Comment on "Heterogeneity and Aggregate Fluctuations: Insights from TANK Models" (Debortoli and Galí, NBER Macro Annual 2024)

Matthew Rognlie

August 2024

Following Campbell and Mankiw (1989), the two-agent model, combining hand-to-mouth and permanent-income households, has become a popular alternative to the representative-agent model. Adding a measure of hand-to-mouth agents, who are unable to borrow and lend at the margin, is a simple way to match high marginal propensities of consume (MPCs) in the data. This approach is especially natural when studying fiscal policy (see, e.g., Galí, López-Salido and Vallés 2007), where MPCs and the associated Keynesian multiplier have long been of interest. "TANK" (two-agent New Keynesian) models are tractable and well-suited for such questions.

More recently, "HANK" (heterogeneous-agent New Keynesian) models have emerged as another alternative to the representative-agent paradigm. They have been used to revisit the transmission of monetary policy (e.g. Kaplan, Moll and Violante 2018), fiscal policy (e.g. Auclert, Rognlie and Straub 2024), and many other core macro topics.

HANK and TANK have important similarities. They both break Ricardian equivalence, can match high MPCs, and have heterogeneous households. But HANK and TANK are also very different: HANK builds on the incomplete-market Bewley-Huggett-Aiyagari tradition, featuring a continuum of heterogeneous households with different incomes and asset positions, while TANK's heterogeneity is limited to just two agents. A tradeoff emerges between richness and tractability: HANK can speak to the detailed interactions between heterogeneity and macro, but TANK requires less computation and can sometimes be solved analytically.

It is natural to ask: if we opt for TANK rather than HANK, just how much are we giving up? Maybe not that much, Debortoli and Galí argue in this paper. In particular, if we are interested in the aggregate effects of standard macro shocks—monetary shocks, productivity shocks, and fiscal shocks—then Debortoli and Galí argue that TANK is an effective substitute for HANK. In their words, "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."

I think this overreaches—and that, indeed, TANK is a far weaker substitute for HANK than this paper indicates. Although the paper's findings of a close fit between TANK and HANK appear impressive, they are to some extent accidental, and specific to the experiments that the paper considers. As I will show, reasonable alternative shocks—such as government spending shocks, financed by debt that is not paid off too quickly—open up a large gap between TANK and HANK.

This is not just a question of quantitative magnitude. It points to a deep, *qualitative* distinction between TANK and HANK, which is that TANK's departure from the representative-agent paradigm is entirely static—limited to hand-to-mouth agents that spend what they receive within a period—while HANK introduces rich, dynamic changes in spending patterns. The difference between TANK and HANK is most apparent when these dynamics are taken into account. I summarize this as the Lesson 1 that I take away from examining this paper.

Lesson 1. *The differences between HANK and TANK are clearest when the question of interest has a strong dynamic component.*

For instance, both HANK and TANK can be calibrated to have similar MPCs, and might behave similarly in response to a government spending shock that is paid for within a few quarters. But they become very different when the debt used to finance spending is more persistent—especially if we look at a cumulative measure of output, rather than just the multiplier at date 0. The more dynamic the shock and measurement, the more HANK vs. TANK matters.

The paper largely misses this point. It emphasizes short-lived shocks, like a monetary shock with a quarterly persistence of 0.5, or a deficit-financed fiscal shock where the debt is paid back on average within 2.5 years.

Instead, the paper spends much of its time on distributional questions, especially what it calls the interest rate exposure channel (how sensitive is each household's cash-on-hand to changes in interest rates?) and the income distribution channel (how do changes in the aggregate markup affect the distribution of income between households?). There is no doubt these channels are important—but they can easily be made *too* important with the wrong assumptions. This paper is a case in point: its rule for real wages implies extremely countercyclical profits, such that an extra \$1 in real output leads to an additional \$3 in wage income and $-\$2$ in dividend income. Needless to say, the paper's HANK and TANK models are extremely sensitive to how this income is distributed. But the premise is absurd—profits are not really so countercyclical, and this fictitious mechanism distracts from other, more economically meaningful ones.

This is a common issue in both HANK and TANK, and I take it as Lesson 2 for the literature.

Lesson 2. *In New Keynesian models with heterogeneity, it is easy to make assumptions with extreme and implausible distributional consequences.*

This often happens when, like here, papers assume sticky prices with flexible wages. This choice was natural in the early days of HANK and TANK, since the baseline representative-agent New Keynesian model ("RANK") made the same assumption. But one side effect of this assumption is a countercyclical profit margin—*very* countercyclical in most calibrations. This is innocuous enough in RANK models, where the same agent receives all income anyway. But it starts to matter a great deal once there is heterogeneity, and it can easily dominate the analysis.

The same is true for the typical assumption that all debt is short-term: often irrelevant in a RANK model, but important once agents have heterogeneous debt positions. The paper's large interest rate exposure channel follows directly from its assumption that when rates change, all debt service is immediately repriced at the new rate.

The paper correctly recognizes these distributional effects as important, and engineers a series of TANK models to match them—but again, they are made far *too* important, especially the income distribution channel, by the paper's assumptions. In my view, the focus on these effects distracts from a more robust comparison of TANK and HANK, as the paper misses not only Lesson 1 but also the following.

Lesson 3. *In a simple benchmark where income shares are unaffected by aggregate output, the general equilibrium response of aggregate consumption to real interest rates is identical in RANK, TANK, and HANK.*

This important neutrality result is from Werning (2015). In its exact form, it requires that (1) all agents' labor incomes and dividends expand proportionally with output, (2) asset positions held coming into date 0 are claims on future dividends, and (3) the elasticity of intertemporal substitution is 1. Quantitatively, however, it is robust to moderate departures from these assumptions. The result holds despite the fact that partial equilibrium consumption responses to interest rates are very different in RANK, TANK, and HANK—because, in the language of Kaplan et al. (2018), the "indirect" (GE) effects through income exactly offset the varying "direct" (PE) effects of interest rates.

