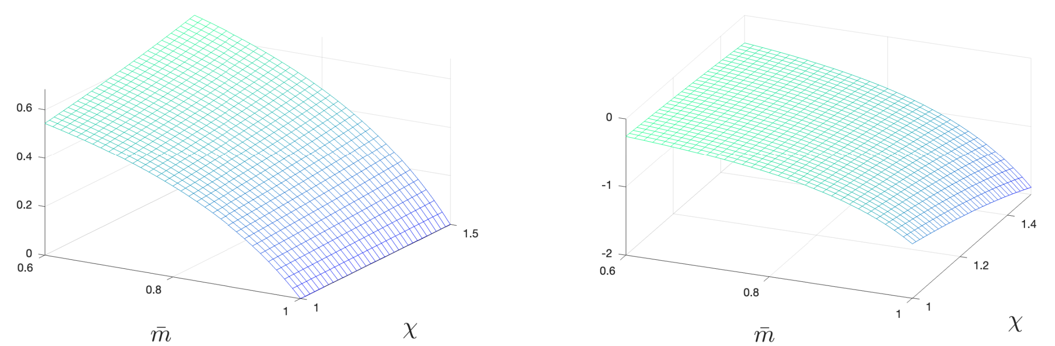
Lemma 1 in Section 3 analytically shows that unequal exposure and cognitive discounting are complements in amplifying the business cycle implications of TFP shocks (and by extension, government spending shocks). Lemma 1 focuses on the limit case of $s = 1$, i.e., the case without idiosyncratic risk, because this allows us to derive the sign of the cross derivative unambiguously. While this is not possible anymore with $s < 1$, we verify numerically that the results in Lemma 1 extend to the case with idiosyncratic risk. We rely on our baseline calibration outlined above (with $s = 0.946$) and set the Taylor rule coefficient to $\phi_w = 2$.

Panel (a) of Figure 3 plots our sufficient statistic for the business cycle impact of a given TFP shock, α, (derived in Proposition 3) for different values of χ and $\bar{m}$. Panel (b) then shows its cross derivative, $\frac{\partial^2 \alpha}{\partial \chi \partial \bar{m}}$ for varying degrees of χ and $\bar{m}$. Just as in the case of $s = 1$ (Lemma 1), the cross derivative is always negative. Hence, also with $s < 1$, unequal exposure and cognitive discounting are complements, as they mutually reinforce their impact on the business cycle impact of a given TFP shock.

While the cross derivative is robustly negative for our baseline calibration of $s = 0.946$, theoretically the sign can flip for a low enough s. Thus, we numerically look for the cutoff $\bar{s}$. We find $\bar{s} \approx 0.82$, which corresponds to a risk of 55% at annual frequency that an unconstrained household becomes hand-to-mouth which is far below (i.e., more extreme than) typical calibrations for s. We conclude that the negative cross-derivative is robust. In Section 4, we rely on an empirically-disciplined idiosyncratic risk process for the quantification of the interaction in our quantitative behavioral HANK model.

Figure 3: Amplification in the behavioral TANK model with $s < 1$

(a) Output response α (b) Interaction $\frac{\partial^2 \alpha}{\partial \chi \partial \bar{m}}$



Notes: This figure shows the output gap response to a TFP shock, α, (panel (a)), how it depends on unequal exposure χ and cognitive discounting $\bar{m}$ (panels (b) and (c)), and the interaction of the two (panel (d)), for the case $s = 0.946$.

A.10 EIS $1/\gamma = 1/2$

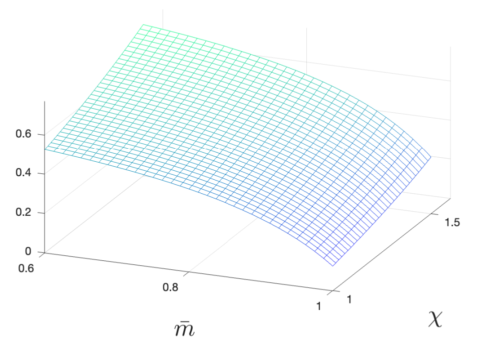
For tractability, we focus on an elasticity of intertemporal substitution of $1/\gamma = 1$ in Section 3. Yet, this assumption is not completely without loss of generality, as the exact irrelevance result of TFP shocks in rational-expectation models hinges on it. Nevertheless, our results remain qualitatively unchanged when considering $1/\gamma = 1/2$ as we numerically verify in the following.

Figure 4 plots the sufficient statistic for the business cycle impact, α (panel (a), its derivative with respect to χ (panel (b), its derivative with respect to $\bar{m}$ (panel (c)), and their cross derivative (panel (d)) each for varying χ and $\bar{m}$ values.[38] All four panels show that the signs are exactly the same as in the case with unit EIS described in Proposition 4 and in Lemma 1. Note that we also do not rely on an EIS of 1 in the quantitative behavioral HANK model in Section 4. We thus conclude that the qualitative insights derived in Section 3 do not hinge on the EIS of 1 assumption.

