Heterogeneity and Aggregate Fluctuations: Insights from TANK Models

Davide Debortoli
ICREA-UPF, CREI and BSE

Jordi Galí
CREI, UPF and BSE

May 2024

Abstract

We analyze the merits and limitations of simple tractable New Keynesian models (RANK and TANK) in accounting for the aggregate predictions of Heterogenous Agent New Keynesian models (HANK). By means of comparison of a number of nested HANK models, we investigate the role played by (i) idiosyncratic income risk, (ii) a binding borrowing constraint, and (iii) a portfolio choice between liquid and illiquid assets. We argue that the effects of household heterogeneity can be largely understood looking at the differential behavior of two types of households: hand-to-mouth and unconstrained. We find that a suitably specified and calibrated TANK model (which abstracts from idiosyncratic income risk) can capture reasonably well the aggregate implications of household heterogeneity and the main channels through which it operates. That ability increases in the presence of a policy rule that emphasizes inflation stability. In the limiting case of a strict inflation targeting policy, heterogeneity becomes irrelevant for the determination of aggregate output.

JEL Classification Numbers: E32, E52

Keywords: monetary policy, idiosyncratic income risk, incomplete markets, representative household, New Keynesian model, HANK models

---

Email: `davide.debortoli@upf.edu` and `jgali@crei.cat`. This is a heavily revised version of a paper previously circulated under the tile "Monetary policy with Heterogeneous Agents: Insights from TANK models". We thank Antonio Giribaldi, Kadir Özen and Danila Smirnov for excellent research assistance. We have benefited from comments by Gianluca Violante, Florin Bilbiie, John Leahy, Valerie Ramey, Johannes Wieland, Matt Rognlie, Paolo Surico, and participants at the CREi Faculty Lunch, NBER SI EFG, NBER Macro Annual, CREI Workshop.on Heterogeneity and Economic Fluctuations, and CIGS Conference on Macroeconomic Policy and Theory. Galí acknowledges the financial support from the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant 882332-HEMPEF). Both authors acknowledge financial support from the Agencia Estatal de Investigación (AEI), through the Severo Ochoa Programme for Centres of Excellence in R&D (Barcelona School of Economics CEX2019-000915-S) and grant PID2020-116268GB-I00 (Debortoli), and from the Generalitat de Catalunya, through the CERCA Programme.

1 Introduction

Differences in income, wealth, education and, more generally, economic fortune are a pervasive feature of modern economies. Yet, macroeconomists have largely ignored such heterogeneity for decades, under the widespread belief that it is largely irrelevant for understanding *aggregate* outcomes and their interaction with macro policies. Consistently with that view, and given its analytical convenience, the assumption of an infinitely-lived representative household became a staple of macro models, without raising any eyebrows. The representative agent New Keynesian (henceforth, RANK) model is a prominent example of that approach.

An emerging class of models, often referred to as HANK (for Heterogeneous Agent New Keynesian), has challenged the dominance of the representative household paradigm. Heterogeneity in those models is usually introduced by assuming that households experience idiosyncratic income shocks that cannot be insured against, due to incomplete financial markets. The presence of *some* assets allows households to partly smooth their consumption, while giving rise to a non-degenerate wealth distribution. The latter thus becomes one of the model's state variables, which evolves over time in response to aggregate shocks, also influencing how the economy responds to those shocks. The previous features are then combined with a supply block that is similar (if not fully identical) to that characterizing the standard New Keynesian model. In particular, the supply block assumes monopolistically competitive firms as well as nominal rigidities, thus allowing monetary policy to have real effects.[1]

In the present paper we seek to advance our understanding of the implications of heterogeneity for aggregate fluctuations in the New Keynesian (NK) model. Our ultimate goal is twofold. Firstly, we want to understand the mechanisms through which heterogeneity, in the form of idiosyncratic income risk, affects aggregate fluctuations and the transmission of shocks in HANK models. As discussed below, we do so by comparing the properties of a number of nested HANK models, a strategy that allows us to isolate the role played by different elements found in conventional versions of those models. Secondly, we want to investigate whether there are simple tractable models that can capture reasonably well the mechanisms that we identify

[1] See, e.g. Kaplan et al (2018) and Auclert et al. (2023) for examples of such models. Our focus in the present paper is on models with household heterogeneity, thus abstracting from firm heterogeneity. The latter is at the core of the literature on Ss pricing and investments policies, firm dynamics, etc. and may potentially have important implications for aggregate behavior.

as most relevant in richer models.[2]

Throughout, our analysis relies heavily on the distinction between hand-to-mouth and unconstrained households. That distinction arises endogenously in conventional HANK models. It is also central to Two-Agent New Keynesian (TANK) models, though in the latter it is introduced in a rather stark way.[3] While the tractability and transparency of TANK models is generally viewed as an advantage relative to conventional HANK models, it is not clear that they can capture well the aggregate implications of the latter. This motivates a key objective of our analysis, namely, the evaluation of the ability of TANK models to approximate the aggregate equilibrium dynamics generated by HANK models.

To be clear, we are not the first to compare the properties of a baseline HANK model to those of simpler, more tractable frameworks.[4] The key difference lies in the particular approach we adopt, which stresses the differential behavior of hand-to-mouth vs unconstrained households. It is in that sense that we analyze the properties of HANK models through the lens of their TANK counterparts, thus motivating our title. We believe this provides an interesting and useful perspective to understanding the mechanisms at work in HANK models.

Our analysis starts by laying out a HANK model that we use as a baseline throughout the paper. Our baseline HANK model describes an economy with a continuum of households subject to idiosyncratic income shocks. All households have access to two assets: bonds and stocks. Bonds are fully liquid, while stocks are fully or partly illiquid. Holdings of the two assets are subject to constraints: a borrowing constraint in the case of bonds, a non-negativity constraint in the case of stocks.

We view the presence of idiosyncratic income shocks (in the absence of complete markets) as a defining feature of HANK models, relative to tractable models like RANK or TANK. It is also a main factor behind its nontrivial equilibrium dynamics, since it gives rise to a time-varying wealth distribution which becomes an infinite-dimensional state variable. By way of contrast, we define a *tractable* model as one that abstracts from the presence of idiosyncratic income risk, assuming instead a small number of household types. For each type, a time-invariant set of

[2]HANK models may also be used for understanding the *distributional* impact of shocks or policy interventions. Our focus here is exclusively on its implications for aggregate fluctuations.

[3]See, e.g., Galí et al. (2007) and Bilbiie (2008). The idea of partitioning households between hand-to-mouth and unconstrained can be traced back to Campbell and Mankiw (1989).

[4]See, e.g. Auclert, Rognlie and Straub (2023) for a recent example.

households that are identical, both ex-ante and ex-post, is assumed. The equilibrium conditions of such tractable models, linearized around a steady state, can be solved analytically, arguably rendering them more suitable for use in the classroom or for communication with policy makers.

Our approach in the present paper consists of (i) analyzing the equilibrium properties of three versions of the HANK model[5] (ii) proposing a tractable counterpart for each of them and (iii) assessing the extent to which the equilibrium properties of the tractable model provide a reasonable approximation to those of the corresponding HANK economy.[6]

We proceed sequentially by considering models with increasing complexity. Thus, we start by analyzing a version of the HANK model (which we refer to as HANK-I) in which borrowing constraints are not binding in equilibrium and stocks are fully illiquid, a framework we took as a benchmark in earlier work (Debortoli and Galí (2024)). We show that the New Keynesian model with a representative agent (RANK), which also displays no binding borrowing constraints, provides a good approximation to the aggregate behavior of the HANK-I model.

Next, we consider a version of the HANK model (HANK-II) similar to HANK-I except for the fact that the borrowing constraint on bonds is binding in equilibrium for a (time-varying) fraction of households. In the equilibrium of HANK-II we can identify two types of households at any point in time: those which are unconstrained and those for which the borrowing constraint is binding. We refer to the latter as hand-to-mouth since their marginal propensity to consume is one. We compare the equilibrium properties of the previous model to those of simple TANK models, in which a time-invariant subset of households behaves in a hand-to-mouth fashion. First, we show that the standard version of the TANK model, which we refer to as TANK-I, fails to capture two key channels determining the response of aggregate consumption to aggregate shocks in HANK-II: (i) the *interest rate exposure channel*, whereby changes in interest rates have a direct effect on the net cash-on-hand and, hence, consumption of hand-to-mouth households, and (ii) *the income distribution channel*, which captures the impact of changes in the average price markup on the distribution of income between unconstrained and hand-to-mouth households, given that the average labor productivity and hence the relative importance of labor vs capital income differs across those two household types. Those missing channels

[5]See below for a detailed explanation of each version.

[6]See Faia and Shabalina (2024) for recent work in the same spirit, using a larger range of HANK models.

prevent the TANK-I model from approximating well some of the aggregate properties of HANK-II. We then propose a straightforward modification of TANK-I, which we label TANK-II, where the hand-to-mouth (i) own a fraction of the existing (illiquid) stocks, (ii) have a lower labor productivity than the unconstrained and (iii) are permanently against an implicit borrowing constraint, thus servicing the interest on a constant level of debt. We show that a suitably calibrated version of the TANK-II model approximates well the aggregate properties and key underlying mechanisms of the HANK-II model.

Finally, we analyze a version of the HANK model (labeled as HANK-III) which is arguably closer to the baseline HANK models found in the recent literature. The main difference with respect to HANK-II is that HANK-III relaxes the assumption of *fully* illiquid stocks, by allowing adjustments in the holdings of the illiquid asset, subject to a transaction cost and a no short-sale constraint. The resulting model is similar to versions of the HANK model found in the literature, generally referred to as *two-asset* HANK models. A property of the equilibrium of HANK-III is that at any point in time three different types of households coexist: (i) the unconstrained, (ii) those for whom the borrowing constraint on the liquid asset is binding, but not the short-sale constraint on the illiquid asset, and (iii) those for which both constraints are binding. Following the literature we refer to (ii) and (iii) as the *wealthy hand-to-mouth* and the *poor hand-to-mouth*, respectively. We show that the introduction of portfolio adjustment costs, and the emergence of wealthy hand-to-mouth agents alters significantly some key properties of the model: in particular, the response to technology shocks. The main reason for this is that in HANK-III changes in dividends are not immediately converted into cash-on-hand, since that conversion requires liquidation of part of the illiquid asset.

We propose a tractable counterpart to HANK-III, which we refer to as TANK-III. In the latter, and relative to TANK-II, we introduce in a parsimonious way the distinction between poor and wealthy hand-to-mouth households found in HANK-III, while abstracting from the presence of idiosyncratic income risk. The resulting model matches reasonably well the predictions of HANK-III.

The bulk of our analysis below is carried out under the assumption of an exogenous real interest rate. We choose that approach so that the response of aggregate variables to different shocks is not affected by any particular assumption regarding the monetary policy rule, which would generally lead to different paths of the real rate across environments that differ in terms of the behavior of variables that the central bank responds to. When we relax that assumption and assume instead a more realistic Taylor-type rule as a description of monetary policy we find that the similarity in the aggregate properties of the different models considered (RANK, TANK, and HANK) increases dramatically. The intuition for that result is straightforward: all the previous models share a common supply block, which features a New Keynesian Phillips curve displaying the "divine coincidence" property. Policies that tend to stabilize inflation (like the assumed Taylor rule) also close the gap between output and its natural counterpart, which is invariant to heterogeneity. As a result, equilibrium output tends tend to converge across models. Furthermore, we show that in the limiting case of a strict inflation targeting, policy heterogeneity becomes completely irrelevant for the determination of aggregate output.

Our paper is related to two main strands of the literature. On the one hand, the HANK literature which explores the implications of introducing household idiosyncratic income risk and incomplete markets into an otherwise standard New Keynesian framework with nominal rigidities. Some examples are the works of Guerrieri and Lorenzoni (2017), McKay et al. (2016), McKay et al. (2016), Farhi and Werning (2017), Gornemann et al. (2016), Kaplan et al. (2018), McKay and Reis (2016), Werning (2015), Auclert (2017), Auclert et al. (2023), Luetticke (2017), and Ravn and Sterk (2021), among others.

On the other hand, the paper builds on a literature that develops simple, tractable models that assume some stylized form of ex-ante household heterogeneity with regard to access to financial markets. That literature was pioneered by Campbell and Mankiw (1989), who proposed a simple two-agent framework with unconstrained and hand-to-mouth households, thus departing from the representative household formalism dominant at that time. Galí et al. (2007) and Bilbiie (2008) are early examples embedding that structure into the New Keynesian model, giving rise to the so-called TANK models. More recent contributions include Bilbiie (2020), Broer et al. (2020), and Cantore and Freund (2021), among others.[7]

Our paper connects the two literatures in two respects. First, we rely on the distinction

[7]Similarly, Bilbiie (2018) uses a TANK model to illustrate the "direct" and "indirect" effects of monetary policy shocks emphasized by Kaplan et al. (2016) in a more general HANK model. Farhi and Werning (2017a) use a variety of TANK models to analyze the size of fiscal multipliers in a liquidity trap and in currency unions. Ravn and Sterk (2021) build a tractable heterogeneous agent model with nominal rigidities and labor market frictions, giving rise to endogenous unemployment risk.

between hand-to-mouth vs unconstrained households –the hallmark of TANK models– in order to better understand the mechanisms at work in HANK models. Secondly, we assess the ability of tractable models in the spirit of TANK to approximate the aggregate properties of richer HANK models.

The remainder of the paper is organized as follows. Section 2 introduces the baseline elements that are common across all the HANK models considered. Section 3 analyzes a HANK model without binding borrowing constraints (HANK-I), and its tractable RANK counterpart. Section 4 introduces an occasionally binding borrowing constraint (HANK-II), and compares its properties to a standard TANK model as well as a modified version of the latter (TANK-II). Section 5 considers a HANK model with liquid bonds and partially illiquid stocks (HANK-III), in comparison to a suitably modified TANK model (TANK-III). Section 6 analyzes the consequences of endogenizing monetary policy. Section 7 relates some our findings to the existing literature. Section 8 concludes.

2 A Baseline HANK Model

In this section we describe the elements of a baseline HANK framework that are shared across the different models considered below.

2.1 Households

We assume a continuum of infinitely-lived households, indexed by $j \in [0,1]$. Each household seeks to maximize utility $\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(C_t(j), \mathcal{N}_t(j))$, where $C_t(j) \equiv \left(\int_0^1 C_t(i,j)^{1-\frac{1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}}$ is an index of the quantities consumed of the different available goods $i \in [0,1]$, with ϵ denoting the elasticity of substitution among goods. $\mathcal{N}_t(j)$ denotes hours worked. We assume $U(C, \mathcal{N}) \equiv \frac{C^{1-\sigma}-1}{1-\sigma} - \frac{\mathcal{N}^{1+\varphi}}{1+\varphi}$, where σ is the coefficient of relative risk aversion and φ is the inverse of the Frisch labor supply elasticity.

Household j's labor income in period t is given by $\Xi_t(j) W_t \mathcal{N}_t(j)$, where W_t is the (real) wage per efficiency unit of labor and $\Xi_t(j) \equiv \exp\{\zeta_t(j)\}$ is an exogenous idiosyncratic shock to the household's supply of efficiency units per hour worked, with $\int_0^1 \Xi_t(j) dj = 1$. For brevity we refer to $\Xi_t(j)$ as the idiosyncratic income shock.

There are two assets in the economy. The first asset is a one-period riskless real bond, with holdings by household j denoted by $B_t(j)$.[8] Bonds are assumed to be *fully liquid*, i.e. they can be bought and sold in a competitive market with no transaction costs, yielding a safe gross real return R_t. Bond holdings are subject to the borrowing constraint

$$R_t B_t(j) \geq \underline{B} \tag{1}$$

where $\underline{B}$ indicates the borrowing limit.

The second type of asset ("stocks") are shares in firms' equity, which generate an aggregate dividend D_t every period. In two of the models considered (HANK-I and II), stocks are illiquid, with each household being allocated a fraction of firms' profits according to a rule specified below.

By contrast, in HANK-III stocks are held directly by a competitive financial intermediary. Households can borrow or lend from the financial intermediary at a competitive interest rate R_t, and subject to a borrowing constraint. In addition they can maintain in the latter an equity account to which they can add or withdraw funds subject to a portfolio adjustment cost, following the formalism in Kaplan et al. (2018) and Auclert et al. (2023).

The resulting period budget constraint for household j can thus be written as:

$$C_t(j) + B_t(j) = R_{t-1} B_{t-1}(j) + \Xi_t(j) W_t \mathcal{N}_t(j) + F_t(j)$$

where $F_t(j)$ denotes the net additions to the household's cash-on-hand associated with equity holdings (dividends and/or sales/purchases of equity positions).

We assume a wage schedule given by

$$W_t = \mathcal{M}_w C_t^{\sigma} N_t^{\varphi} \tag{2}$$

where $C_t \equiv \int_0^1 C_t(j)dj$ and $N_t \equiv \int_0^1 \mathcal{N}_t(j)dj$ denote aggregate consumption and hours, respectively. Note that $C_t^{\sigma} N_t^{\varphi}$ can be interpreted as an "average" marginal rate of substitution, with $\mathcal{M}_w > 1$ thus playing the role of an average wage markup, determined by workers' market

[8] In our baseline model, we assume a real bond in order to avoid large reallocations of wealth resulting from unexpected inflation, since the latter is very sensitive to the assumed properties of the Phillips curve.

power.[9] Given the wage, firms determine the quantity of hours hired, which are assumed to be distributed uniformly across all households, i.e. $\mathcal{N}_t(j) = N_t$ for all $j \in [0,1]$.[10] Accordingly, household j takes its labor income $\Xi_t(j)W_tN_t$ as exogenous.

2.2 Firms

The supply side, common to all the models analyzed below, is kept as simple as possible. In particular we make assumptions that guarantee that it is not affected by the presence of heterogeneity. This allows us to focus on the impact of the latter on aggregate demand (which coincides with aggregate consumption in our simple model).

On the production side, we assume a continuum of firms, indexed by $i \in [0,1]$. Each firm produces a differentiated good with the linear technology

$$Y_t(i) = A_t N_t(i) \tag{3}$$

where $N_t(i)$ is the quantity of labor (expressed in efficiency units) hired by firm i, and $A_t \equiv \exp\{a_t\}$ is an exogenous technology parameter common to all firms, which follows the $AR(1)$ process $a_t = \rho_a a_{t-1} + \varepsilon_t^a$.