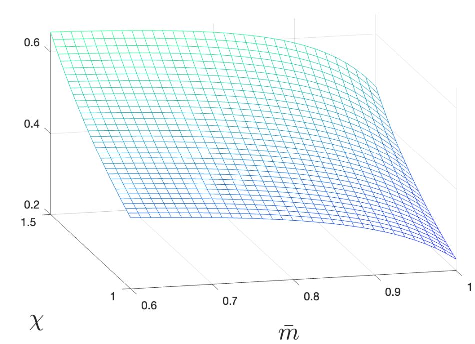
[38] We set the inverse Frisch elasticity to $\varphi = 2$ and idiosyncratic risk $s = 0.946$.

Figure 4: Amplification in the behavioral TANK model with $\frac{1}{\gamma} = \frac{1}{2}$

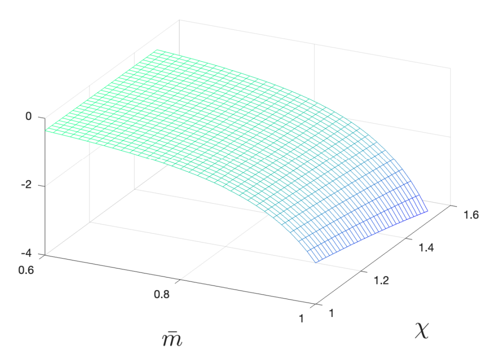
(a) Output response α



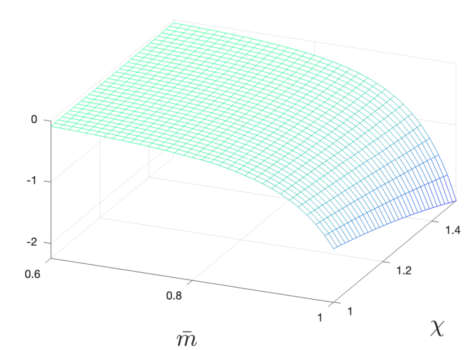
(b) Role of unequal exposure $\frac{\partial \alpha}{\partial \chi}$



(c) Role of cognitive discounting $\frac{\partial \alpha}{\partial \bar{m}}$



(d) Interaction $\frac{\partial^2 \alpha}{\partial \chi \partial \bar{m}}$



Notes: This figure shows the output gap response to a TFP shock, α, (panel (a)), how it depends on unequal exposure χ and cognitive discounting $\bar{m}$ (panels (b) and (c)), and the interaction of the two (panel (d)), for the case $1/\gamma = 1/2$.

A.11 Microfounding $\bar{m}$

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

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

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

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

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

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

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

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

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

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

Furthermore, we have

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

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

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

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

$$X_{t+1} = \bar{m}\left(\Gamma X_t + \varepsilon_{t+1}\right), \tag{128}$$

as assumed. 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 Estimating cognitive discounting

In this section, we discuss a new approach how we can estimate households' degree of cognitive discounting, $\bar{m}$, using survey data. Our approach is complementary to approaches that estimate cognitive discounting indirectly by estimating DSGE models with cognitive discounting (Afsar et al., 2024; Benchimol et al., 2025), as well as the experimental approach in Roth et al. (2021). The empirical estimates in those existing papers align with our findings.

Estimation approach. Let $\mathcal{I}_t$ denote the information set available to agents at time t. The full-information rational expectation of a variable x_{t+1} is

$$\mathbb{E}_t x_{t+1} \equiv \mathbb{E}[x_{t+1} \mid \mathcal{I}_t],$$

and the corresponding behavioral expectation under *cognitive discounting* is given by

$$\mathbb{E}^{BR}_t x_{t+1} = \bar{m}\,\mathbb{E}_t x_{t+1}, \tag{129}$$

where $\bar{m}$ is the cognitive discounting parameter which we want to estimate empirically.

Under full-information rational expectations, the forecast error for x_{t+1} is

$$\eta_{t+1} \equiv x_{t+1} - \mathbb{E}_t x_{t+1}, \tag{130}$$

so that

$$x_{t+1} = \mathbb{E}_t x_{t+1} + \eta_{t+1}, \tag{131}$$

with

$$\mathbb{E}_t[\eta_{t+1}] = 0. \tag{132}$$

More generally, the forecast error is orthogonal to all variables measurable with respect to $\mathcal{I}_t$:

$$\mathbb{E}\big[\eta_{t+1} z_t\big] = 0 \quad \text{for all } z_t \text{ such that } z_t \in \mathcal{I}_t. \tag{133}$$

In particular,

$$\mathbb{E}\big[\eta_{t+1} x_t\big] = 0, \quad \mathbb{E}\big[\eta_{t+1} \mathbb{E}_t x_{t+1}\big] = 0. \tag{134}$$

Our empirical strategy is to regress a measure of $\mathbb{E}_t^{BR} x_{t+1}$ as well as x_{t+1} on x_t.[39] First, consider the linear regression:

$$\mathbb{E}_t^{BR} x_{t+1} = \delta_0 + \delta_1 x_t + \varepsilon_t^B, \tag{135}$$

where δ_0 and δ_1 are the population regression coefficients and ε_t^B is the projection error satisfying $\mathbb{E}[x_t\, \varepsilon_t^B] = 0$.

Second, consider the linear regression of x_{t+1} itself on x_t:

$$x_{t+1} = \rho_0 + \rho_1 x_t + \varepsilon_t^x, \tag{136}$$

where ρ_0 and ρ_1 are the population regression coefficients and ε_t^x is the corresponding projection error, with $\mathbb{E}[x_t\, \varepsilon_t^x] = 0$.