Each firm sets the price of its good optimally each period, subject to a quadratic adjustment cost $\frac{\xi}{2}P_tY_t\left(\frac{P_t(i)}{P_{t-1}(i)}-1\right)^2$ where $\xi > 0$, and a sequence of demand constraints $Y_t(i) = (P_t(i)/P_t)^{-\epsilon}Y_t$, where Y_t denotes aggregate output. Profit maximization, combined with the symmetric equilibrium conditions $P_t(i) = P_t$ and $Y_t(i) = Y_t$ for all $i \in [0,1]$, implies:

$$\Pi_t(\Pi_t - 1) = \mathbb{E}_t\left\{\Lambda_{t,t+1}\left(\frac{Y_{t+1}}{Y_t}\right)\Pi_{t+1}(\Pi_{t+1}-1)\right\} + \frac{\epsilon-1}{\xi}\left(\frac{\mathcal{M}}{\mathcal{M}_t}-1\right) \tag{4}$$

where $\Pi_t \equiv P_t/P_{t-1}$ is (gross) price inflation rate, $\mathcal{M}_t$ is the gross price markup, with $\mathcal{M} \equiv \epsilon/(\epsilon-1) > 1$ being its desired (or flexible price) counterpart, τ is a constant employment subsidy and $\Lambda_{t,t+1}$ is the firm's stochastic discount factor. Aggregate profits are given by

[9] Strictly speaking, the interpretation of $C_t^\sigma N_t^\varphi$ as an average MRS will be valid *in equilibrium* only if $\sigma = 1$. The assumed wage equation guarantees that idiosyncratic shocks are not reflected in unrealistic differences in the quantity of labor supplied across households. In addition, it simplifies the analysis by making labor income exogenous to individual decisions. Finally, it is one of the assumptions that guarantee that the supply block of the model corresponds to that in the standard New Keynesian model, thus being insulated from the presence of heterogeneity.

[10] Under our assumptions, if $\mathcal{M}_w \gg 1$, all households will be willing to supply the amount of labor demanded by firms, as long as shocks are not too large.

$D_t = Y_t\Delta(\Pi_t) - (1-\tau)W_tN_t - T_t$ where $\Delta(\Pi_t) \equiv 1 - (\xi/2)(\Pi_t - 1)^2$ and T_t are lump-sum taxes on firms that finance the employment subsidy. We set τ so that $\mathcal{M}(1-\tau) = 1$, which implies that the zero inflation steady state is efficient and involves zero profits. The latter property guarantees that the distribution of wealth and consumption in the stochastic steady state is not affected by assumptions on the allocation of profits across households, and instead depends exclusively on the presence of idiosyncratic income shocks.

A first-order approximation of (4) around the zero-inflation steady state yields the inflation equation

$$\pi_t = \beta\mathbb{E}_t\{\pi_{t+1}\} - \lambda\widehat{\mu}_t \tag{5}$$

where $\widehat{\mu}_t \equiv \log(\mathcal{M}_t/\mathcal{M})$ and $\lambda \equiv (\epsilon - 1)/\xi$. Noting that $\mathcal{M}_t = A_t/(1-\tau)W_t$ and using the wage schedule above we obtain

$$\widehat{\mu}_t = -(\sigma + \varphi)\widehat{y}_t + (1 + \varphi)a_t \tag{6}$$

where $y_t \equiv \log(Y_t/Y)$. Moreover, we can determine the (log) natural output (i.e. the equilibrium level of output under flexible prices), denoted by y_t^n, by setting $\widehat{\mu}_t = 0$ in (6). This yields $y_t^n \equiv \frac{1+\varphi}{\sigma+\varphi}a_t$. Thus, under our assumptions, natural output is not affected by the presence of idiosyncratic income risk, the set of assets available, or the existence of binding borrowing constraints.

Combining (5) and (6) yields a version of the familiar New Keynesian Phillips curve

$$\pi_t = \beta\mathbb{E}_t\{\pi_{t+1}\} + \kappa(y_t - y_t^n) \tag{7}$$

where $y_t - y_t^n$ is the output gap, and $\kappa \equiv \lambda(\sigma + \varphi)$. Note that equation (7) is invariant to the presence of household heterogeneity and the nature of the latter.

2.3 Monetary Policy

Regarding monetary policy, we assume that the central bank controls the real interest rate R_t. In our baseline specification we assume that the central bank keeps the real interest rate constant in the face of aggregate shocks (other than monetary policy shocks). Under that approach the response of aggregate variables to different (non monetary) shocks is not affected by any particular assumption regarding the monetary policy rule, which would generally lead to different paths of the real rate across environments associated with different responses of aggregate variables.[11] On the other hand, when we analyze the effects of monetary policy shocks we assume an exogenous process for the real rate, given by

$$R_t = R \exp\{v_t\}$$

where $v_t = \rho_v v_{t-1} + \varepsilon_t^v$.

The assumptions of a real bond and an exogenous process for its real return jointly imply that we can solve for equilibrium output without any reference to the supply side of the model and, hence, independently of the specific form of the Phillips curve. In section 6 we bring back (7) into the analysis when studying the consequences of endogenizing monetary policy, for which purpose we assume a simple Taylor-type rule for the nominal rate while allowing also for nominal bonds.

2.4 Baseline Calibration and Solution Method

The baseline calibration of our economy is summarized in Table 1. Each period is assumed to correspond to a quarter. We set the coefficient of risk aversion $\sigma = 1$, which corresponds to log utility, and the (inverse) Frisch elasticity of substitution $\varphi = 1$. In addition, we set the average wage markup $\mathcal{M}_w = 1.10$, the elasticity of substitution among good varieties $\epsilon = 11$, which implies an average price markup of $\mathcal{M} = 1.10$, and the price adjustment cost parameter ξ so that the resulting slope of the Phillips Curve is $\kappa = 0.10$, in line with available estimates.

Following Auclert et. al (2021), we calibrate the parameters of the K-state Markov process for idiosyncratic income using the Rouwenhorst method in order to match the volatility and persistence of an AR(1) process $\zeta_t(j) = \rho_\zeta \zeta_{t-1}(j) + \varepsilon_t^\zeta(j)$, where $\varepsilon_t^\zeta(j) \sim N(0, \sigma_\zeta \sqrt{1-\rho_\zeta^2})$, with $\rho_\zeta = 0.966$ and $\sigma_\zeta = 0.5$.

For each model considered below, we calibrate the discount factor β so that the real risk-free rate is 2 percent (in annual terms) in the steady state. This results in a discount factor $\beta = 1/R = 0.995$ in the RANK and TANK models, and to $\beta = (0.9937, 0.9838, 0.9905)$ for HANK-I, HANK-II and HANK-III, respectively.

[11] A similar approach is followed by Woodford (2011) when studying the size of the fiscal multiplier in a New Keynesian model.

For the economy without a binding borrowing constraint of section 3 we set the borrowing limit at $\underline{B} = -Y \exp\{\min_j \zeta(j)\}/r$, which constitutes the "natural debt limit," given aggregate output and interest rate at their steady state values (Y, r).[12] For the remaining economies, we set $\underline{B} = -2Y$, which implies a steady state share of hand-to-mouth households of 30 percent, as suggested by the evidence in Kaplan et al (2014).

The calibration of parameters pertaining to the determination of $F_t(j)$ is discussed below in the respective sections.

Regarding the numerical solution method, we build a grid for individual assets of 500 points, equally distanced (in logs) between the lower bound described above, and an upper bound set to 300 times quarterly income for the model of section 3 and to 50 times income for the other models.

For given values of the real interest rate, we solve for the households' policy functions using the endogenous gridpoints method described in Carroll (2006), which are then used to calculate the implied equilibrium asset distribution. We solve for the steady state iterating on the value of the discount factor β so that the stationary asset distribution implied by the households' choices satisfies the market clearing condition $\int B_t(j)dj = 0$ at an (annualized) steady state real rate of 2 percent.

For the transition dynamics, we adopt the sequence space Jacobian approach described in Auclert et. al. (2021). This amounts to finding the first-order approximation of the equilibrium responses to arbitrary sequences of anticipated monetary policy and technology shocks, i.e. under perfect foresight over a finite horizon (set to $T = 300$ quarters). Unless otherwise noted, we set the persistence parameters $\rho_r = 0.5$ for monetary policy shocks, and $\rho_a = 0.9$ for technology shocks. Due to certainty equivalence, the resulting dynamics are equivalent to the ones that would be obtained solving the linearized rational expectations model, e.g. as in Reiter (2009) and Ahn et. al. (2018).[13] Also, by construction, the approximate responses to positive and negative aggregate shocks are fully symmetric, and proportional to the size of the shocks. Most importantly, the assumption of perfect-foresight (or certainty equivalence) with respect

[12]For sufficiently small fluctuations in the previous two variables, the fraction of constrained households in equilibrium can be made arbitrarily close to zero. In our simulations, the fraction of constrained consumer is negligible (below 0.1 percent) both in steady state, and in response to aggregate shocks.

[13]See also Boppart, Krusell and Mitman (2018) for a related perfect-foresight sequence-based approach.

to aggregate shocks implies that idiosyncratic income shocks are the only source of individual uncertainty.

3 A HANK Economy without Binding Borrowing Constraints

In this section we consider a version of the HANK model with a natural debt limit, fully illiquid stocks and a profit allocation rule. This framework, which builds on our earlier work (Debortoli and Galí (2024)), helps us identify the specific role of idiosyncratic income shocks as a factor underlying aggregate fluctuations. For brevity we refer to this version of the HANK model as HANK-I.

We assume aggregate operating profits D_t are distributed among households according to the rule

$$\begin{aligned} F_t(j) &= \left[\vartheta + (1-\vartheta)\Xi_t(j)\right] D_t \\ &\equiv \Theta_t(j) D_t \end{aligned}$$

where parameter ϑ defines the fraction of profits that are distributed uniformly across households in the form of dividends, while $1-\vartheta$ is the corresponding fraction which is distributed as "bonuses" in proportion to each household's productivity $\Xi_t(j)$.[14] As discussed below, the setting of ϑ is potentially important in determining how a given change in income is allocated across households in response to a shock, with the consequent implications for aggregate consumption.

The previous assumption on the allocation of profits allows us to write the period budget constraint of a typical household as:

$$C_t(j) + B_t(j) = R_{t-1}B_{t-1}(j) + \Xi_t(j)W_tN_t + \Theta_t(j)D_t$$

The assumption of a natural debt limit, together with $\lim_{C\to 0} U_{c,t} = +\infty$, implies that the borrowing constraint is never binding in equilibrium for any household. In turn, the latter fact implies that the following Euler equation holds for all t and $j \in [0,1]$:

$$1 = \beta R_t \mathbb{E}_t\{(C_{t+1}(j)/C_t(j))^{-\sigma}\} \tag{8}$$

[14] This is the interpretation favored by Kaplan et al (2018), which assume $\vartheta = 0$.

In order to understand how the dynamics of aggregate consumption in the HANK-I model differ from those of RANK it is useful to derive the log-linear approximation to (8). As shown in the Appendix, (8) can be approximated and then aggregated across households to yield the log-linear Euler equation for aggregate consumption

$$c_t = \mathbb{E}_t\{c_{t+1}\} - \frac{1}{\sigma}\widehat{r}_t - \frac{\sigma+1}{2}\widehat{v}_t \tag{9}$$

where $c_t \equiv \log C_t$ with $C_t \equiv \int_0^1 C_t(j)dj$ denoting aggregate consumption, $\widehat{r}_t \equiv \frac{1}{\beta R}(\frac{R_t - R}{R})$, and

$$v_t \equiv \int \frac{C_t(j)}{C_t} v_t(j)dj$$

where $v_t(j) \equiv var_t\{c_{t+1}(j)\}$ is a measure of "individual consumption risk," and $\widehat{v}_t \equiv v_t - \mathbb{E}\{v_t\}$.

The impact of idiosyncratic income risk is captured by the *risk shifter* term v_t in the Euler equation for aggregate consumption. We assume that due to the presence of idiosyncratic income risk, variations in v_t are of the same order of magnitude as variations in aggregate variables resulting from aggregate shocks. This is in contrast with a model with a representative household, in which by construction $var_t\{c_{t+1}\}$ is of second order relative to aggregate consumption and other aggregate variables and is thus generally ignored when approximating the Euler equation for aggregate consumption.

Iterating forward we obtain

$$\begin{aligned} \widehat{c}_t &= -\frac{1}{\sigma}\sum_{k=0}^{\infty}\mathbb{E}_t\{\widehat{r}_{t+k}\} - \frac{\sigma+1}{2}\sum_{k=0}^{\infty}\mathbb{E}_t\{\widehat{v}_{t+k}\} \\ &= -\frac{1}{\sigma(1-\rho_r)}\widehat{r}_t - \frac{\sigma+1}{2}\sum_{k=0}^{\infty}\mathbb{E}_t\{\widehat{v}_{t+k}\} \end{aligned} \tag{10}$$

where $\widehat{c}_t \equiv \log C_t/C$ and where we have used the fact that $\lim_{T\to\infty}\mathbb{E}_t\{c_{t+T}\} = c$, where c is the mean of the ergodic distribution for $c_t(j)$. Note that, in addition to the real interest rate, aggregate consumption depends inversely on current and anticipated values of the risk shifter which capture the extent of precautionary savings. Our simulations of a calibrated version of HANK-I discussed below allow us to assess the importance of that factor in accounting for fluctuations in aggregate consumption.

3.1 A Tractable Counterpart: RANK

As a tractable counterpart to the HANK-I economy analyzed above we propose the standard RANK model. The latter corresponds to a special case of the HANK-I model above with $\Xi_t(j) = 1$ for all t and $j \in [0,1]$. The representative household's optimality condition is given by

$$1 = \beta R_t \mathbb{E}_t\{(C_{t+1}/C_t)^{-\sigma}\}$$

or, in log-linearized form,

$$c_t = \mathbb{E}_t\{c_{t+1}\} - \frac{1}{\sigma}\widehat{r}_t$$

Iterating forward and imposing $\lim_{T\to\infty} \mathbb{E}_t\{c_{t+T}\} = c$ yields:

$$\widehat{c}_t = -\frac{1}{\sigma(1-\rho_r)}\widehat{r}_t$$

Thus, under our assumptions equilibrium output in the RANK economy is a function of the exogenous state variable, $\widehat{r}_t$, and is invariant to the specification of the supply side. Output does not display any endogenous persistence.

The gap between aggregate consumption in the HANK-I and RANK models is thus given by

$$\widehat{c}_t^{HA} - \widehat{c}_t^{RA} = -\frac{\sigma+1}{2}\sum_{k=0}^{\infty}\mathbb{E}_t\{\widehat{v}_{t+k}\} \tag{11}$$

i.e., it depends exclusively on current and expected future values of the risk shifter. Accordingly, the presence of idiosyncratic risk will affect the aggregate response of consumption and output to a given aggregate shock only if the latter has an impact on the risk shifter $\widehat{v}_t$. The latter is an endogenous variable which cannot be solved in closed form, so we need to (numerically) solve for the equilibrium of the HANK-I model to evaluate quantitatively the size of that gap. In doing so, and in addition to the calibrated values introduces above, we assume $\vartheta = 0.5$, implying that half of aggregate profits are distributed to households in proportion to their idiosyncratic labor productivity.

Panel (a) of Figure 1 shows the impulse response of equilibrium output to monetary policy and technology shocks in HANK-I (red line with circles) and RANK (blue line with crosses). In the case of a monetary policy shock we consider a 25 basis point reduction in the real rate on impact (which corresponds to a 1 percentage point in annualized terms). The presence of idiosyncratic risk in HANK-I leads to an amplification of the output effects of the shock: the effects are stronger on impact, and more persistent. The difference is, however, quantitatively very small.

The previous assessment is confirmed by Panel (b) of Figure 1, which displays the simulated time series for (log) output generated by HANK-I and RANK in response to a sequence of monetary policy surprises drawn from a normal distribution. In a way consistent with the impulse responses, we see that the volatility of output under HANK-I is slightly larger (by a 1.22 factor), but the correlation between the two is very high (0.97), pointing to a limited impact of the additional dynamics resulting from changes in the wealth distribution, an endogenous state variable present in HANK-I but not in RANK.

In Debortoli and Galí (2024), we sought to understand the reasons behind the small gap between the output responses in the two models. Given (11), that finding must ultimately be associated with a small response of the risk shifter to a shock. The basic intuition for that small response can be obtained from the following approximation derived in that paper:

$$v_t(j) \simeq \psi_t(j)^2 \sigma_\zeta^2$$

where $\psi_t(j)$ is the elasticity of individual consumption with respect to the idiosyncratic shock. Hence

$$v_t \simeq \sigma_\zeta^2 \int \frac{C_t(j)}{C_t} \psi_t(j)^2 dj \tag{12}$$

where $\psi_t(j)$ is household j's elasticity of consumption with respect to the innovation in the idiosyncratic income process and σ_ζ^2 is the variance of the latter. In response to an aggregate shock the following approximation holds:

$$\frac{dv_{t+k}}{d\varepsilon_t} \simeq \sigma_\zeta^2 \int \frac{C_t(j)}{C_t} \frac{d\psi_{t+k}(j)^2}{d\varepsilon_t} dj$$

As shown in Figure 2, drawn from Debortoli and Galí (2024), $\psi_t(j)^2$ is *decreasing and (strongly) convex in the level of consumption*, capturing the fact that the consumption of households with less liquid wealth and closer to the natural debt limit is more responsive to shocks that change that wealth (i.e. their MPCs are higher). Accordingly, and in response to shocks that shift the wealth distribution in either direction, $\frac{d\psi_{t+k}(j)^2}{d\varepsilon_t}$ is quantitatively significant only for poor households. Since the weight of those households in aggregate consumption is small, the dynamic response of the aggregate risk shifter becomes muted, thus accounting for the small gap between $\widehat{c}_t^{HA}$ and $\widehat{c}_t^{RA}$.