By the definition of linear projections, the slopes δ_1 and ρ_1 can be written in terms of covariances as

$$\delta_1 = \frac{\mathrm{Cov}\big(\mathbb{E}_t^{BR} x_{t+1},\, x_t\big)}{\mathrm{Var}(x_t)}, \tag{137}$$

$$\rho_1 = \frac{\mathrm{Cov}(x_{t+1},\, x_t)}{\mathrm{Var}(x_t)}. \tag{138}$$

Using the definition of cognitive discounting (129), we can rewrite δ_1 in (137) as

$$\delta_1 = \frac{\mathrm{Cov}\big(\mathbb{E}_t^{BR} x_{t+1},\, x_t\big)}{\mathrm{Var}(x_t)} = \frac{\mathrm{Cov}(\bar{m}\, \mathbb{E}_t x_{t+1},\, x_t)}{\mathrm{Var}(x_t)} \tag{139}$$

$$= \bar{m}\, \frac{\mathrm{Cov}(\mathbb{E}_t x_{t+1},\, x_t)}{\mathrm{Var}(x_t)}. \tag{140}$$

[39] Throughout, we assume that x_t is covariance-stationary with finite second moments, so that variances and covariances that appear below are well defined and time-invariant.

Next, we relate $\text{Cov}(\mathbb{E}_t x_{t+1}, x_t)$ to $\text{Cov}(x_{t+1}, x_t)$. Using (131),

$$\text{Cov}(x_{t+1}, x_t) = \text{Cov}(\mathbb{E}_t x_{t+1} + \eta_{t+1},\ x_t) \tag{141}$$

$$= \text{Cov}(\mathbb{E}_t x_{t+1}, x_t) + \text{Cov}(\eta_{t+1}, x_t)\,. \tag{142}$$

From condition (133), we have $\text{Cov}(\eta_{t+1}, x_t) = 0$, so

$$\text{Cov}(x_{t+1}, x_t) = \text{Cov}(\mathbb{E}_t x_{t+1}, x_t)\,. \tag{143}$$

Substituting (143) into (140) yields

$$\delta_1 = \bar{m}\,\frac{\text{Cov}(x_{t+1}, x_t)}{\text{Var}(x_t)} = \bar{m}\,\rho_1, \tag{144}$$

where the last equality uses the definition of ρ_1 in (138).

Equation (144) implies the identification result

$$\bar{m} = \frac{\delta_1}{\rho_1}. \tag{145}$$

Hence, we regress x_{t+1} on x_t to obtain an estimate of ρ_1, and regress $\mathbb{E}_t^{BR} x_{t+1}$ on x_t to obtain an estimate of δ_1. The ratio of the two then yields our estimate of $\bar{m}$.

The regressions (135) and (136) are linear projections, which exist under the mild assumptions of covariance stationarity and finite second moments. In order to use them to estimate $\bar{m}$, we need to impose that, first, the forecast error η_{t+1} under full-information rational expectations is orthogonal to x_t and that, second, the definition of cognitive discounting as specified in (129).

Data. To measure (potentially) behavioral expectations, we use the New York Fed Survey of Consumer Expectations (SCE) which is a monthly panel eliciting, among many other variables, households' one-year-ahead inflation expectations. To make these expectations comparable to our model, we assume that expectations are linear such that we can back-out the three-month-ahead expectations from the twelve-months-ahead expectations as: $\mathbb{E}_t^{BR} x_{t+3} = \frac{1}{4}\mathbb{E}_t^{BR} x_{t+12}$. Our sample period is from 2013M6 until 2024M9. As our measure of x_t, we use the quarter-on-quarter CPI inflation rate at monthly frequency.

We can then directly estimate $\bar{m}$ at quarterly horizon by estimating the regressions:

$$\mathbb{E}_t^{BR} x_{t+3} = \delta_0^{(3)} + \delta_1^{(3)} x_t + \varepsilon_t^{B,3}, \tag{146}$$

$$x_{t+3} = \rho_0^{(3)} + \rho_1^{(3)} x_t + \varepsilon_t^{x,3}. \tag{147}$$

Under rational expectations and the definition of cognitive discounting, the same covariance argument as above implies

$$\bar{m} = \frac{\delta_1^{(3)}}{\rho_1^{(3)}}. \tag{148}$$

The SCE asks households for their point forecasts and also asks them to assign probabilities to different potential outcomes. The SCE also reports the implied mean forecasts from this second approach. We use both measures as our measure $\mathbb{E}_t^{BR} x_{t+3}$. Throughout, we consider standard OLS regressions with heteroskedasticity-robust standard errors, as well as Huber robust regressions that control for outliers (Huber, 1964). Additionally, we also consider a winsorized sample where we drop the bottom 1% and the top 1% of expectations.

In addition to the SCE, we use the Survey of Consumers (SoC) from the University of Michigan. The SoC does not have the same panel structure as the SCE. Instead, we use average expectations across all consumers, as well as the average expectations within the four income quartiles. In addition to inflation expectations, using the SoC allows us to consider expectations about changes in the unemployment rate.

The SoC 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. (2024) to translate these categorical unemployment expectation into numerical expectations.

Focus on group e, where e could be the full sample or one of the four income quartiles, and let $q_t^{e,D}$, $q_t^{e,S}$ and $q_t^{e,U}$ denote the shares within 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 remain "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

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

This leaves us with one degree of freedom, namely the "roughly-the-same" threshold a. We make two assumptions. First, a is independent of the group. The second assumption is that we set $a = 0.35$ which means that if a household expects the change in unemployment to be less than 0.35 percentage points – which corresponds to one standard deviation of actual unemployment rate changes – she reports that she expects unemployment to be about the same as it is at the time of the survey (we verify that our results are qualitatively robust with respect to our choice of a).

We can then run the same regressions as for inflation expectations. When using the SoC, our sample period is substantially longer, ranging from 1979Q4 until 2025Q3. When focusing on unemployment, we drop the year 2020 due to the pandemic- and lockdown-related enormous spikes in unemployment.

Results. Table 4 shows the regression results for the SCE. The first column shows the estimate of ρ_1 from regression (147). Columns (2) to (7) show the estimates of δ_1 from regression (146) for the different specifications, discussed above. In particular, columns (2)-(4) show the estimates when using density-implied forecasts and when considering all observations with heteroskedasticity-robust standard errors, when using Huber weights, as well as when dropping the bottom and top 1%, respectively. Columns (5)-(7) show the estimates when considering point forecasts. Standard errors are in parenthesis. The third row provides the implied cognitive discounting parameter $\bar{m}$, computed following equation (148). We see that in all cases, the estimated $\bar{m}$ is clearly below 1, and ranges between 0.51 and 0.81. We also test for all specifications the null hypothesis that $\delta_1 = \rho_1$, in which case, the implied $\bar{m}$ would not be statistically significantly different from 1, and, hence, we could not rule out that expectations are rational. We find that we can reject the null of rational expectations at all conventional levels in all specifications (p-values range between 0.000 and 0.002).

Table 4: Estimating $\bar{m}$ using inflation expectations from the SCE

	$\hat{\rho}_1$	$\hat{\delta}_1^{D,all}$	$\hat{\delta}_1^{D,Huber}$	$\hat{\delta}_1^{D,winsorized}$	$\hat{\delta}_1^{P,all}$	$\hat{\delta}_1^{P,Huber}$	$\hat{\delta}_1^{P,winsorized}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coefficient	0.445	0.282	0.225	0.228	0.316	0.254	0.362
s.e.	(0.103)	(0.005)	(0.003)	(0.003)	(0.041)	(0.003)	(0.008)
Implied $\bar{m}$		0.63	0.51	0.51	0.71	0.57	0.81
N	156	169,603	169,603	163,517	172,342	172,340	168,791

Table 5 shows the results when considering the SoC. The results overall are similar to the SCE when considering point forecasts with an estimated average cognitive discounting parameter of 0.77. Columns (2)-(5) further indicate that there is some heterogeneity by income: higher income households tend to cognitively discount less with the implied $\bar{m}$ ranging between 0.72 for the lowest and 0.80 for the highest income quartile.

A similar picture emerges when considering unemployment change expectations, as shown

Table 5: Estimating $\bar{m}$ using inflation expectations from the SoC

	$\hat{\delta}_1^{all}$	$\hat{\delta}_1^{\text{1st Quartile}}$	$\hat{\delta}_1^{\text{2nd Quartile}}$	$\hat{\delta}_1^{\text{3rd Quartile}}$	$\hat{\delta}_1^{\text{4th Quartile}}$
	(1)	(2)	(3)	(4)	(5)
Coefficient	0.342	0.321	0.345	0.349	0.356
s.e.	(0.064)	(0.063)	(0.066)	(0.065)	(0.066)
Implied $\bar{m}$	0.77	0.72	0.77	0.78	0.80
N	183	183	183	183	183

in Table 6. Over the full sample, the implied cognitive discounting parameter is 0.66.[40] The estimates are more heterogeneous than for inflation expectations, as documented in columns (2)-(5): the estimated $\bar{m}$ increases monotonically with income from 0.52 in the lowest quartile to 0.85 in the top quartile.

Table 6: Estimating $\bar{m}$ using unemployment change expectations from the SoC

	$\hat{\delta}_1^{all}$	$\hat{\delta}_1^{\text{1st Quartile}}$	$\hat{\delta}_1^{\text{2nd Quartile}}$	$\hat{\delta}_1^{\text{3rd Quartile}}$	$\hat{\delta}_1^{\text{4th Quartile}}$
	(1)	(2)	(3)	(4)	(5)
Coefficient	0.328	0.259	0.298	0.358	0.426
s.e.	(0.021)	(0.019)	(0.020)	(0.028)	(0.033)
Implied $\bar{m}$	0.66	0.52	0.60	0.72	0.85
N	179	179	179	179	179

These estimates are consistent with other approaches used in the literature. For example, as Gabaix (2020) highlights, another approach is by matching estimated IS equations. Fuhrer and Rudebusch (2004), for example, estimate an IS equation and find that the coefficient in front of $\mathbb{E}_t \widehat{y}_{t+1}$ (what through the lens of our analytical framework in Section 3 would be $\psi_f \equiv \bar{m}\delta$) is approximately 0.65, which together with $\delta > 1$, would imply a $\bar{m}$ below 0.65. Benchimol et al. (2025) estimate a ψ_f for the US for three different episodes: 1996Q1–2004Q1, 2004Q2–2011Q4, and 2012Q1–2019Q4, and find values of 0.77, 0.71, and 0.73, respectively. Through the lens of our analytical framework in Section 3, these values indicate estimates of $\bar{m}$ somewhere between 0.65 and 0.77, depending on δ.