We can apply the previous reasoning to understand the impact of idiosyncratic risk on the responses to a monetary policy shock. In particular, note that an expansionary monetary policy shock has two effects that tend to offset each other, above and beyond the intertemporal substitution in response to a change in the real interest rate that is already captured by the RANK model. First, the interest rate reduction implies a redistribution from (rich) lenders to (poor) borrowers, which reduces wealth dispersion. We refer to this as the *interest rate exposure* channel. On the other hand, the monetary expansion raises *average* labor income but lowers aggregate dividends. For households with productivity above a certain threshold, total (non-interest) income increases, but this is not the case for low productivity households, which experience a decline in that income. To see this formally, define a household's income (excluding interest) as:

$$Y_t(j) \equiv \Xi_t(j) W_t N_t + \Theta_t(j) D_t$$

which we can rewrite in terms of aggregate output and the markup as follows:

$$Y_t(j) = \left[\Xi_t(j) \frac{\mathcal{M}}{\mathcal{M}_t} + \Theta_t(j) \left(1 - \frac{\mathcal{M}}{\mathcal{M}_t} \right) \right] Y_t$$

Differentiating the previous expression and using (6) we obtain:

$$dY_t(j)/Y = \left[\Xi_t(j) - \vartheta(1 - \Xi_t(j))(\sigma + \varphi) \right] dy_t + \vartheta(1 - \Xi_t(j))(1 + \varphi) da_t \tag{13}$$

Accordingly, in response to an expansionary monetary policy shock which raises aggregate output, $dY_t(j) > 0$ if and only if $\Xi_t(j) > \frac{\vartheta(\sigma+\varphi)}{1+\vartheta(\sigma+\varphi)}$. Thus, through this mechanism and to the extent that $\theta > 0$, income is redistributed from poor/low productivity households to rich/high productivity households. This is what we call the *income distribution* channel.[15]

Thus, and as long as $\vartheta > 0$, the interest rate exposure and income distribution channels work in opposite directions and thus tend to neutralize the impact of the shock on the wealth distribution and, as a result, on the difference in the response of aggregate consumption and output between RANK and HANK-I. In addition, whatever net redistribution there is it affects

[15] Note that when $\vartheta = 0$ we have $Y_t(j) = \Xi_t(j) Y_t$; hence, aggregate shocks have no redistributive effects and that channel is not operative.

mostly the consumption risk of poor households, so the impact on the aggregate risk shifter and aggregate consumption are small. This is consistent with the small difference uncovered in the impulses responses shown in Figure 1.[16]

In order to assess the importance of each channel, the green dashed line in the Figure shows the response of aggregate output when setting $\vartheta = 0$, i.e. when we turn off the income distribution channel. The tiny gap between that response and the one implied by our baseline calibration suggests a very small role for that channel in shaping the aggregate response to the shock. In that case, the interest rate exposure channel redistributes resources from the rich to the poor, thus reducing the overall consumption risk and amplifying the response of output relative to RANK, albeit in a small amount given the low weight of poor households' consumption in the aggregate.

Let us now turn to the effects of technology shocks. The second panel in Figure 1 shows that a one percent positive technology shock has no effect on output in the RANK model since the central bank keeps the real rate constant, thus preventing aggregate demand from increasing. By contrast, the output effect of the same shock in HANK-I is positive, albeit very small. The intuition for the increase in output in HANK-I is as follows. *At any given initial level of output*, the shock raises dividends while reducing labor income by the same amount. As (13) makes clear that adjustment does not affect everyone equally: it raises the income of households with productivity below the mean ($\Xi_t(j) < 1$), for which dividends account for a larger share of their income, while lowering it for the remaining households. Thus, the shock implies a redistribution of income from rich to poor households. The reduction in consumption risk for the poor more than offsets the small increase in that risk experienced by the rich, implying an overall decline in precautionary savings, with the consequent expansion in aggregate demand and output captured in Figure 1. Again, the effect is quantitatively small because the reduction in precautionary savings affects mostly households with a low consumption share to begin with. The response in the counterfactual case with $\vartheta = 0$ is represented by the dashed green line, which overlaps perfectly with the zero response associated with RANK, for in that case there is no income redistribution and hence no change in consumption.

[16]The absolute difference between the two IRFs increases when the autoregressive coefficient is calibrated to a higher value —see the discussion by Rognlie (2024) contained in the present volume. However, the difference in relative terms remains an order of magnitude smaller than the IRF itself.

Next we assess the role of idiosyncratic income risk as a source of endogenous persistence. As is well known, in the basic RANK model there are no state variables beyond the exogenous variables themselves. This is not the case in HANK models generally, in which, as discussed above, the distribution of wealth is itself a state variable. In order to assess to what extent the dynamic response of that distribution is capable of generating endogenous persistence (i.e. persistence beyond that of the exogenous driving forces), we report in Figure 3 the dynamic response of output to *purely transitory* monetary policy and technology shocks in the HANK-I model, next to the corresponding responses in RANK. The Figure makes clear that while the presence of idiosyncratic risk generates persistence in the output response, the effect beyond the initial period is quantitatively small. In other words, the endogenous response of the wealth distribution to an aggregate shock has quantitatively small implications for aggregate output.[17]

We conclude from the previous exercises that RANK provides a good approximation to the response of aggregate output to both monetary and technology shocks implied by the HANK-I model, i.e. a version of the HANK model without binding borrowing constraints. Next we study whether the introduction of occasionally binding borrowing constraints in the HANK model leads to a different conclusion.

4 A HANK Economy with Binding Borrowing Constraints

In the previous section we analyzed a version of the HANK model without binding borrowing constraints, in which the consumption Euler equation held for all households at all times. This allowed us to derive an approximate aggregate consumption Euler equation and to insulate the role of idiosyncratic income risk. By contrast, in the present section, and following much of the HANK literature, we assume a borrowing limit tighter than the natural debt limit. As a result, the borrowing constraint is binding in equilibrium for a non-negligible fraction of households, which we label as hand-to-mouth. That fraction evolves endogenously, with its value in period t denoted by λ_t^H.

[17]On the other hand, tiny quantitative differences that are highly persistent may end up having significant cumulative effects, often described by means of a cumulative multiplier statistic. That statistic may capture differences that grow very fast *in percent terms* with the horizon if the reference response in the denominator gets close to zero.

More specifically, the borrowing limit is now assumed to be given by

$$\underline{B} = -\psi Y$$

The remaining assumptions are the same as in HANK-I. We refer to the resulting model as HANK-II.

Households who are unconstrained in period t, satisfy the Euler equation

$$1 = \beta R_t \mathbb{E}_t\{(C_{t+1}(j)/C_t(j))^{-\sigma}\} \tag{14}$$

which thus holds for all $j \in \mathcal{U}_t \equiv \{j \in [0,1] : R_t B_t(j) > -\psi Y\}$.

As shown in the appendix, one can approximate and then average the resulting equation across unconstrained households to yield the log-linearized Euler equation:

$$\widehat{c}_t^U = \mathbb{E}_t\{\widehat{c}_{t+1}^U\} - \frac{1}{\sigma}\widehat{r}_t - \frac{\sigma+1}{2}\widehat{v}_t^U - \widehat{h}_t^U \tag{15}$$

where $\widehat{c}_t^U \equiv \log C_t^U / C^U$ with $C_t^U = \frac{1}{1-\lambda_t^H}\int_{j\in\mathcal{U}_t} C_t(j)dj$ and $v_t^U \equiv \frac{1}{1-\lambda_t^H}\int_{j\in\mathcal{U}_t}\frac{C_t(j)}{C_t^U}v_t(j)dj$ measuring respectively average consumption and average consumption risk (the latter weighed by relative consumption) of households who are unconstrained in period t. In addition we have $h_t^U \equiv \mathbb{E}_t\{(C_{t+1}^U - C_{t+1|t}^U)/C_t^U\}$ where $C_{t+1|t}^U \equiv \frac{1}{1-\lambda_t^H}\int_{j\in\mathcal{U}_t} C_{t+1}(j)dj$ is the average consumption in $t+1$ of households who were unconstrained in period t. Note that h_t^U emerges as a result of changes in the composition of subset $\mathcal{U}_t$, associated with the fact that some households that are unconstrained at t become constrained at $t+1$, and viceversa, so that in general we have $C_{t+1}^U \neq C_{t+1|t}^U$. We refer to this additional term in the Euler equation as the *composition shifter*. The presence of both the risk shifter $\widehat{v}_t^U$ and the composition shifter $\widehat{h}_t^U$ is tied to the existence of idiosyncratic income risk. In the absence of the latter, $\widehat{v}_t^U$ would be of second order and $\widehat{h}_t^U$ would be zero.

Iterating (15) forward we can write the gap between average consumption of the unconstrained in HANK-II and that in the RANK model as

$$\widehat{c}_t^U - \widehat{c}_t^{RA} = -\frac{\sigma+1}{2}\sum_{k=0}^{\infty}\mathbb{E}_t\{\widehat{v}_{t+k}^U\} - \sum_{k=0}^{\infty}\mathbb{E}_t\{\widehat{h}_{t+k}^U\} \tag{16}$$

whose quantitative significance we seek to evaluate below.

For the remaining households, for which the borrowing constraint is binding, consumption is given by

$$C_t(j) = \Xi_t(j) W_t N_t + \Theta_t(j) D_t + R_{t-1} B_{t-1}(j) + \frac{\psi Y}{R_t} \tag{17}$$

which holds for all $j \in \mathcal{H}_t \equiv \{j \in [0,1] : R_t B_t(j) = -\psi Y\}$. Note that, at the margin, any change in household j's current income $\Xi_t(j) W_t N_t + \Theta_t(j) D_t$ while constrained leads to a one-for-one change in consumption (i.e. the MPC is one). Thus, and following the literature, we refer to that subset of households as *hand-to-mouth*.

Averaging (17) over $j \in \mathcal{H}_t$ yields:

$$\begin{aligned} C_t^H &= \frac{1}{\lambda_t^H} \int_{j \in \mathcal{H}_t} C_t(j) dj \\ &= \Xi_t^H W_t N_t + \Theta_t^H D_t - \psi Y \left(\Omega_{t-1}^H - \frac{1}{R_t} \right) \\ &= \left[\Xi_t^H \frac{\mathcal{M}}{\mathcal{M}_t} + \Theta_t^H \left(1 - \frac{\mathcal{M}}{\mathcal{M}_t} \right) \right] Y_t - \psi Y \ (\widehat{R}_t + \Omega_{t-1}^H - 1) \end{aligned} \tag{18}$$

where $\widehat{R}_t \equiv \frac{R_t - 1}{R_t}$, $\Xi_t^H \equiv \frac{1}{\lambda_t^H} \int_{j \in \mathcal{H}_t} \Xi_t(j) dj$ and $\Theta_t^H \equiv \frac{1}{\lambda_t^H} \int_{j \in \mathcal{H}_t} \Theta_t(j) dj$ are, respectively, the average productivity of and the average fraction of aggregate profits accruing to hand-to-mouth households in period t, while $\Omega_{t-1}^H \equiv -\frac{1}{\lambda_t^H} \int_{j \in \mathcal{H}_t} \frac{R_{t-1} B_{t-1}(j)}{\psi Y} dj$ denotes the average debt maintained in period $t-1$ by period t hand-to-mouth households, expressed as a ratio to the debt limit ψY.

Equation (18) reveals explicitly the role of the interest rate exposure channel and the income distribution channel in determining the consumption of constrained households. First, note that to the extent that most currently constrained households were either already constrained or close to being constrained in the previous period we would expect Ω_{t-1}^H to be positive and close to 1. As a result, $\partial C_t^H / \partial R_t < 0$, revealing the presence of the interest rate exposure channel. Moreover, and given that *average* productivity among hand-to-mouth households will (naturally) be below the mean (i.e., $\Xi_t^H < 1$), we will have $\Theta_t^H > \Xi_t^H$ (as long as $\vartheta > 0$). Accordingly, an increase in the average markup will, ceteris paribus, raise hand-to-mouth consumption, i.e. $\partial C_t^H / \partial \mathcal{M}_t > 0$. The reason is that it implies an increase in dividends and a reduction in labor income, and thus a redistribution of income towards low productivity, constrained households (for whom dividend income is relatively more important).

Aggregate consumption can thus be written as:

$$\begin{aligned} C_t &= \lambda_t^H C_t^H + (1-\lambda_t^H) C_t^U \\ &= \Phi_t Y_t - \lambda_t^H \psi Y \ (\widehat{R}_t + \Omega_{t-1}^H - 1) + (1-\lambda_t^H) C_t^U \end{aligned} \tag{19}$$

where C_t^U satisfies (15) and $\Phi_t \equiv \lambda_t^H \left[\Xi_t^H \frac{\mathcal{M}}{\mathcal{M}_t} + \Theta_t^H \left(1 - \frac{\mathcal{M}}{\mathcal{M}_t} \right) \right]$ can be interpreted as a time-varying "slope of the Keynesian cross."

What are the aggregate consequences of introducing a borrowing constraint, by setting a debt limit tighter than the natural debt limit, so that such constraint becomes binding for a non-negligible fraction of households every period? We address this question by computing the impulse responses to monetary and technology shocks in a calibrated version of HANK-II with $\psi = 2$. The implied borrowing limit is consistent with 30% of households being constrained in the stochastic steady state, consistently with Kaplan et al. (2014). The settings for the remaining parameters are left unchanged. Figure 4 displays the resulting impulse responses, together with those generated by HANK-I.

The first panel of Figure 4 shows that the response of output to an expansionary monetary policy shock. Perhaps surprisingly, we see that the presence of a binding borrowing constraint does not amplify significantly the response of output, and dampens it after a certain horizon. The previous finding seems at odds with the fact that in HANK-II a significant fraction of households behave in a hand-to-mouth fashion, thus generating a large direct multiplier effect captured by coefficient Φ_t –the slope of the Keynesian cross in (19). At least two factors account for this result. First, the *direct* effects of interest rate changes, working through intertemporal substitution by unconstrained households are now smaller, since the latter account only for a fraction $1 - \lambda_t^H$ of all households. Secondly, because of the tighter borrowing constraint the distribution of wealth across households in HANK-II is less dispersed than in HANK-I. As a result, the interest rate exposure channel is more muted.

Things are substantially different in response to a technology shock, since in that case the absence of a monetary policy response neutralizes the interest rate exposure channel, as well as the direct effect working through intertemporal substitution by unconstrained households. As a result, the higher multiplier associated with the presence of hand-to-mouth households ends up becoming the key factor, leading to the amplified response to the positive technology shock relative to HANK-I shown on the right panel of Figure 4.

To sum up, the introduction of occasionally binding borrowing constraints, with the consequent emergence of hand-to-mouth households, does not necessarily amplify the effects of shocks. Whether this is the case or not depends on the nature of the shock as well as on the strength of potential offsetting effects (including an eventual endogenous response of monetary policy, not modeled here). Next we look for a tractable framework that can account for all these properties of HANK-II.

4.1 A Tractable Counterpart to HANK-II

4.1.1 The Standard TANK Model (TANK-I)

The key difference between HANK-I and HANK-II is the presence of a fraction of households that do not satisfy the Euler equation and instead behave in a hand-to-mouth fashion, consuming their current income (net of debt service payments). The "standard" TANK model (Galí et al. (2007), Bilbiie (2008)), which assumes ex-ante heterogeneity with two types of households (unconstrained and hand-to-mouth), may naturally be viewed –and has often been portrayed in the literature– as a tractable framework to approximate the equilibrium dynamics of a HANK model with binding borrowing constraints like HANK-II. Standard versions of the TANK model, however, fail to capture some of the mechanisms at work in HANK-II, as we show next.

Consider a standard version of the TANK model, where a constant fraction λ^H of households hold no financial assets, and just consume their current labor income each period, i.e. $C_t^H = W_t N_t$. The remaining fraction $1-\lambda^H$ are unconstrained. Most importantly, households of a given type are identical, both ex-ante and ex-post; in particular they experience no idiosyncratic income shocks. Henceforth, we refer to that version of the TANK model as TANK-I (in order to distinguish it from alternative versions considered below).

Accordingly, in the TANK-I economy aggregate consumption is given by:

$$
\begin{aligned}
C_t &= \lambda^H W_t N_t + (1-\lambda^H) C_t^U \\
&= \lambda^H \frac{\mathcal{M}}{\mathcal{M}_t} Y_t + (1-\lambda^H) C_t^U
\end{aligned}
\tag{20}
$$

where C_t^U satisfies

$$\hat{c}_t^U = \mathbb{E}_t\{\hat{c}_{t+1}^U\} - \frac{1}{\sigma}\hat{r}_t \tag{21}$$

A comparison of (19) and (20) uncovers several differences between HANK-II and TANK-I. Two of those differences are linked to standard features of TANK models: (i) a *constant* fraction of hand-to-mouth households and (ii) the absence of (first order effects of) precautionary savings, as reflected in the missing shifter terms in (21), due to the absence of idiosyncratic risk. Thus, to the extent that variations in the fraction of hand-to-mouth households and in precautionary savings are significant factors underlying fluctuations in aggregate consumption, TANK models will have little chance to approximate the equilibrium properties and mechanisms in HANK-II.

Other shortcomings, however, are specific to the standard version of the TANK model described above (i.e. TANK-I), and may be amenable to modification. In particular, it is clear that (20) fails to capture the interest rate exposure channel revealed by (19), since hand-to-mouth households are not indebted in TANK-I. Secondly, (20) points to a negative relation between aggregate consumption and the markup, for a given level of output, due to the negative effect of a higher markups on labor income, a variable tightly connected to aggregate consumption due to the presence of hand-to-mouth households. Thus, the sign of the income distribution channel at work in HANK-II (and captured in (19)) is reversed in TANK-I.

Given the previous differences, it may not be surprising that TANK-I fails to approximate well the aggregate properties of HANK-II. This is illustrated in Figure 5a which displays the response of aggregate output to monetary policy and technology shocks in both models (red and green lines, respectively). Note that the only new parameter in TANK-I relative to RANK is λ^H, which we set to 0.30, the mean value of λ_t^H in HANK-II.

As Figure 5a makes clear, the output response to an expansionary monetary policy shock is *highly* amplified in TANK-I (green line with crosses) relative to HANK-II (red line with circles), almost trebling the effect on impact. In the case of technology shocks the difference is even starker since the sign of the output response in TANK-I is reversed relative to HANK-II, due to the fall in labor income. Given the previous discussion and findings, one can hardly view TANK-I as providing a reasonable approximation to HANK-II.