Gabaix (2020) further shows that one can rely on the regression approach of Coibion and Gorodnichenko (2015) to obtain a bound on $\bar{m}$, which is given by

$$\bar{m} \geq \left(\frac{1}{1 + b^{CG}} \right)^{1/h}, \tag{149}$$

where $1/h$ is a horizon-adjustment that has to be made when the expectation horizon that

[40] The estimated ρ_1 is 0.499 (not shown in the table).

is used to estimate b^{CG} is longer than one quarter. The seminal work by Coibion and Gorodnichenko (2015) finds $b^{CG} > 0$, consistent with $\bar{m} < 1$, i.e., underreaction. Their baseline estimate is $b^{CG} = 1.1$ for $h = 3$, implying a lower bound on $\bar{m}$ of 0.78. More 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 lower bounds for $\bar{m} \in [0.82, 0.83]$ for unemployment and $\bar{m} \in [0.73, 0.92]$ for inflation, and are largely consistent with our baseline value of 0.81. Note, however, that these estimates in Coibion and Gorodnichenko (2015) and Angeletos et al. (2021) pertain to professional forecasters and should therefore be seen as upper bounds on rationality.

Afsar et al. (2024) estimate cognitive discounting by estimating different versions of the New Keynesian model, treating $\bar{m}$ as an additional parameter to be estimated. They estimate $\bar{m}$ to be 0.46, both in models with and without backward-looking components (price indexation, as well as habit formation).

C The distributional effects of TFP changes in the data

In this section, we estimate how changes in aggregate TFP affect consumption inequality.

Data. We construct different measures of consumption inequality, directly borrowing from Coibion et al. (2021).[41] The data source is the Survey of Consumer Expenditures (CEX) as well as the Nielsen survey, covering the sample 1980-2015 at quarterly frequency. Throughout our analysis, we focus on nondurable consumption to be consistent with our model. For each data source, we consider two measures of consumption inequality: the share of total consumption of the bottom 50% as well as the Gini index. This leaves us with four measures of consumption inequality in total. We interpret a higher consumption share of the bottom 50% as well as a decrease in the Gini index as a reduction in consumption inequality.

As our measure of TFP changes, ΔTFP, we use the change in utilization-adjusted TFP from Fernald (2014), and regress these raw changes in TFP on a constant, four lags of real GDP, four lags of CPI inflation, as well as four lags of TFP changes. We then use the residual of that regression as our measure of ΔTFP. By residualizing the raw TFP changes, we purge those raw changes from endogenous fluctuations due to aggregate fluctuations to recover a measure of ΔTFP that is plausibly exogenous. Despite this, there may still be concerns that ΔTFP is endogenous; accordingly we interpret our results mainly as dynamic correlations and not necessarily as causal relations.

[41] We use the replication package accessed through the website of the American Economic Association, downloaded in June 2025.

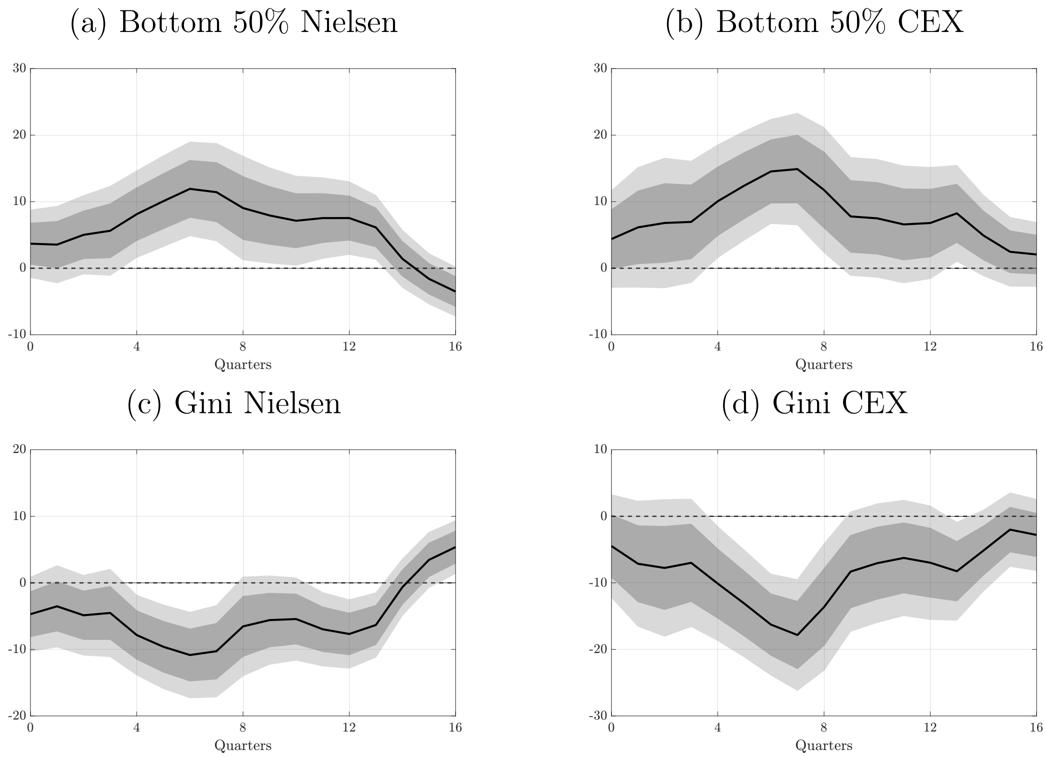
Empirical specification. To examine how changes in TFP affect consumption inequality, we estimate local projections, following Jordà (2005). In particular, we estimate the following regression:

$$\frac{1}{4}\sum_{j=0}^{3} y_{t+h+j} = \alpha_h + \beta_h \Delta TFP_t + \Gamma_h X_t + u_{t,h}, \tag{150}$$

where y_{t+h} denotes the dependent variable $h \geq 0$ periods after the change in TFP, with the dependent variable being the four different consumption inequality measures discussed above. We use the average over four quarters of the dependent variable to smooth the responses, following Fieldhouse and Mertens (2023). If we do not do this, the estimated IRFs are less smooth, but are otherwise identical. α_h is a horizon-specific intercept, ΔTFP_t denotes the (residualized) change in TFP in period t. The coefficient of interest is β_h, which captures the response of the dependent variable h periods after the TFP change. Controls X_t contain four lags of ΔTFP, as well as four lags of the dependent variable, the federal funds rate, CPI inflation, the CPI index, real GDP per capita growth, and real GDP per capita. All our data, except for the inequality measures and TFP, are obtained from the FRED data base.

Results. Figure 5 shows the results. Panels (a) and (b) report responses for the consumption share of the bottom 50%—from Nielsen and CEX, respectively—while panels (c) and (d) report responses for the Gini coefficient—again from Nielsen and CEX, respectively. The coefficients are normalized such that a coefficient of 10, for example, indicates that the respective variable increases by 10% of a standard deviation in response to a negative one standard deviation TFP change. The dark gray areas indicate 68% confidence intervals and the light gray areas 90% confidence intervals, with heteroskedasticity-robust standard errors. Figure 5 shows a rather clear pattern: In response to a negative TFP change, consumption inequality decreases robustly across all four different inequality measures. This empirical finding is consistent with our model predictions. Importantly, the inequality response indicates that households with lower consumption levels decrease their consumption by less than households with higher consumption levels in response to a negative TFP change. This unequal exposure of households to TFP changes emerges endogenously in our model whenever we target the previously documented unequal exposure of households to monetary policy shocks. Despite the caveat of potential endogeneity issues, we view these findings as empirical confirmation of a key channel in our model: in response to a negative TFP change, consumption inequality decreases as households with higher consumption levels reduce their consumption expenditures more strongly.

Figure 5: Consumption inequality after negative change in TFP



Notes: This figure shows the estimated impulse response functions of the respective variables to a one standard deviation decrease of residualized TFP changes.

D Appendix to the quantitative model

This Section provides additional results stemming from our full, quantitative behavioral HANK model discussed in Section 4.

D.1 Monetary policy in the quantitative model

In this Section, we analyze the transmission of monetary policy in our full, quantitative HANK model using the baseline calibration in Section 4.1, with one exception. In line with our analytical results and the literature on forward guidance in HANK models (see, e.g., (McKay et al., 2016; Farhi and Werning, 2019)), we assume that monetary policy directly controls the real interest rate. Specifically, we consider the following real-rate rule (in log-deviations from steady state) which is a slightly modified version relative to Section 3:

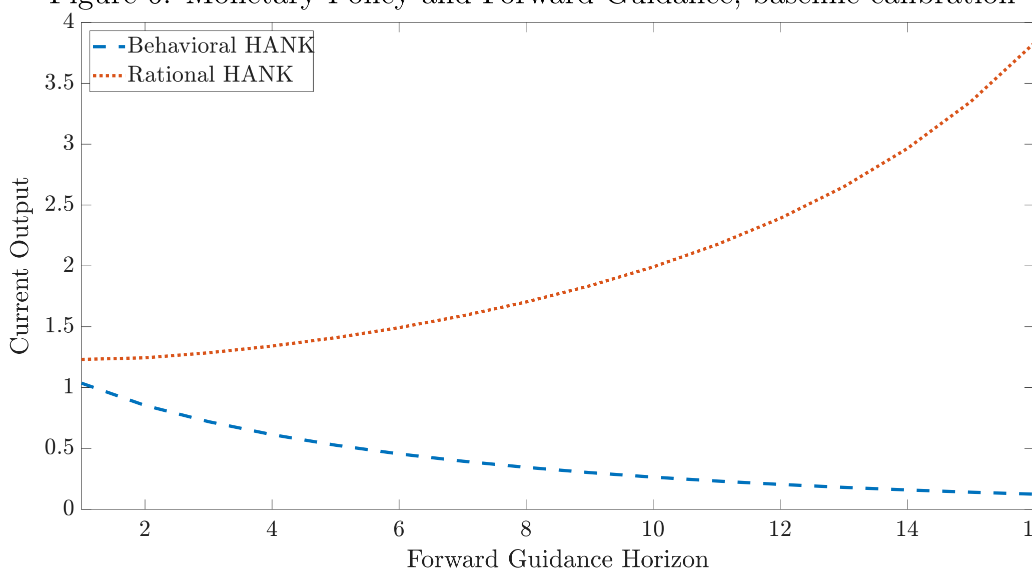
$$i_t = \begin{cases} E_t \pi_{t+1} + \epsilon^r_{t-k}, & \text{if } t < k + 10, \\ \theta_\pi \pi_t & \text{otherwise.} \end{cases} \tag{151}$$

The first line in (151) implies that—apart from the period in which the exogenous real-rate change occurs—monetary policy keeps the real rate at its steady-state level until 10 periods after the change materializes. The second line in equation (151) says that thereafter, policy leans against persistent inflation by reverting to the baseline Taylor rule from Section 4. This assumption ensures that all model variants that we consider are well behaved.