4.1.2 A Modified TANK Model: TANK-II

Next we propose a simple modification of the TANK model that has a better chance to approximate well the aggregate predictions of HANK-II. In our modified model, which we refer to as TANK-II, we make three assumptions that seek to mirror some features of HANK-II, in a way not captured by the standard TANK model. First, we assume that hand-to-mouth households are permanently against the borrowing constraint introduced in HANK-II, i.e. $R_t B_t^H = -\psi Y$ for all t. Secondly, we assume that the productivity of hand-to-mouth households is given by $\Xi^H < 1$, and hence is lower than that of unconstrained households.[18]Finally, we assume that dividends are allocated to all households (including the hand-to-mouth) according to the same rule assumed in HANK-II, thus implying $D_t^H = \Theta^H D_t$, where $\Theta^H \equiv \vartheta + (1-\vartheta)\Xi^H > \Xi^H$.

Under the previous assumptions, consumption of hand-to-mouth households in TANK-II is given by

$$C_t^H = \Xi^H W_t N_t + \Theta^H D_t - \psi Y \ \widehat{R}_t$$

implying the following expression for aggregate consumption:

$$C_t = \lambda^H \left[\Xi^H \frac{\mathcal{M}}{\mathcal{M}_t} + \Theta^H \left(1 - \frac{\mathcal{M}}{\mathcal{M}_t} \right) \right] Y_t - \lambda^H \psi Y \ \widehat{R}_t + (1-\lambda^H) C_t^U \tag{22}$$

where C_t^U satisfies (21). Note that, in contrast with (20), consumption equation (22) captures, at least qualitatively, both the interest rate exposure and income distribution channels at work in HANK-II, as revealed by a comparison with the expression for aggregate consumption in the latter in (19), which we reproduce here for convenience:

$$C_t = \lambda_t^H \left[\Xi_t^H \frac{\mathcal{M}}{\mathcal{M}_t} + \Theta_t^H \left(1 - \frac{\mathcal{M}}{\mathcal{M}_t} \right) \right] Y_t - \lambda_t^H \psi Y \ (\widehat{R}_t + \Omega_{t-1}^H - 1) + (1-\lambda_t^H) C_t^U$$

Note that TANK-II will be a good approximation to HANK-II if $\lambda_t^H \simeq \lambda^H$, $\Omega_{t-1}^H \simeq 1$, $\Xi_t^H \simeq \Xi^H$ and $v_t^U \simeq v^U$ for all t. Next we turn to a quantitative assessment of the goodness of that approximation.

As in previous sections, we assess the goodness of the approximation by comparing the impulse responses of output in the HANK-II and TANK-II models to monetary and technology shocks. In order to generate the impulse responses for TANK-II we set $\lambda^H = 0.30$, $\Xi^H = 0.56$

[18]Given our normalization, $\lambda^H \Xi^H + (1-\lambda^H)\Xi^U = 1$.

and $\Theta^H = 0.78$, which match the steady state values of their (time-varying) counterparts in HANK-II.[19] We also set $\psi = 2$, the value assumed in HANK-II.

As shown in Figure 5a, and in contrast with the predictions of TANK-I discussed above, the output response to a monetary policy shock in TANK-II (shown in blue) matches closely that of HANK-II. For the technology shock the match is also reasonably good, especially in comparison with TANK-I, which even fails to get the sign right. The reason for the difference is the presence in TANK-II of the interest rate exposure and income distribution channels. As discussed above, those channels play an important role in shaping the aggregate properties of HANK-II, but are absent from TANK-I.

Figure 5b plots the time series for (log) output generated by TANK-II and HANK-II in response to monetary policy (top panel) and technology shocks (bottom panel). In the case of monetary policy shocks, the gap between the two time series is hardly noticeable, with the ratio of volatilities equal to 1.10 and implied correlations being very close to unity. In the case of technology shocks the pattern is more different, with a larger ratio of volatilities (1.97), though it is not clear that the latter is much meaningful since the absolute impact of the shock is tiny in the two cases, as Figure 5a makes clear. The correlation is still very high (0.998), pointing to an insignificant role of changes in the wealth distribution as an additional state variable in HANK-II.

Next we analyze whether our modified TANK framework is able to capture the impact of a tightening of borrowing constraints predicted by HANK. We illustrate this in Figure 6, which reports the impulse responses generated by HANK-II and TANK-II when we tighten the borrowing limit by setting $\psi = 0.8$, so that the fraction of constrained households in steady state increases from 0.30 to 0.50. The responses are expressed as a gap relative to their counterparts in the absence of borrowing constraints (i.e. HANK-I and RANK, respectively), thus allowing us to isolate the role of the tightening of the budget constraint, independently of initial differences between HANK-I and RANK.

In the case of monetary policy shocks the tightening of the borrowing limit shifts down the

[19] In his discussion, Wieland (2024) argues that there is little advantage to using a tractable framewok if one has to solve the HANK model in order to calibrate its parameters. Note, however, that one could instead rely on alternative strategies to calibrate those parameters, e.g. using direct independent evidence on each parameter, or matching the empirical impulse responses.

impulse response predicted by HANK-II, i.e. it dampens the impact of monetary policy. A similar downward shift is observed in the case of TANK-II. With regard to technology shocks, we see that the TANK-II model can also capture well the amplification of the output response predicted by the HANK-II model under a tighter borrowing constraint. We conclude that the TANK-II model can also capture reasonably well the impact of a change in the environment like the tightening of a borrowing constraint.

In order to further understand the extent to which the TANK-II model provides a good approximation to HANK-II and its underlying mechanisms we consider the following decomposition of aggregate consumption

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

Accordingly, the response of aggregate consumption at different horizons to a shock in period t can be decomposed as follows:

$$\frac{dC_{t+k}}{d\varepsilon_t} = \lambda^H \frac{dC_{t+k}^H}{d\varepsilon_t} + (1-\lambda^H)\frac{dC_{t+k}^{RA}}{d\varepsilon_t} + \frac{d\Upsilon_{t+k}}{d\varepsilon_t} \tag{23}$$

where C_{t+k}^{RA} denotes consumption in the corresponding RANK model.[20] The third term is a residual component resulting from variations in λ_t^H as well as changes in the risk and composition shifters caused by the shock. Note that this residual component is absent in TANK-II, since the latter assumes subsets of unconstrained and hand-to-mouth households that are time invariant in size and composition, and displays no precautionary savings.

Figure 7a displays the decomposition of the impulse responses to a monetary policy shock into the components shown in (23) for both the HANK-II and TANK-II models. The figure highlights the similarity in that decomposition across the two models, suggesting that not only the TANK-II model is successful in approximating the *aggregate* properties of HANK-II but also in capturing the underlying mechanisms. Fig 7b shows the corresponding results for the

[20] Formally, note that

$$\begin{aligned}
\frac{d\Upsilon_{t+k}}{d\varepsilon_t} &= (1-\lambda^H)\left(\frac{dC_{t+k}^U}{d\varepsilon_t} - \frac{dC_{t+k}^{RA}}{d\varepsilon_t}\right) + \frac{d\lambda_{t+k}^H}{d\varepsilon_t}\left(C^H - C^U\right) \\
&= -(1-\lambda^H)\left(\frac{\sigma+1}{2}\sum_{l=k}^{\infty}\frac{dv_{t+l}^U}{d\varepsilon_t} + \sum_{l=k}^{\infty}\frac{dh_{t+l}^U}{d\varepsilon_t}\right) + \frac{d\lambda_{t+k}^H}{d\varepsilon_t}\left(C^H - C^U\right)
\end{aligned}$$

technology shock. In this case, and given the small output response to the shock, the residual component (shown as "other" in the Figure) is relatively more important, even though still small in absolute terms. That component cannot be captured by the TANK-II model. The latter captures well, however, the size and pattern of the consumption response of hand-to-mouth households.

4.1.3 The Effects of Fiscal Policy

As a complementary exercise, we analyze how heterogeneity may affect the transmission of fiscal shocks. To that end, we modify the models considered above by introducing an exogenous government spending shock (dG_t), which is assumed to be financed through lump-sum taxes — which for simplicity are set to be identical for all households- or by issuing debt (B_t^g), according to the rule $dB_t^g = \rho_b(dB_{t-1}^g + dG_t)$, with $\rho_b \in [0, 1)$.

Similarly to Auclert et. al. (2023), we consider two alternative scenarios: (i) a balanced-budget rule (i.e. $\rho_b = 0$) and (ii) deficit-financed spending (with $\rho_b = 0.9$). In both cases, we consider the impulse response to a government spending shock equal to 1% of steady-state GDP and with persistence $\rho_g = 0.8$, and starting from the same steady-state considered above —i.e. with $B = G = 0$.

Results are summarized in Figure 8. As is well known, and due to Ricardian equivalence, in a RANK economy (black dashed-line) the effects of a government spending shock do not depend on whether it is financed with current taxes (left panel) or debt (right panel). Also, under the maintained assumption that the central bank keeps a constant real interest rate, an increase in government spending translate one-to-one into an increase in output —i.e. a multiplier equal to one.[21] Things are different in the HANK-II economy (red line), where we find that the effects of government spending shocks are dampened (relative to RANK) under a balanced budget rule, but amplified (at least on impact) under deficit financing.

Interestingly, the same results are obtained in the TANK-II model (blue line), where those patterns can be easily rationalized. Under a balanced budget rule, the government needs to raise taxes, which other things equal lead to a decline in consumption of financially constrained households, and thus to a reduction in aggregate consumption.[22]. Such a crowding-out effect

[21] See Woodford (2011)

[22] More precisely, since hand-to-mouth households have low productivity, the increase in taxes is larger than

implies that the government spending multiplier is lower than one. Conversely when government spending is deficit-financed, constrained households do not internalize that future taxes will increase, and their consumption increases due to the increase in labor income. Thus, aggregate consumption increases, and the government spending multiplier is larger than one.[23]

4.1.4 Caveats

Two potential aspects of the HANK-II model analyzed above can be criticized on empirical grounds. First, the model assumes an extreme dichotomy with regard to the extent of assets' liquidity: fully liquid bonds, fully illiquid stocks, with profits allocated according to an-hoc rule. In actual economies, most assets can be bought and sold, even though for some assets such transactions may be subject to significant costs. That possibility, assumed away in the HANK models considered above, opens the door for resorting to the sale of less liquid assets for the purposes of consumption smoothing once the borrowing limit is attained. Secondly, and relatedly, the micro evidence points to the need to distinguish between poor hand-to-mouth and wealthy hand-to-mouth households, based on whether they have or do not have some illiquid (or less liquid) assets that they can deplete once they have attained their borrowing limit (e.g. Kaplan et al. (2014)).

Next we analyze a version of the HANK model that seeks to overcome those limitations, and propose a tractable counterpart to it.

5 A HANK Economy with Binding Borrowing Constraints and Portfolio Choice (HANK-III)

In this section we take as a starting point the HANK-II model developed above and modify it to allow for endogenous changes in individual equity holdings. That possibility gives an extra margin to equity holders through which they can smooth consumption, even when they have reached their borrowing limit and cannot use bonds for that purpose. Our assumption that such equity changes are subject to a portfolio adjustment cost which limits the extent to which they are effectively used for consumption smoothing purposes. As a result households whose

the increase in labor income

[23] A similar result was obtained by Galí, López-Salido and Vallés (2007).

borrowing constraint is binding will still display high MPCs even when their equity holdings are positive. Following Kaplan et al. (2014) we refer to those households as the "wealthy hand-to-mouth".

Our model builds on the formalism proposed in Kaplan et al. (2018). In particular, we assume that households are not allowed to hold firms' stocks directly; instead they hold bonds and/or equity issued by financial intermediaries, who in turn invest the proceeds into the available assets (firms' stocks in our case, as we abstract from capital and government debt). We refer to this version of HANK as HANK-III.

Bonds, denoted by $B_t(j)$, can be adjusted at no cost and yield a gross real return R_t. Negative values of $B_t(j)$ can be interpreted as loans from financial intermediaries to household j. As before, we assume a borrowing constraint given by $R_t B_t(j) \geq -\psi Y$. On the other hand, household j's equity position, denoted by $E_t(j)$, yields a stochastic gross return R^e_{t+1} (defined below) and may be adjusted at a cost given by $\chi_t(j)$ given by

$$\chi_t(j) \equiv \frac{\chi_1}{\chi_2} \left| \frac{R^e_t E_{t-1}(j) - E_t(j)}{R^e_t E_{t-1}(j) + \chi_0} \right|^{\chi_2} \left(R^e_t E_{t-1}(j) + \chi_0 \right)$$

with $\chi_0 \geq 0$, $\chi_1 \geq 0$ and $\chi_2 > 1$. Note that $R^e_t E_{t-1}(j) - E_t(j)$ can be interpreted as net withdrawals from the equity account (net addition to that account, if negative). Note also that a passive strategy consisting of reinvestment of initial balances plus returns is costless.[24] Finally, and most importantly, we assume that individual equity holdings cannot be negative, i.e. we impose $E_t(j) \geq 0$ for all t and $j \in [0,1]$ ("short-sale constraint").

The period budget constraint of household j can thus be written as

$$C_t(j) + B_t(j) \leq R_{t-1} B_{t-1}(j) + \Xi_t(j) W_t N_t + F_t(j) \tag{24}$$

where $F_t(j) \equiv R^e_t E_{t-1}(j) - E_t(j) - \chi_t(j)$ denotes additions to cash-on-hand linked to equity holdings.

When the short-sale constraint is not binding (i.e. when $E_t(j) > 0$) there is an additional optimality condition that the household must satisfy, given by:

$$1 = \beta \mathbb{E}_t \left\{ \left(\frac{C_{t+1}(j)}{C_t(j)} \right)^{-\sigma} \left(\frac{R^e_{t+1} - \frac{\partial \chi_{t+1}(j)}{\partial E_t(j)}}{1 + \frac{\partial \chi_t(j)}{\partial E_t(j)}} \right) \right\} \tag{25}$$

[24]The term χ_0 in the denominator is added in order to avoid infinite adjustment costs when $E_{t-1}(j) = 0$.

At each point in time, we can partition households into three groups: the unconstrained, the "wealthy hand-to-mouth," and the "poor hand-to-mouth." Unconstrained households satisfy $R_t B_t(j) > -\psi Y$ and $E_t(j) > 0$. For the wealthy hand-to-mouth it is also the case that $E_t(j) > 0$, but their borrowing constraint is binding, i.e. $R_t B_t(j) = -\psi Y$. Finally, both constraints are binding in the case of the poor hand-to-mouth, i.e. $R_t B_t(j) = -\psi Y$ and $E_t(j) = 0$.

A representative financial intermediary takes bonds and equity from households and invests them into firms' stocks, which are traded at a price Q_t. It faces an intermediation cost ωB_t —which can be viewed as the cost of liquidity transformation— incurred at maturity. The financial intermediary solves the following problem:

$$\begin{aligned} &\max_{S_t, B_t} \; \mathbb{E}_t \left\{ \Lambda_{t,t+1} [(Q_{t+1} + D_{t+1}) \, S_t - (R_t + \omega) \, B_t] \right\} \\ s.t. \;\; &Q_t S_t = E_t + B_t \end{aligned}$$

where S_t denotes the quantity of firms' stocks, $\Lambda_{t,t+1}$ is the relevant stochastic discount factor, and where Q_t, R_t, $E_t \equiv \int_0^1 E_t(j)\, dj$ and the distribution of D_{t+1} are taken as given. The solution to the problem above implies the following no-arbitrage condition:

$$\mathbb{E}_t \left\{ \Lambda_{t,t+1} \left[\left(\frac{Q_{t+1} + D_{t+1}}{Q_t} \right) - (R_t + \omega) \right] \right\} = 0$$

In equilibrium $S_t = 1$, and the ex-post return R^e_{t+1} on equity is given by

$$R^e_{t+1} = \frac{Q_{t+1} + D_{t+1} - (R_t + \omega)\, B_t}{E_t}$$

and, hence,

$$\mathbb{E}_t \left\{ \Lambda_{t,t+1} \left[R^e_{t+1} - (R_t + \omega) \right] \right\} = 0$$

implying the steady state relation

$$R^e = R + \omega$$

In our quantitative exercise we set $\omega = 0.002$, which implies an annualized equity premium of 0.8 percent. Given the steady state real interest rate, this is consistent with a steady state return on equity $R^e = 1.0071$ (as in Kaplan et al. (2018) and Auclert et al. (2023)) and a value of total assets equal to 3.20 times annual GDP. For the portfolio adjustment cost function, we set the curvature parameter $\chi_2 = 2$ (i.e. a quadratic function), as well as $\chi_0 = 2.55$ and $\chi_1 = 9.60$ so that the fraction of constrained households (for which $R_t B_t(j) = -\psi Y$) is 30 percent in steady state, of which 25 percent hold equity (the wealthy hand-to-mouth) and 5 percent hold no equity (the poor hand-to-mouth). This calibration also implies that the total amount of liquid and illiquid assets equal 0.25 and 2.9 times annual GDP, which are close to the values reported in Kaplan et. al (2018). The remaining parameters are set to the same values shown in Table 1.

Figure 9 displays the impulse responses of output to monetary policy and technology shocks generated by HANK-III (red solid lines with diamonds), as well as HANK-II (red dashed lines). In the case of monetary shocks, we see that the possibility of a portfolio adjustment amplifies significantly (but not dramatically) the response of output. In order to get some intuition for this result, consider the budget constraint facing hand-to-mouth households, shown in (24).

Averaging the budget constraint over poor hand-to-mouth households in period t we get:

$$C_t^P = \Xi_t^P W_t N_t + F_t^P - \psi Y \ (\widehat{R}_t + \Omega_{t-1}^P - 1)$$

where $F_t^P \equiv R_t^e E_{t-1|t}^P - \chi_t^P$ denotes average withdrawals from the equity account by wealthy hand-to-mouth households, where $E_{t-1|t}^P$ denotes average equity holdings in period $t-1$ among period t poor hand-to-mouth. Note that $F_t^P \simeq 0$ to the extent that most poor hand-to-mouth households in period t were also in that group in period $t-1$.

Similarly for wealthy hand-to-mouth households:

$$C_t^W = \Xi_t^W W_t N_t + F_t^W - \psi Y \ (\widehat{R}_t + \Omega_{t-1}^W - 1)$$

where $F_t^W \equiv R_t^e E_{t-1|t}^W - E_t^W - \chi_t^W$ denotes average net withdrawals from the equity account by wealthy hand-to-mouth households, with obvious notation. Combining both expressions and defining average hand-to-mouth consumption as $C_t^H \equiv (\lambda_t^P C_t^P + \lambda_t^W C_t^W)/(\lambda_t^P + \lambda_t^W)$, we can write:

$$C_t^H = \Xi_t^H W_t N_t + F_t^H - \psi Y \ (\widehat{R}_t + \Omega_{t-1}^H - 1) \tag{26}$$

where $\Xi_t^H \equiv (\lambda_t^P \Xi_t^P + \lambda_t^W \Xi_t^W)/(\lambda_t^P + \lambda_t^W)$, $F_t^H \equiv (\lambda_t^P F_t^P + \lambda_t^W F_t^W)/(\lambda_t^P + \lambda_t^W)$, and $\Omega_{t-1}^H \equiv (\lambda_t^P \Omega_{t-1}^P + \lambda_t^W \Omega_{t-1}^W)/(\lambda_t^P + \lambda_t^W)$.

In the case of HANK-II we had an identical expression for average hand-to-mouth consumption but with $F_t^H \equiv \Theta_t^H D_t$, which evolves exogenously. By contrast, in HANK-III F_t^H is an endogenous variable. The difference between the two models lies in the fact that wealthy hand-to-mouth households in HANK-III can smooth fluctuations in their cash-on-hand by adjusting their equity balance (albeit at a cost). This is not possible in HANK-II since stocks are not tradable, which makes hand-to-mouth households' consumption vary one-for-one with their current income, $\Xi_t^H W_t N_t + \Theta_t^H D_t$, which they take as exogenous. As discussed above, in HANK-II the decline in dividends in response to an expansionary monetary policy shock has a negative effect on the consumption of the hand-to-mouth through the income distribution channel, which partly offsets the positive impact of the interest rate exposure channel. By contrast, in HANK-III the decline in dividends does not directly impact their cash-on-hand unless it is reflected in lower withdrawals from the equity account. As a result the relative importance of the interest rate exposure channel is enhanced, leading to a stronger response of aggregate consumption and output.

In the case of technology shocks, the difference between HANK-II and HANK-III is more dramatic. As shown in Figure 9, aggregate output *falls* in response to a *positive* technology shock in HANK-III, which contrasts with the more conventional increase in output predicted by HANK-II. The intuition for that result is as follows. A positive technology shock tends to lower employment and labor income, which by itself should lower consumption of hand-to-mouth households (poor and wealthy). In HANK-II, this is more than compensated by the increase in dividend income, causing a mild expansion. This is not the case in HANK-III, because poor hand-to-mouth households do not benefit at all from the higher dividends, while the wealthy hand-to-mouth cannot freely convert dividends into available cash-on-hand. Accordingly, the overall cash-on-hand of hand-to-mouth households declines and so does their consumption, causing aggregate demand and output to fall.

Note that the previous finding of a fall in output in response to a positive technology shock, which contrasts with existing evidence, should not be held against the empirical merits of HANK-III since it hinges critically on our (counterfactual) assumption of a constant real rate.[25] In section 6 below we show how the sign of that response switches from negative to

[25] Galí (1999) and Basu et al. (2006)) find that employment decreases in response to a positive technology

positive when we assume a more realistic monetary policy response.

5.1 A Tractable Counterpart to HANK-III

Next we consider a version of a TANK model which aims at capturing in a stylized way the main features of HANK-III. We refer to this model as TANK-III. As in TANK-II, we assume a constant fraction λ^H of identical hand-to-mouth households whose consumption is given by

$$C_t^H = \Xi^H W_t N_t + \Theta^H D_t - \psi Y \ \widehat{R}_t \tag{27}$$

where Ξ^H denotes the productivity of hand-to-mouth households. The difference with TANK-II is that here Θ^H, representing the fraction of aggregate profits accruing to each hand-to-mouth household, is a free parameter, decoupled from Ξ^H.[26]

We can write aggregate consumption as

$$\begin{aligned} C_t &= \lambda^H C_t^H + (1-\lambda^H) C_t^U \\ &= \lambda^H \Xi^H W_t N_t + \lambda^H \Theta^H D_t - \lambda^H \psi Y \ \widehat{R}_t + (1-\lambda^H) C_t^U \end{aligned} \tag{28}$$

We can compare the previous expression with its counterpart in HANK-III:

$$C_t = \lambda_t^H \Xi_t^H W_t N_t + \lambda_t^H F_t^H - \lambda_t^H \psi Y \ [\widehat{R}_t + \Omega_{t-1}^H - 1] + (1-\lambda_t^H) C_t^U \tag{29}$$

As in the analysis of HANK-II and TANK-II, we see that the first, third and fourth terms on the right hand side of (28) will be a good approximation to their counterparts in (29) if $\lambda_t^H \simeq \lambda^H$, $\Omega_{t-1}^H \simeq 1$, $\Xi_t^H \simeq \Xi^H$ and $v_t^U \simeq v^U$ for all t. Again, this will be true if variations over time in λ_t^H and Ξ_t^H, as well as the gap between Ω_{t-1}^H and 1, are sufficiently small, and if the impact of the shock on aggregate precautionary savings is small.

This leaves us with the second term in (28) and (29). In principle they are not directly comparable since the wealthy hand-to-mouth in the TANK-III model cannot adjust their equity holdings, in contrast with their counterparts in HANK-III. Our strategy is to calibrate Θ^H in order to minimize the gap between $\Theta^H D_t$ and F_t^H in response to a unit increase in aggregate

shock, but not output.

[26] Recall that in TANK-II we had

$$\Theta^H \equiv \vartheta + (1-\vartheta)\Xi^H > \Xi^H$$

where the inquality followed from the fact that Ξ^H was less than one.

dividends D_t, as implied by HANK-III. With that goal in mind we set $\Theta^H = \frac{\partial F_t^H/\partial\varepsilon_t}{\partial D_t/\partial\varepsilon_t}$ where $\partial F_t^H/\partial\varepsilon_t$ and $\partial D_t/\partial\varepsilon_t$ respectively denote the impact responses of F_t^H and D_t to a shock ε_t. Since the latter statistic as implied by our calibrated HANK-III model is (slightly) different across the two shocks considered, we take a simple average between the two in our calibration below.

In Figure 10 we display the responses of aggregate output to monetary policy and technology shocks in a calibrated version of TANK-III, where we set $\lambda^P = 0.05$, $\lambda^W = 0.25$, $\Xi^P = 0.43$, and $\Xi^W = 0.66$, which correspond to their steady state counterparts in our calibrated HANK-III model. The previous settings in turn imply an average productivity for hand-to-mouth households of $\Xi^H = 0.62$. Following the procedure discussed above we set $\Theta^H = 0.58$. The previous calibration implies $\lambda^H\Xi^H - \lambda^H\Theta^H = 0.04 > 0$, a negative relation between the average markup and aggregate hand-to-mouth consumption, given output, in contrast with our calibrated TANK-II model (see previous section). This is a consequence of a relatively lower dividend income share for hand-to-mouth households, which implies a lower income for that group when markups go up. The fact that $\lambda^H\Xi^H \simeq \lambda^H\Theta^H$ in our calibration implies that such an income distribution channel is, however, very weak quantitatively. Furthermore, the fact that $\lambda^H\Xi^H = 0.18$ is relatively small implies that qualitatively similar results can be obtained in a version of the TANK-III model that makes the extreme assumption of $\Theta^H = 0$.

For the sake of comparison, Figure 9 also displays the corresponding impulse responses generated by TANK-II. Notably, the TANK-III model is able to capture, at least qualitatively, the difference in the responses originated by the introduction of portfolio choice in its HANK counterpart and, in particular, the amplification of the effects of monetary policy shocks, as well as the reversal of sign in the response to a technology shock.

Finally, Figure 10 displays a decomposition of the response of aggregate consumption to monetary policy and technology shocks in HANK-III and TANK-III into the components associated with the responses of a hypothetical representative household with no precautionary savings, the poor hand-to-mouth, the wealthy hand-to-mouth and (in the case of HANK-III) a residual that combined the effects associated with changes in the risk and composition shifters and the imperfect correlation between dividends and equity withdrawals. In the case of the monetary policy shock the decomposition is very similar between the two models, suggesting that not only the aggregate effects but the channels at work are also similar. This is less so in the case of technology shocks, in particular given the substantial role of the residual component.

6 Endogenous Monetary Policy

The analysis in the previous sections has assumed an exogenous real interest rate path. As discussed above, the reason for doing this was to make sure that the economy's aggregate response to those shocks was not affected by the choice of a monetary policy rule, since different assumptions regarding the latter would generally imply different real rate paths, preventing us from insulating the impact of heterogeneity itself. In the present section we relax that (admittedly unrealistic) assumption by allowing for an endogenous response of monetary policy.

In particular we consider a simple Taylor-type rule which has the central bank adjust the nominal rate i_t according to:

$$i_t = r + \phi_\pi \pi_t + \phi_y y_t + v_t \tag{30}$$

where v_t is an exogenous monetary policy shifter following the $AR(1)$ process $v_t = \rho_v v_{t-1} + \varepsilon_t^v$.[27]

In addition, and also in the spirit of making the models considered more realistic, we assume bonds are nominal, implying their ex-post real return has an unanticipated component, associated with unexpected inflation. The previous two changes imply that equilibrium output is no longer invariant to the evolution of prices, so we need to include the New Keynesian Phillips curve (7) as part of the set of equations describing the economy's equilibrium.

Figure 12 displays the response of aggregate output in several of the calibrated models considered above to a monetary policy and a technology shock, under the assumption that $\phi_\pi = 1.5$ and $\phi_y = 0.5/4 = 0.125$, as in Taylor (1993). The remaining parameters for each model are calibrated as before. As the Figure makes clear, the assumption of an endogenous response reduces even further the gap between the predictions of RANK, TANK and HANK models regarding the aggregate output response to both monetary policy and technology shocks, thus strengthening the view of a limited role for the presence of idiosyncratic income risk as a factor shaping aggregate fluctuations. Panel (b) displays the simulated path of log output generated by the different models considered in response to a sequence of monetary policy and

[27]Note that the mean of (log) output is normalized to zero under our calibration. Hence, y_t can be interpreted a deiviations from steady state.

technology shocks, respectively. The similarity in the predicted paths is striking, making it hard to distinguish more than a single trajectory.

What is the explanation behind that finding? It follows from two properties of our model. First, as discussed in section 2, the natural level of output, y_t^n, is invariant to the presence of idiosyncratic income risk, being determined by the supply block of the underlying NK framework, which is common across all the models considered. Secondly, the New Keynesian Phillips curve (7), which is shared by all the models analyzed above, displays the *divine coincidence* property, namely, stabilization of inflation implies stabilization of the output gap, and viceversa.[28] It follows that a monetary policy rule that tends to stabilize inflation, as it is the case with rule (30), will reduce the distance between the equilibrium output path generated by any of the model economies considered above (RANK, TANK, or HANK) and their common natural output path. As a result, the distance between their respective implied equilibrium output paths will also shrink.

The previous reasoning can be taken to the limit, and applied to the case of a strict inflation targeting policy. We state its implication in the form a simple proposition.

Proposition *[heterogeneity irrelevance for aggregate output under strict inflation targeting]*: under a strict inflation targeting policy (i.e. $\pi_t = 0$ for all t) all the HANK, TANK and RANK models considered above generate an identical equilibrium output path, which corresponds to that of natural output, which is common across models. In that case equilibrium output is invariant to the presence of heterogeneity.

Proof: it follows directly from equation (7).

A caveat is warranted regarding the previous irrelevance result: the fact that equilibrium output is identical across models under strict inflation targeting does not mean that this is also the case for other variables, including the real interest rate and the distribution of consumption.

7 Caveats and Further Discussion

Our analysis suggests that the role of heterogeneity in shaping economic fluctuations, as stressed in the HANK literature, can be largely understood through simpler frameworks that focus on

[28]See, e.g. Blanchard and Galí (2007) for an early discussion of the divine coincidence.

the distinction between two types of households —unconstrained and hand-to-mouth— but which abstract from the presence of idiosyncratic risk.

A few considerations are in order regarding the relevance and the limitations of our results.

First, our main result applies to environments where idiosyncratic income risk and the associated precautionary savings motive play a limited role for the transmission of aggregate shocks. This is the case in the HANK models considered above, where the "risk-shifter" is largely insensitive to aggregate shocks. In this respect, our quantitative results are consistent with the empirical findings of Berger et. al (2023), who use U.S. household survey data on consumption (CEX) to measure the aggregate implications of imperfect risk sharing in a broad class of HANK models, and find that *wedges* capturing deviations from perfect risk sharing only account for 7% of output fluctuations.

Larger fluctuations in the "risk-shifter" would naturally arise in the presence of countercyclical income risk, an aspect that has been ignored in our analysis, but that has been shown to be empirically relevant for understanding the cyclical properties of income and wealth distribution (see e.g. Bayer et. al. (2019) and Patterson (2023)).[29] A separate question is to understand to what extent fluctuations in income risk translate into fluctuations in consumption risk. For instance, Acharya and Dogra (2020) consider a heterogeneous agent model with CARA preferences, where all agents display a low marginal propensity to consume, and thus where cyclical income risk have quantitatively small effects on the behavior of aggregate consumption. Bilbiie (2023) considers a two-agent model with cyclical idiosyncratic risk, modeled as a time varying probability that a (rich) unconstrained household could become a (poor) hand-to-mouth household in the future period. In that environment, rich households experience a large drop in consumption when hit by a negative idiosyncratic shock, and the precautionary savings motive plays a more prominent role.

Second, our analysis has focused mainly on the effects of monetary policy and technology shocks, while abstracting from other sources of economic fluctuations. In particular, Auclert et. al. (2023) study the effects of fiscal shocks through the lens of an *intertemporal* Keynesian cross logic, where a key role is played by the *intertemporal* marginal propensities to consume out of

[29] A simple way to incorporate the role of cyclical income risk into our TANK models would be to consider a time-varying difference in the productivity of hand-to-mouth and unconstrained households, e.g.. letting $\Xi_t^H - \Xi_t^U$ depend on aggregate output.

income shocks (iMPCs), and their interactions with public deficits. Determining whether simple TANK models can account for the empirical evidence on iMPCs remains an open question. For instance, Fagereng et. al. (2021), using rich tax-registry data for Norwegian households, estimate large MPCs out of lottery wins on impact (about 0.5), but which persist for several years. As argued by Auclert et. al. (2023), this finding could be rationalized by certain HANK models, but is inconsistent with representative agent models —featuring a low MPC at all horizons— and with TANK models —where the MPC falls abruptly after one period. Sahm et. al. (2010), Borusyak et al. (2024), and Orchard et. al. (2023) estimate the MPCs out of the 2008 rebate using U.S. survey data, and find a smaller MPC on impact (below 0.3) that remain positive for at most few months. Similar findings emerge in Boehm et al. (2024) using a randomized experiment involving a debit card gift to a subset of bank customers. Some of these findings can be matched by the simple TANK models discussed above, which would then provide a good approximation to study the effects of fiscal shocks.

Third, we have argued that household heterogeneity plays a limited role for aggregate fluctuations when the central bank seeks to stabilize inflation, as it is the case when it follows an empirically plausible Taylor rule. This result rationalizes the empirical findings of Bayer et. al. (2024) and Bilbiie et. al. (2023) who estimate medium-scale heterogeneous agent models, and conclude that household heterogeneity does not fundamentally alter our understanding of the causes and consequences of aggregate fluctuations.[30] Also, our results are broadly consistent with the findings in McKay and Wolf (2023), who argue that many of the redistributive channels at work in HANK economies operate in opposite directions, and tend to offset each other, so that the response of aggregate consumption is not too dissimilar to what would arise in a representative agent model, even though the transmission channels could be different.

Fourthly, for the sake of simplicity we have assumed flexible wages and a constant aggregate wage markup throughout our analysis. Introducing some form of wage stickiness (real or nominal) would affect the relationship between price inflation, the average markup, and output. This would alter the response of profits to shocks and hence how the income distribution channel operates, under an exogenous real rate. Furthermore, under a more general policy rule, the introduction of sticky wages would no longer stabilization imply a "divine coincidence" between

[30]Similar findings are obtained by Faia and Shabalina (2024) in recent work.

the stabilization of price inflation and that of the output gap. Accordingly, the proposition in section 6 would no longer hold as stated, though it is likely to hold if the central bank stabilizes "composite inflation."[31] We leave an extension to the case of sticky wages for future research.

Importantly, note that our analysis has deliberately abstracted from heterogeneity impacting the economy through supply side channels. An interesting question that we leave for future research is whether HANK economies where heterogeneity affects the supply side of the economy (e.g. due to segmented labor markets, and/or the presence of heterogenous firms) could also be approximated by simpler alternative frameworks.[32] It should be clear, however, that to the extent that the presence of heterogeneity affects the natural level of output, the irrelevance proposition found above will no longer obtain.

Lastly, our analysis has refrained from normative considerations, such as the implications of heterogeneity for the optimal design of monetary policy. Several studies, using both tractable and rich quantitative models, have shown that stabilizing inflation is no longer optimal in the presence of inequality, as monetary policy may be used to partially offset the redistributional effects of aggregate disturbances (see e.g. Acharya et. al. (2023), Bhandari et. al. (2021), Challe (2020), Davila and Schaab (2023) and Smirnov (2023)).

8 Concluding Comments

The emergence of HANK models has been viewed as a challenge to the heretofore dominance of the representative household paradigm in the modelling of aggregate fluctuations and their interaction with macro policies.

In the present paper we have sought to understand the role of idiosyncratic income risk –the key source of heterogeneity in existing HANK models– in shaping aggregate fluctuations by comparing the aggregate properties of three different versions of a HANK model to those of three tractable counterparts that abstract from idiosyncratic risk. In our effort to understand the mechanisms at work in the different HANK models and to design a tractable counterpart to each of them we have stressed the distinction between unconstrained and hand-to-mouth

[31] See chapter 6 in Galí (2015).

[32] See, e.g. Andreolli et al. (2024), for an analysis of a TANK economy where a non-homotheticity in prefereces, combined with the interaction of household heterogeneity (à la TANK) and firm heterogeneity (in the composition of their workforce) leads to an amplification of monetary policy shocks.

households, a distinction which is the hallmark of TANK models. For each HANK model considered, we have found a suitably specified and calibrated tractable model that captures reasonably well its implications for aggregate output and the main channels through which aggregate shocks are transmitted. That similarity increases dramatically in the presence of a policy rule that emphasizes inflation stability. Finally, we have shown that heterogeneity becomes irrelevant for the determination of aggregate output in the limiting case of a strict inflation targeting policy.