We then study monetary policy news shocks with horizon $k \in [0, 16]$. Figure 6 plots the response of current output (at time t) as a function of the forward-guidance horizon k for the behavioral HANK model (blue dashed line) and its rational-expectations counterpart (orange dotted line). In the behavioral HANK model, the effect of announced policy shocks on current output declines quickly with the horizon and converges to zero. An announced real-rate change 5 periods ahead moves current output only about half as much as an equally sized contemporaneous real-rate change. Thus, also the quantitative behavioral HANK model resolves the forward guidance puzzle.

This is in stark contrast to the rational HANK model, in which the effectiveness of announced monetary policy shocks strongly increases in their horizon, in line with our analytical insight in Section 3. For example, an announced real-rate change 13 periods ahead is about twice as powerful as a contemporaneous change of the same size. We conclude that, without simultaneously accounting for behavioral expectations, the empirically documented unequal exposure substantially exacerbates the forward guidance puzzle.

Figure 6: Monetary Policy and Forward Guidance, baseline calibration

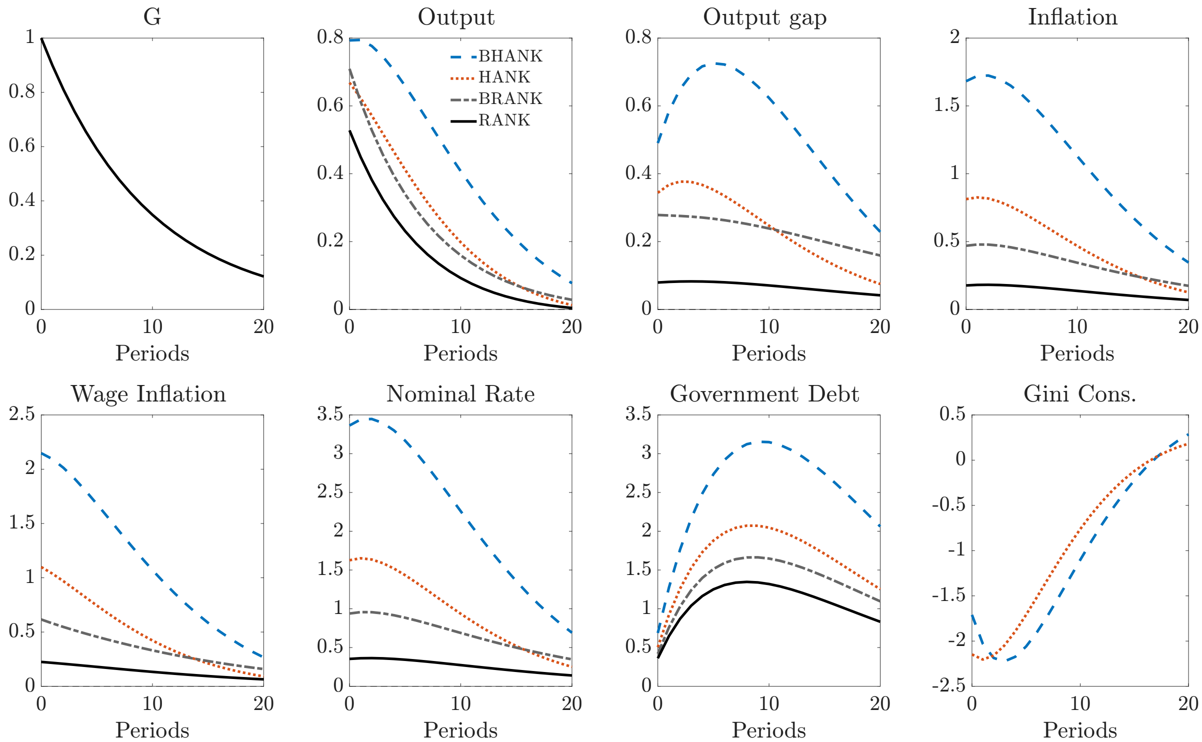


Note: This figure shows the response of current output in period 0 to a 200 basis point anticipated i.i.d. monetary policy shocks occurring at different horizons k.

D.2 Government spending in the quantitative model

This Section presents the impulse responses in response to an increase in government spending. In particular, we assume that government spending follows an exogenous AR(1) process akin to equation 7, with a persistence of 0.9. Figure 7 shows the impulse responses for the same four models as Figure 1 does for the TFP shock.

Figure 7: Government spending shock



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

D.3 Additional results

In this Section, we present two additional results. First, Table 7 decomposes the amplification of inflation on impact of our behavioral HANK model with varying degrees of unequal exposure and varying degrees of cognitive discounting relative to the acyclical HANK model similar as Table 2 in Section 4.4. See Section 4.3 for details on the decomposition.