References

Acharya, Sushant and Keshav Dogra (2020): “Understanding HANK: Insights From a PRANK,” *Econometrica*, 88 (3), 1113–1158.

Acharya, Sushant, Edouard Challe, and Keshav Dogra (2023): ”Optimal Monetary Policy According to HANK,” *American Economic Review* 113(7), 1741-1782.

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry and Christian Wolf (2018): ”When inequality matters for macro and macro matters for inequality,” *NBER Macroeconomics Annual*, 32(1), pp.1-75.

Aiyagari, Rao (1994): ”Uninsured Idiosyncratic Risk and Aggregate Savings,” *Quarterly Journal of Economics*, Volume 109(3), 659–684.

Alves, Felipe, Greg Kaplan, Benjamin Moll, and Giovanni L. Violante (2020): ”A Further Look at the Propagation of Monetary Policy Shocks in HANK,” *Journal of Money, Credit and Banking* 52(S2), 521-559.

Andreolli, Michele, Natalie Rickard, and Paolo Surico (2024): ”Non-Essential Business Cycles,” unpublished manuscript.

Auclert, Adrien (2019): “Monetary Policy and the Redistribution Channel,” *American Economic Review*, 109 (6): 2333–67.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie and Ludwig Straub (2021): “Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models,” *Econometrica*, 89(5), 2375-2408.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub (2023): ”The Intertemporal Keynesian Cross,” *Journal of Political Economy*, forthcoming.

Basu, Susanto, John Fernald, and Miles Kimball (2006): “Are Technology Improvements Contractionary?,” *American Economic Review*, vol. 96, no. 5, 1418-1448.

Bayer, Christian, Ralph Luetticke, Lien Pham-Do and Volker Tjaden (2015): “Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk,” *Econometrica*, forthcoming.

Bayer, Christian, Benjamin Born, and Ralph Luetticke (2024): ”Shocks, Frictions, and Inequality in U.S. Business Cycles,” *American Economic Review* 114(5), 1212-1247.

Berger, David, Luigi Bocola, and Alessandro Dovis (2023): ”Imperfect Risk Sharing and the Business Cycle,” *Quarterly Journal of Economics* 138(3), 1765-1815.

Bewley, Truman (1983): ”A Difficulty with the Optimum Quantity of Money”, *Econometrica*, 51(5), 1485-1504.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J. Sargent (2021): ”Inequality, Business Cycles, and Monetary-Fiscal Policy,” *Econometrica* 89(6), 2559-2599.

Bilbiie, Florin (2008): ”Limited Asset Market Participation, Monetary Policy and (Inverted) Aggregate Demand Logic,” *Journal of Economic Theory*, 140(1), 162-196.

Bilbiie, Florin (2020): ”The New Keynesian Cross,” *Journal of Monetary Economics* 114, 90-108.

Bilbiie, Florin and Ragot, Xavier (2021): “Optimal Monetary Policy and Liquidity with Heterogeneous Households,” *Review of Economic Dynamics* 41, 71-95.

Bilbiie, Florin O. (2024): “Monetary Policy and Heterogeneity: An Analytical Framework,” mimeo.

Bilbiie, Florin, Giorgio Primiceri and Andrea Tambalotti (2023): ”Inequality and Business Cycles,” mimeo

Blanchard, Olivier, and Jordi Galí (2007): ”Real Wage Rigidities and the New Keynesian Model,” *Journal of Money, Credit and Banking* 39, 35-65.

Boehm, Johannes, Etienne Fize and Xavier Jaravel (2’24): ”Five Facts about MPCs: Evidence from a Randomized Experiment,” unpublished manuscript.

Boppart, Timo, Per Krusell, and Kurt Mitman (2018): “Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative,” *Journal of Economic Dynamics and Control*, 89, 68–92.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2024): ”Revisiting Event Study Designs: Robust and Efficient Estimation,” *Review of Economic Studies*, forthcoming

Broer, Tobias, Niels-Jakob Harbo Hansen, Per Krusell, and Erik Oberg (2020): “The New Keynesian Transmission Mechanism: A Heterogeneous-Agent Perspective,” *The Review of Economic Studies*, 87 (1), 77–101.

Campbell, John Y. and N. Gregory Mankiw (1989): ”Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence,” *NBER Macroeconomics Annual,* 4, 185-216.

Cantore, Cristiano and Lukas B. Freund (2021): "Workers, Capitalists and the Government: Fiscal Policy and Income (Re)Distribution," *Journal of Monetary Economics* 119, 58-74.

Carroll, Christopher D. (2006): "The method of endogenous gridpoints for solving dynamic stochastic optimization problems." *Economics Letters* 91(3), 312-320.

Challe, Edouard and Xavier Ragot (2016): "Precautionary Saving Over the Business Cycle," *The Economic Journal* 129(590), 135-164.

Challe, Edouard (2020): "Uninsured Unemployment Risk and Optimal Monetary Policy in a Zero-Liquidity Economy," *American Economic Journal: Macroeconomics* 12(2), 241-283.

Dávila, Eduardo, and Andreas Schaab (2023): "Optimal Monetary Policy with Heterogeneous Agents: Discretion, Commitment, and Timeless Policy,". No. w30961. National Bureau of Economic Research.

Debortoli, Davide and Jordi Galí (2024): "Idiosyncratic Income Risk and Aggregate Fluctuations," *American Economic Journal: Macroeconomics*, forthcoming.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik (2021): "MPC Heterogeneity and Household Balance Sheets," *American Economic Journal: Macroeconomics* 13(4), 1-54.

Faia, Ester and Ekaterina Shabalina (2024): "What Central Bankers Can Learn from Models with Idiosyncratic Risk," mimeo.

Farhi, Emmanuel and Iván Werning (2017b): "Monetary Policy, Bounded Rationality and Incomplete Markets," mimeo.

Floden, Martin and Jesper Lindé (2001): "Idiosyncratic Risk in the United States and Sweden: Is there a Role for Government Insurance?," *Review of Economic Dynamics* 4(2), 406-437.

Galí, Jordi (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, vol. 89, no. 1, 249-271.

Galí, Jordi (2015): *Monetary Policy, Inflation and the Business Cycle*, Second edition, Princeton University Press.

Galí, Jordi (2018): "The State of New Keynesian Economics: A Partial Assessment," *Journal of Economic Perspectives* 32(3), 87-112.

Galí, Jordi, J. David López-Salido and Javier Vallés (2007): "Understanding the Effects of

Government Spending on Consumption", *Journal of the European Economic Association*, vol. 5 (1), 2007, 227-270.

Gornemann, Nils, Keith Kuester and Makoto Nakajima, (2016): "Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy", mimeo.

Guerrieri, Veronica and Guido Lorenzoni (2017): "Credit Crises, Precautionary Savings, and the Liquidity Trap," *Quarterly Journal of Economics* 132(3), 1427-1467.

Kaplan, Greg, Giovanni L.Violante and Justin Weidner (2014): "The Wealthy Hand-to-Mouth " *Brooking Papers on Economic Activity*, 1, 77-153.

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

Krusell, Per and Anthony A. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106 (5), 867–896.

Nisticó, Salvatore (2016): "Optimal Monetary Policy and Financial Stability in a non-Ricardian Economy," *Journal of the European Economic Association,* 14(5), 1225-1252.

Luetticke, Ralph (2017): "Transmission of Monetary Policy with Heterogeneity in Household Portfolios", mimeo.

McKay, Alisdair and Ricardo Reis (2016) "The Role of Automatic Stabilizers in the U.S. Business Cycle, *Econometrica* 84(1) 141–94.

McKay, Alisdair, Emi Nakamura and Jon Steinsson (2016): "The Power of Forward Guidance Revisited," *American Economic Review*, 106(10), 3133-3158.

McKay, Alisdair and Christian K. Wolff (2023): "Monetary Policy and Inequality," *Journal of Economic Perspectives* 37(1), 121-144.

Orchard, Jacob, Valerie A. Ramey, and Johannes F. Wieland (2023): "Micro MPCs and macro counterfactuals: the case of the 2008 rebates."National Bureau of Economic Research, wp 31584.

Patterson, Christina (2023): "The Matching Multiplier and the Amplification of Recessions," *American Economic Review* 113 (4), 982-1012.

Ravn, Morten and Vincent Sterk (2014): "Job Uncertainty and Deep Recessions," Discussion Papers 1501, Centre for Macroeconomics (CFM).

Ravn, Morten O., and Vincent Sterk. "Macroeconomic fluctuations with HANK & SAM: An analytical approach." Journal of the European Economic Association 19.2 (2021): 1162-1202.

Reiter, Michael (2010): "Solving Heterogeneous Agent Models by Projection and Perturbation," *Journal of Economic Dynamics and Control* 33, 649-665.

Sahm, Claudia R, Matthew D Shapiro, and Joel Slemrod (2010): "The Response to the 2008 Tax Rebate: Survey Evidence and Aggregate Implications," *Tax Policy and the Economy* 24(1): 69-110

Smirnov, Danila (2023): "Optimal Monetary Policy in HANK,". Unpublished manuscript, Universitat Pompeu Fabra.

Taylor, John B. (1993): "Discretion versus policy rules in practice." Carnegie-Rochester Conference Series on Public Policy (Taylor, John B. "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy* (39), 195-214.

Walsh, Carl (2017): "Workers, Capitalists, Wage Flexibility and Welfare," mimeo.

Werning, Iván (2015): "Incomplete Markets and Aggregate Demand," NBER Working Paper No. 21448.

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

Appendix

Derivation of the Approximate Individual and Aggregate Euler Equations

Our starting point is the individual Euler equation

$$C_t(j)^{-\sigma} = \beta R_t \mathbb{E}_t\{C_{t+1}(j)^{-\sigma}\} \tag{31}$$

Substituting a second order approximation of $C_{t+1}(j)^{-\sigma}$ around $C_t(j)$ into (31) yields

$$C_t(j)^{-\sigma} \simeq \beta R_t \mathbb{E}_t \left\{ C_t(j)^{-\sigma} - \sigma C_t(j)^{-\sigma} \left(\frac{\Delta C_{t+1}(j)}{C_t(j)} \right) + \frac{\sigma(\sigma+1)}{2} C_t(j)^{-\sigma} \left(\frac{\Delta C_{t+1}(j)}{C_t(j)} \right)^2 \right\}.$$

Rearranging terms,

$$\mathbb{E}_t \left\{ \frac{\Delta C_{t+1}(j)}{C_t(j)} \right\} \simeq \frac{1}{\sigma}\left(1 - \frac{1}{\beta R_t}\right) + \frac{\sigma+1}{2} v_t(j)$$

where $v_t(j) \equiv \mathbb{E}_t \left\{ \left(\frac{\Delta C_{t+1}(j)}{C_t(j)} \right)^2 \right\} \simeq \mathbb{E}_t\{\xi_{t+1}(j)^2\}$, with $\xi_t(j) \equiv c_t(j) - \mathbb{E}_{t-1}\{c_t(j)\}$ being the innovation in individual consumption.

Rearranging terms, we have:

$$\mathbb{E}_t\{\Delta C_{t+1}(j)\} \simeq \frac{1}{\sigma}\left(1 - \frac{1}{\beta R_t}\right) C_t(j) + \frac{\sigma+1}{2} C_t(j) v_t(j) \tag{32}$$

When all households are unconstrained (as in HANK-I), we can integrate the previous equation over $j \in [0,1]$ and divide the resulting by expression by C_t to obtain:

$$\mathbb{E}_t \left\{ \frac{\Delta C_{t+1}}{C_t} \right\} \simeq \frac{1}{\sigma}\left(1 - \frac{1}{\beta R_t}\right) + \frac{\sigma+1}{2} v_t$$

where

$$v_t \equiv \int \frac{C_t(j)}{C_t} v_t(j) dj$$

The previous equation can be approximated around the stochastic steady state to yield equation (9) in the text. Note that in the stochastic steady state

$$\frac{1}{\sigma}\left(1 - \frac{1}{\beta R}\right) + \frac{\sigma+1}{2} v = 0$$

thus implying $\beta R < 1$. Wealthy households (with high consumption) will have $v_t(j) > v$ and hence will experience a decline in consumption (on average). The opposite will be true for poor households, whose consumption will tend to increase. Consistently with that property, the stochastic steady state is characterized by a well defined distribution of consumption across households (which also corresponds to the ergodic distribution of individual consumption).

When the individual Euler equation only holds for a subset of households $\mathcal{U}_t$ in period t, we can integrate (32) over that subset and rearrange terms to obtain:

$$\mathbb{E}_t\left\{\frac{C^U_{t+1|t}-C^U_t}{C^U_t}\right\}\simeq\frac{1}{\sigma}\left(1-\frac{1}{\beta R_t}\right)+\frac{\sigma+1}{2}v^U_t$$

where $C^U_t=\frac{1}{1-\lambda^H_t}\int_{j\in\mathcal{U}_t}C_t(j)dj$, $C^U_{t+1|t}=\frac{1}{1-\lambda^H_t}\int_{j\in\mathcal{U}_t}C_{t+1}(j)dj$, and $v^U_t\equiv\frac{1}{1-\lambda^H_t}\int_{j\in\mathcal{U}_t}\frac{C_t(j)}{C^U_t}v_t(j)dj$. Equivalently, we can write:

$$\mathbb{E}_t\left\{\frac{C^U_{t+1}-C^U_t}{C^U_t}\right\}\simeq\frac{1}{\sigma}\left(1-\frac{1}{\beta R_t}\right)+\frac{\sigma+1}{2}v^U_t+h^U_t \tag{33}$$

where $h^U_t\equiv\mathbb{E}_t\left\{c^U_{t+1}-c^U_{t+1|t}\right\}$. Note that h_t emerges as a result of changes in the composition of $\mathcal{U}_t$, which imply that some households who are unconstrained a t become constrained at $t+1$, and viceversa, so that in general we have $C^U_{t+1}\neq C^U_{t+1|t}$. Approximating (33) around the stochastic steady state yields equation (15) in the text.

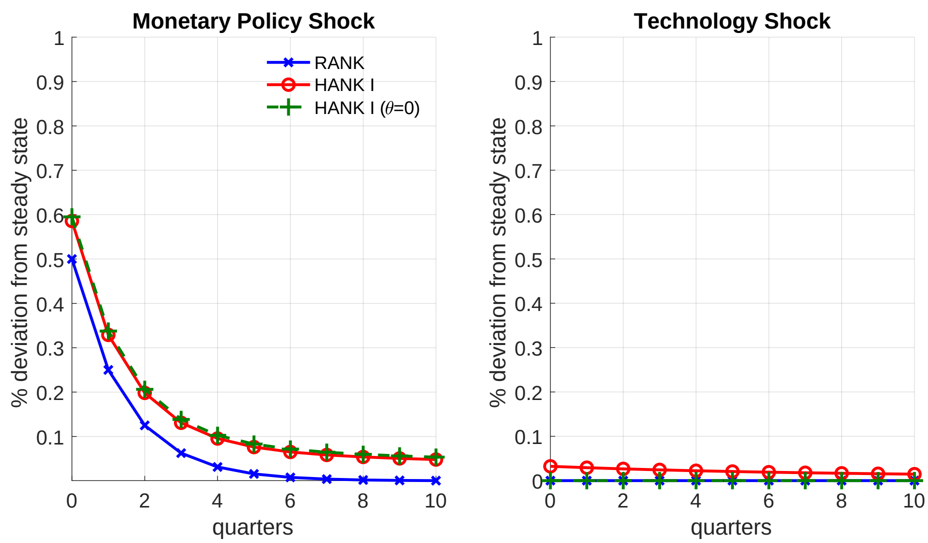
Tables and Figures

Table 1: Calibration

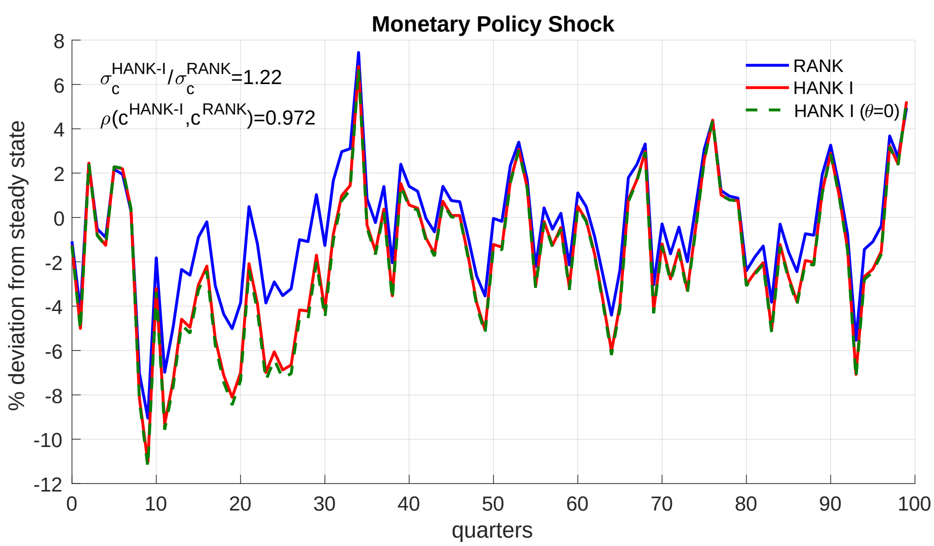
Model parameters		
σ	Coefficient of risk aversion	1
φ	(Inverse) Frisch elasticity	1
$\mathcal{M}$	Average price markup	1.10
$\mathcal{M}_w$	Average wage markup	1.10
κ	Slope of Phillips curve	0.10
Θ	Fraction of profits distributed as lump-sum	0.5
R	Steady state (gross) interest rate	1.005
	$\rightarrow$ Discount factor β:	
	RANK, TANK (I, II, III)	0.995
	HANK (I, II, III)	0.9937, 0.9838, 0.9905
Shocks processes		
ρ_r	Persistence monetary policy shocks	0.5
ρ_a	Persistence technology shocks	0.9
ρ_z	Autocorr. of idiosyn. earnings	0.966
σ_z	Std. dev. of idiosyn. earnings	0.5
Solution Method		
n_z	Gridpoints for idiosyn. earnings	11
n_a	Gridpoints for liquid asset	500
$(\underline{B}, \bar{B})$	Bounds on grid for liquid asset:	
	HANK-I	$(-36.33Y, 300Y)$
	HANK-II and III	$(-2Y, 50Y)$

Figure 1: RANK vs HANK-I

Panel (a)

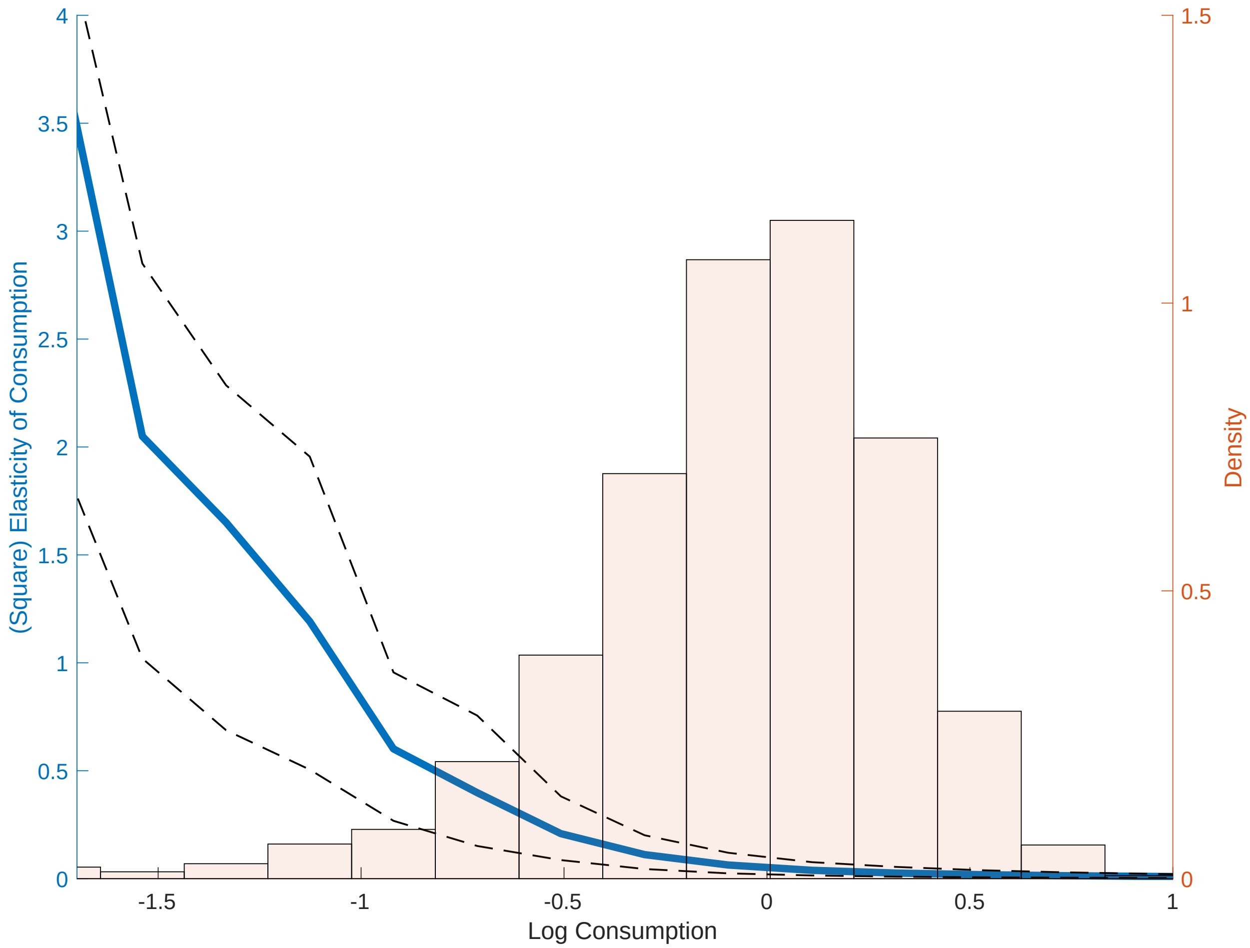


Panel (b)



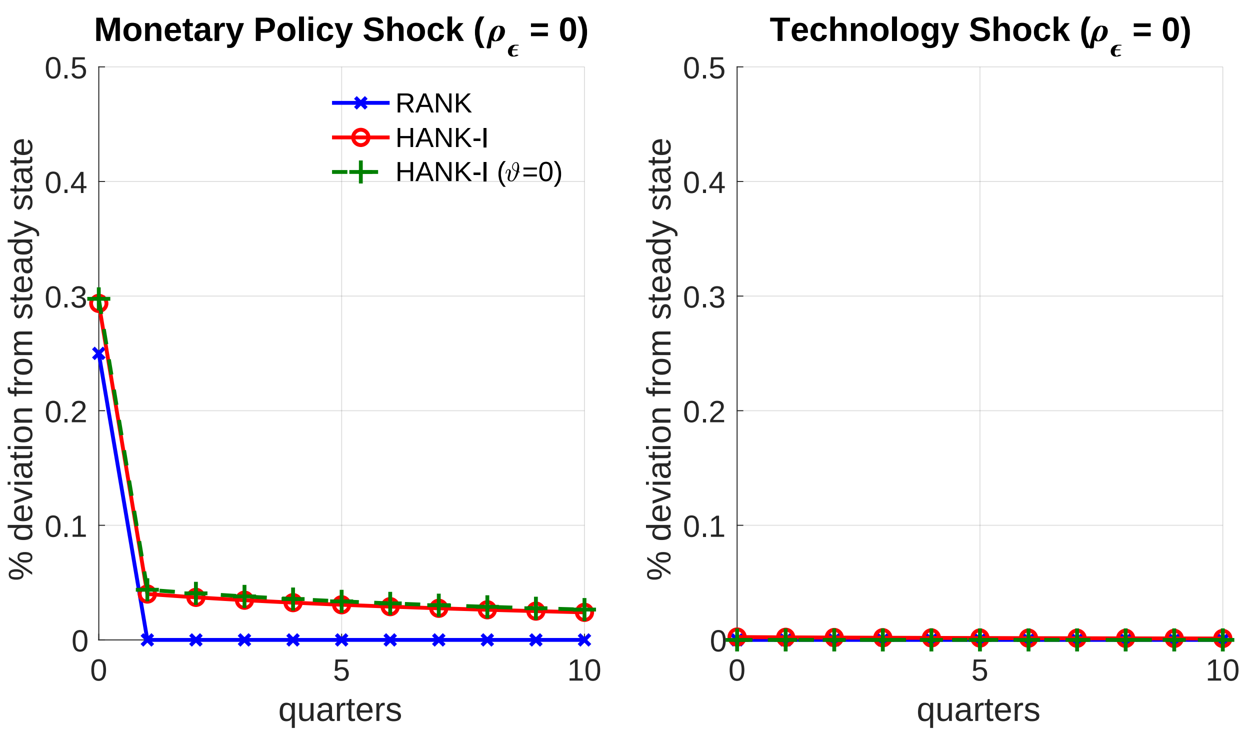
Notes: Panel (a) shows the response of output to a 1 percent decrease in the (annualized) real interest rate (left), and to a 1 percent positive technology shock (right) in a representative agent model (blue line with crosses) and in the heterogeneous agent model with no binding borrowing constraint, with or without an income distribution channel (red line with circles and dashed green line with 'pluses', respectively). Panel (b) shows a simulated path of consumptions in response to monetary policy shocks.

Figure 2: Elasticity of Consumption in Steady State



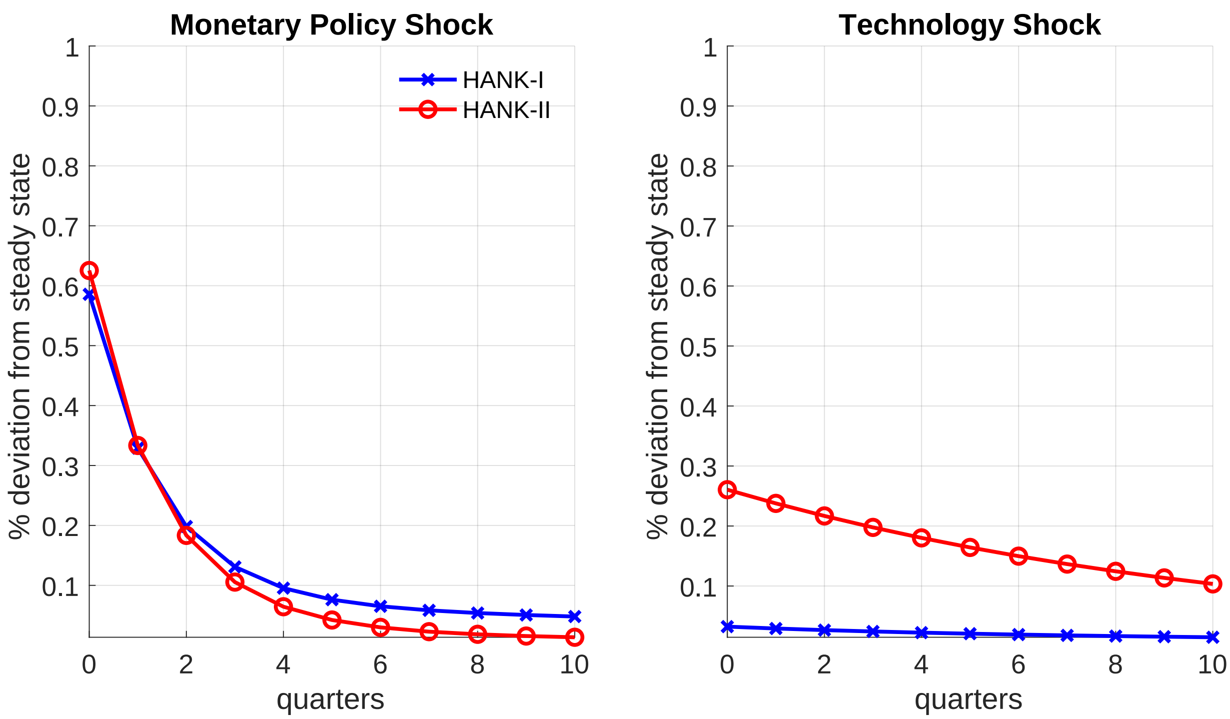
Notes: The figure shows the relationship between log consumption (horizontal axis), and the elasticity of consumption (left vertical axis) in steady state. For each value of consumption, the figure reports the average elasticity (solid blue line), the 5% - 95% interval of the distribution (black dashed lines), while the histogram indicate the steady state distribution (right vertical axis).

Figure 3: Impulse Responses, RANK vs HANK-I (Purely Transitory Shock)



Notes: The figure shows the response of output to a purely transitory 1 percent decrease in the (annualized) real interest rate (left panel), and to a purely transitory 1 percent positive technology shock (right panel) in a representative agent model (blue line with crosses) and in the heterogeneous agent model with no binding borrowing constraint with or without an income distribution channel (red line with circles and dashed green line with 'pluses', respectively).

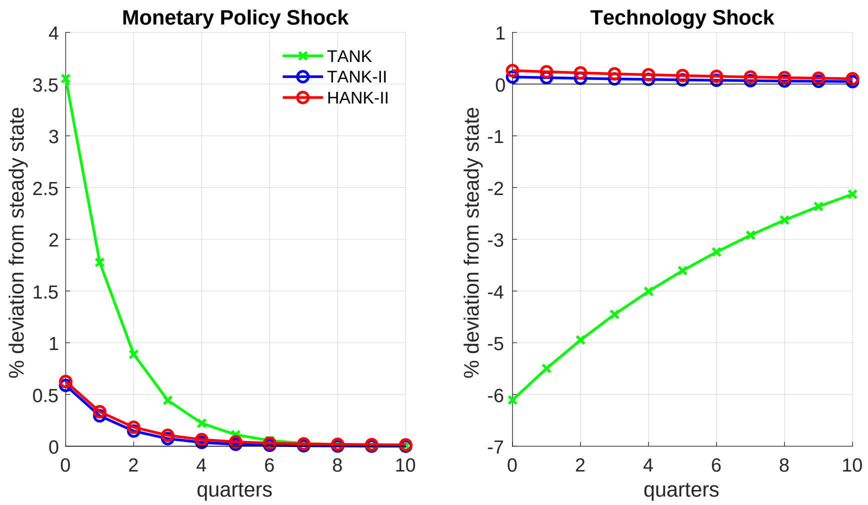
Figure 4: HANK-I vs HANK-II



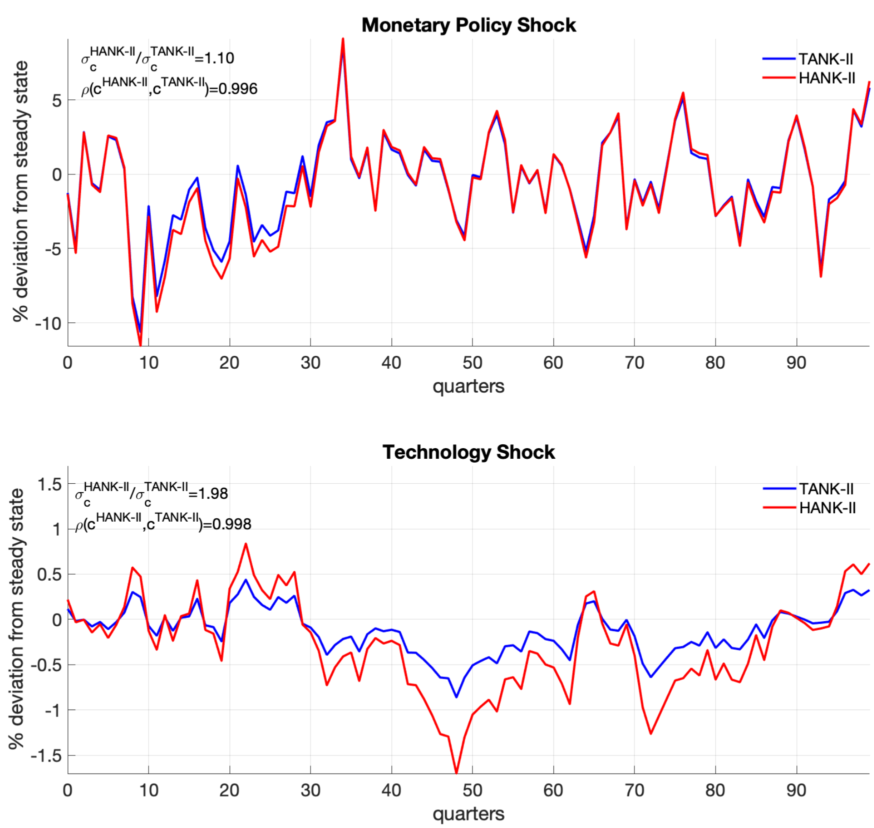
Notes: The figure shows the response of output to a 1 percent decrease in the (annualized) real interest rate (left panel), and to a 1 percent positive technology shock (right panel) in the heterogeneous agent model with no binding borrowing constraint (blue line with crosses), and with binding borrowing constraint for 30 percent of the population (red line with circles).