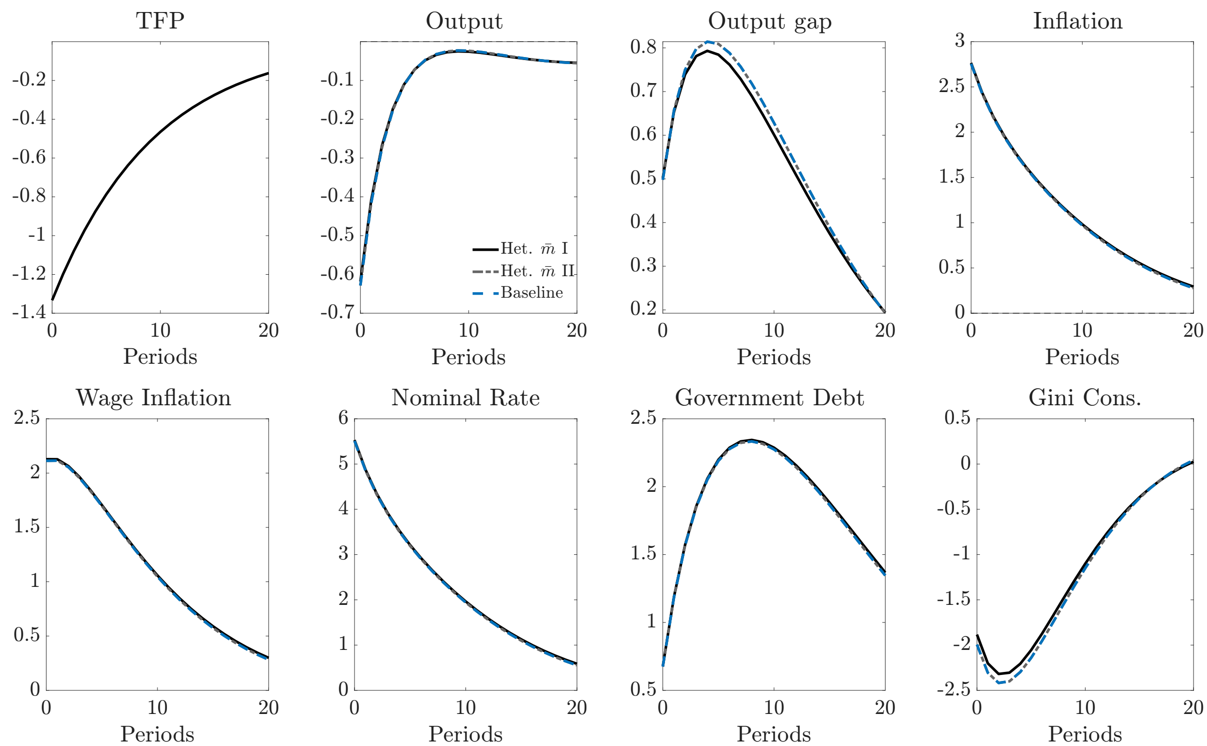
Table 7: Robustness: Unequal exposure and degree of bounded rationality

	Comparison with acyclical HANK			
Scenario	Amplification of π	Unequal exposure	Cognitive discounting	Interaction
Baseline	67%	15%	26%	26%
Less unequal exposure	43%	7%	26%	10%
More unequal exposure	107%	27%	26%	55%
Higher $\bar{m}$	45%	15%	15%	14%
Lower $\bar{m}$	109%	15%	42%	52%
Het. $\bar{m}$, specification 1	69%	15%	27%	27%
Het. $\bar{m}$, specification 2	69%	15%	27%	26%

Note: *Less unequal* refers to $\zeta = -0.425$ (instead of $\zeta = -1.0$); *more unequal* to $\zeta = -2$. Higher $\bar{m}$ refers to $\bar{m} = 0.90$ (instead of $\bar{m} = 0.81$), lower to $\bar{m} = 0.60$. *Het. $\bar{m}$, specification 1* refers to the case in which we target our empirical results based on inflation expectations and *Het. $\bar{m}$, specification 2* refers to the case in which we target our empirical results based on unemployment expectations.

Second, Figure 8 compares the impulse responses of our baseline model (*Baseline*; blue, dashed lines) to the impulse responses in the two behavioral HANK models with heterogeneous degrees of cognitive discounting discussed in Section 4.5. The black, solid lines (*Het. $\bar{m}$ I.*) show the impulse responses in the calibration using our estimates based on inflation expectations, whereas the gray, dashed-dotted lines show the impulse responses in the calibration using our estimates based on unemployment expectations.

Figure 8: Inflationary supply shock: Baseline vs. heterogeneous $\bar{m}$



Note: This figure shows the impulse responses after a productivity shock that decreases potential output by 1%. It compares our baseline model with homogeneous cognitive discounting to two heterogeneous calibrations: *Het. $\bar{m}$ I* which targets our estimates based on inflation expectations and, *Het. $\bar{m}$ II*, which targets our estimates based on unemployment expectations. Output and the output gap are shown as percentage deviations from steady state output, the nominal interest rate and inflation as annualized percentage points, and the government debt level as percentage point deviations of the debt-per-annual GDP level. The lower-right figure shows the change in the consumption Gini index as a percentage deviation from the stationary equilibrium.