Figure 5: Simple Alternatives to HANK-II

Panel (a): Impulse Responses

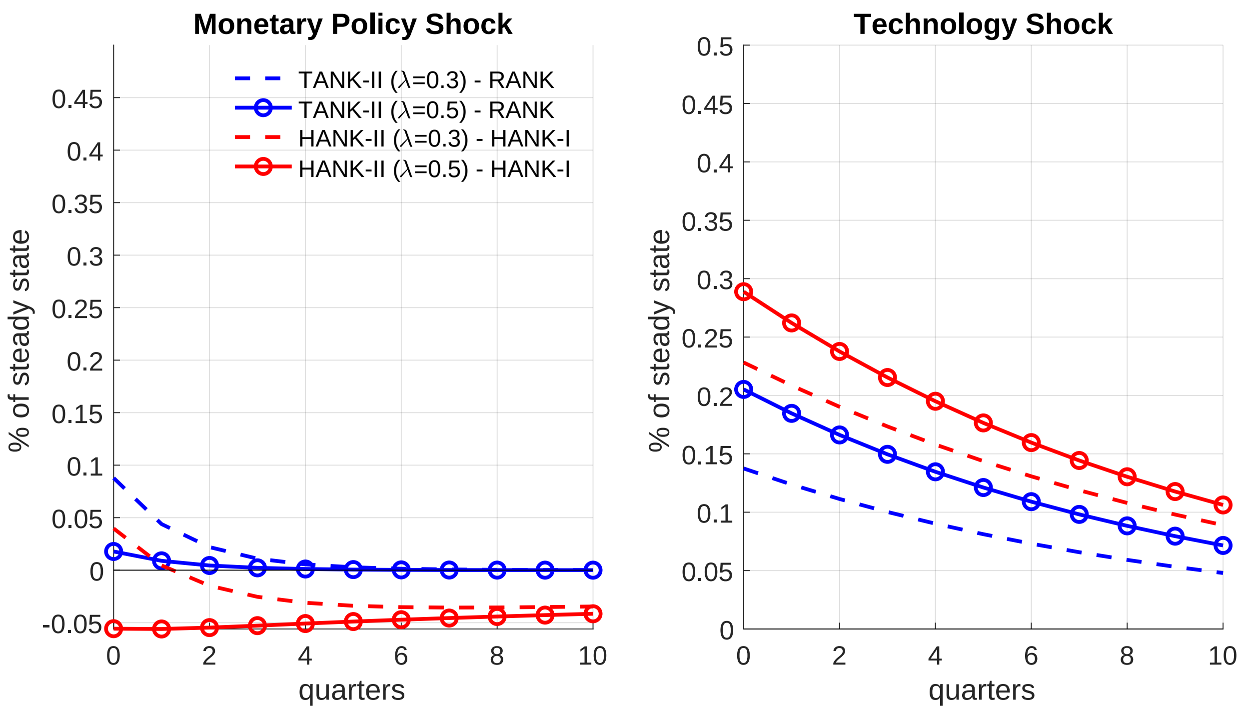


Panel (b): Simulations



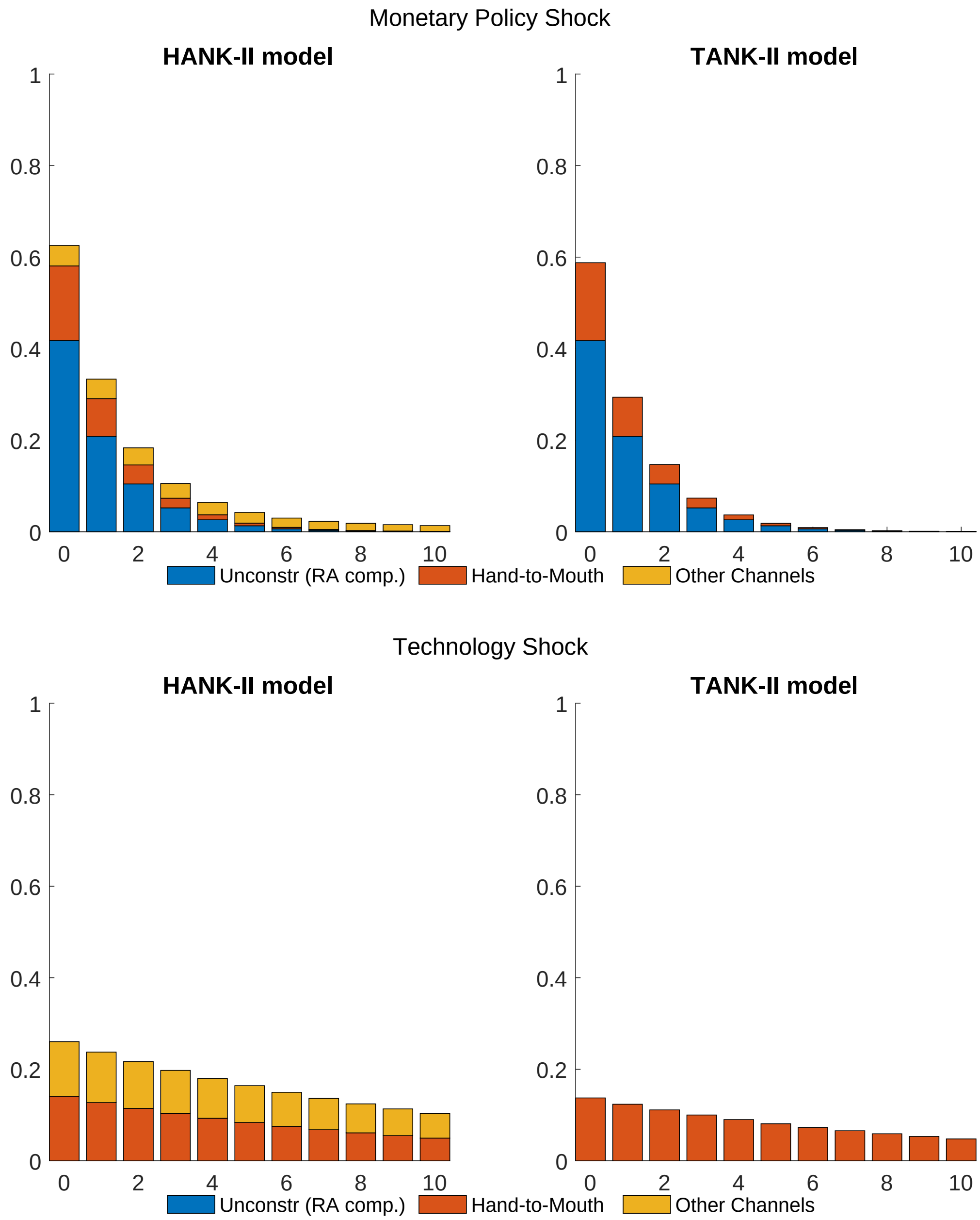
Notes: Panel (a) shows the response of output to a 1 percent decrease in the (annualized) real interest rate (left panel), and to a 1 percent positive technology shock (right panel) in heterogeneous agent model with binding borrowing constraint for 30 percent of the population (red line with circles), and in the TANK (green line with crosses) and TANK-II (blue line with circles) models. Panel (b) shows a simulated path of consumptions in response to monetary policy and technology shocks.

Figure 6: The Role of Binding Borrowing Constraints: HANK-II vs TANK-II



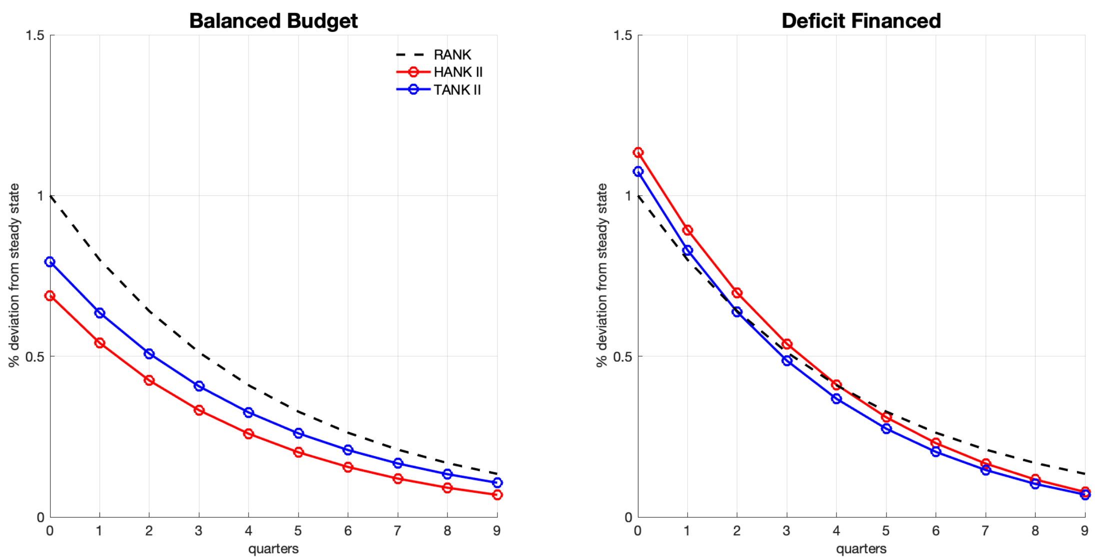
Notes: The figure displays the impulse responses in the HANK-II (red) and TANK-II (blue), for the cases where fraction of constrained agents equals 30 percent (dashed lines) and 50 percent (solid lines with circles), expressed as a gap relative to their counterparts in the absence borrowing constraints (i.e. HANK-I and RANK, respectively).

Figure 7: Decomposition of Transmission Channels: HANK-II vs TANK-II



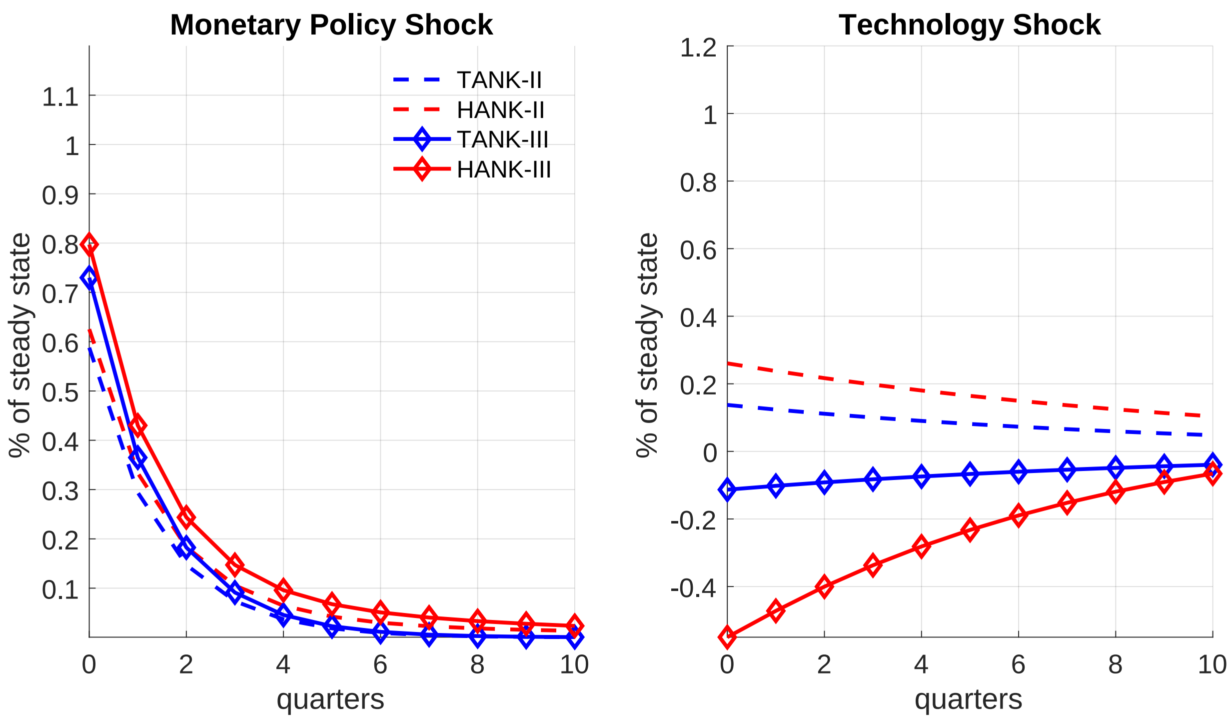
Notes: The figure shows the decomposition of the impulse responses to monetary policy shocks (top panel) and technology shocks (bottom panel) into the three components shown in eq. (23) for both the HANK-II (left column) and TANK-II (right column) models.

Figure 8: The Effects of Government Spending Shocks: HANK-II vs TANK-II



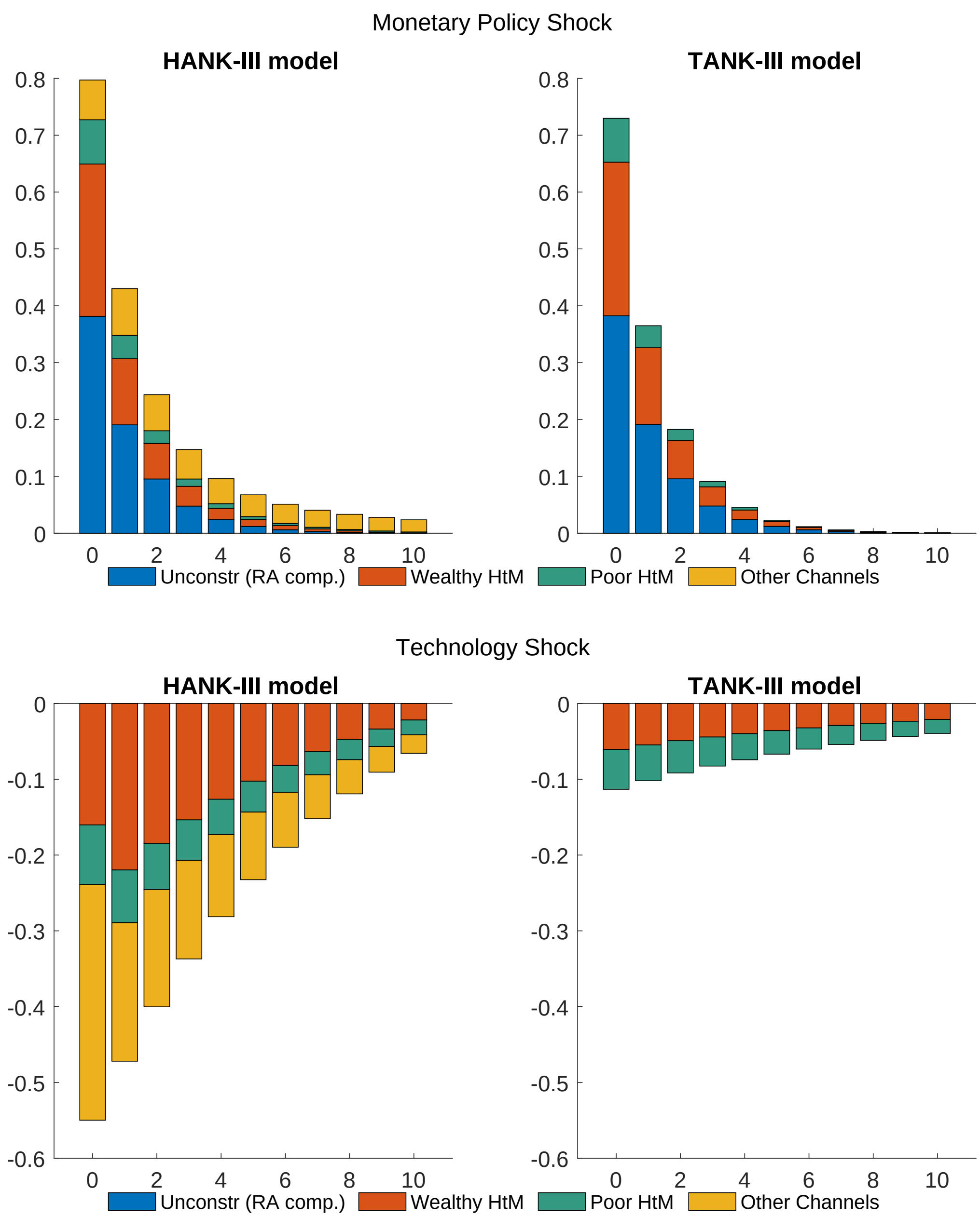
Notes: The figure shows the response of output to a 1 percent of GDP increase in government spending under a balanced budget rule (left panel) and under deficit financing (right panel), in the RANK (black dashed line), HANK-II (red line with circles) and TANK-II (blue line with circles) models.

Figure 9: The Role of Portfolio Adjustment Costs: HANK-III vs TANK-III



Notes: The figure shows the response of output to a 1 percent decrease in the (annualized) real interest rate (left column), and to a 1 percent positive technology shock (right column) in the heterogeneous agent models (red lines) and two-agent models (blue lines), for the case without portfolio adjustment costs (HANK-II and TANK-II, dashed lines), and with portfolio adjustment costs (HANK-III and TANK-III, lines with diamonds).

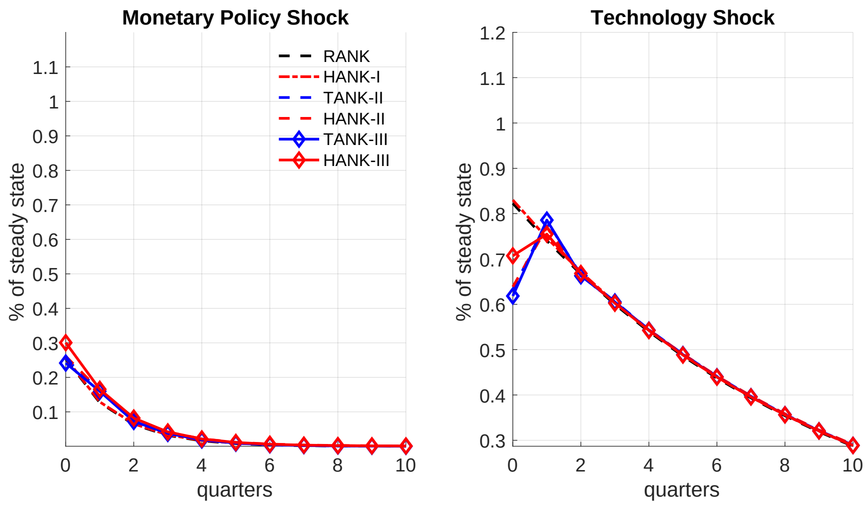
Figure 10: Decomposition of Transmission Channels: HANK-III vs TANK-III



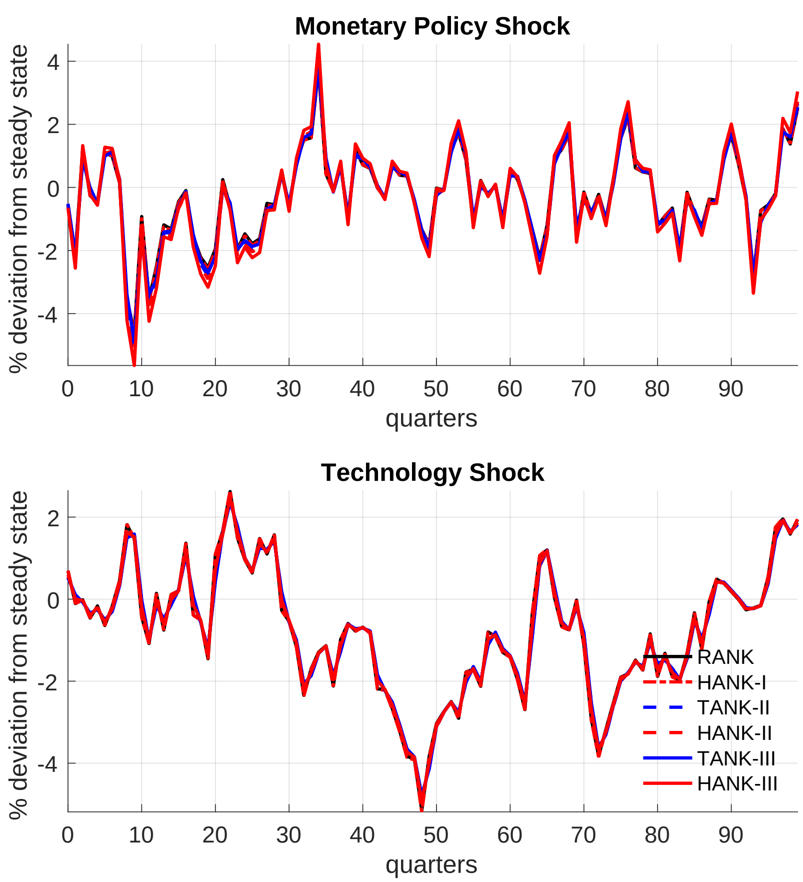
Notes: The figure shows the decomposition of the impulse responses to a monetary policy shocks (top panel) and technology shocks (bottom panel) into the three components shown in eq. (23) for both the HANK-III (left column) and TANK-III (right column) models.

Figure 11: Heterogeneity, Nominal Bonds and Endogenous Monetary Policy

Panel (a): Impulse Responses



Panel (b): Simulations



Notes: Panel (a) shows the response of output to a 1 percent shock to the (annualized) nominal interest rate (left panel), and to a 1 percent positive technology shock (right panel) for the case without portfolio adjustment costs (HANK-II and TANK-II, dashed lines), and with portfolio adjustment costs (HANK-III and TANK-III, lines with diamonds), in the presence of nominal bonds, and assuming the central bank follows a Taylor rule $\hat{i}_t = 1.5\pi_t + 0.5/4y_t$. Panel (b) shows a simulated path of consumptions in response to monetary policy and technology shocks.