In short, monetary policy is not the first place we should look to find a difference between TANK and HANK—or RANK and HANK, for that matter. There are many reasons, of course, why the Werning (2015) neutrality result might not hold in practice—but it is a very natural starting point, and reminds us that *very* different models, with different implications elsewhere, can have the same aggregate response to interest rates. It is thus unfortunate that this paper chooses monetary shocks as its primary experiment.[1]

Summary and outline of comment. In sum, my critique of this paper is the following. Its assumptions imply a counterfactually volatile income distribution (Lesson 2), which is distracting and requires special attention to handle. This volatility obscures the fact that monetary shocks often have similar aggregate consequences regardless of heterogeneity (Lesson 3), which makes them a bad test of the distinction between TANK and HANK. On top of all this, the paper deals mostly with short-lived shocks, missing the dynamic settings where differences between TANK and HANK are most likely to emerge (Lesson 1).

[1] Productivity shocks receive parallel treatment in many places, but given the paper's exogenous-real-rate monetary rule, they only matter due to the effect on distribution, where higher productivity (counterfactually) means much lower wages and higher dividends—making productivity shocks an even less appealing exercise. I discuss these and also fiscal shocks, which were introduced in the final version of the paper, later in this comment.

In the remaining comment, I start in section 1 by documenting the profound differences between HANK and TANK, in both the distribution of MPCs and aggregate MPCs, that emerge when we look beyond one period. Then, in section 2, I go through the paper and discuss its main exercises, showing that mild changes in the analysis—a longer-lived shock, or more cumulative measure of output—result in different conclusions. In section 3, I tweak the models to satisfy Werning (2015) neutrality, and show that there are vast differences in the response to deficit-financed fiscal shocks, even though RANK = TANK = HANK for monetary policy. Finally, in section 4, I discuss the path forward, and the role of tractable models like TANK versus other approaches to make heterogeneity in macro more accessible.

Throughout my discussion, I use exactly the same models and calibrations as in the paper—all unmodified, aside from my Werning (2015) analysis in section 3. My main goal is to show that even with the same models, a closer look gives us different results.[2]

1 Static TANK, dynamic HANK: an overview

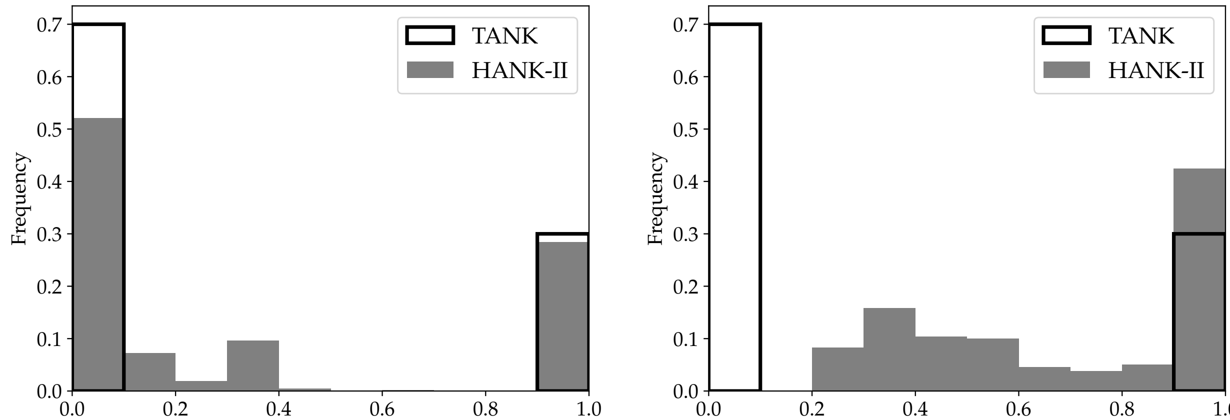
TANK models feature stark MPC heterogeneity: agents are either hand-to-mouth consumers, with MPCs of 1, or unconstrained permanent-income consumers, with MPCs very close to 0. HANK models, in principle, are more general: agents can have a range of MPCs between 0 and 1, depending on their distance to the borrowing constraint.

If we look only at the immediate MPC out of a transfer, however, the contrast is not always so clear. Figure 1(a) compares the distribution of MPCs in Debortoli and Galí's TANK against that in their "HANK-II", where the latter has a standard household side with a binding borrowing constraint. TANK clusters at the extremes: as calibrated, 70% of households are unconstrained with MPCs near 0, and 30% are hand-to-mouth with MPCs of 1. But what about HANK? As targeted by the paper, it features roughly 30% of households with MPCs of 1. Of the remaining 70% of households, around 50% have MPCs below 0.1, and the remaining 20% of households mostly have MPCs in the 0.1–0.4 range. The latter group with intermediate MPCs is special to "HANK"—impossible to achieve in a TANK model—but one might reasonably wonder whether having 20% of the population with mildly elevated MPCs is really such an important distinction.

Figure 1(b) takes a different approach, looking at the distribution of the *expected cumulative* MPC over quarters $t = 0, \dots, 10$ following a transfer. This measure asks: if a household is given \$1 today, what fraction of it do they cumulatively expect to spend, in present value terms, through date 10?[3] Here, TANK has exactly the same distribution as in figure 1(a): the hand-to-mouth still have an MPC of 1, while the unconstrained have such low MPCs that they spend little, even over many quarters. But the HANK distribution has shifted to the right, and now looks very different. In this calibration, everyone expects to consume at least 20% of a transfer through date

[2] Code that replicates the figures in this comment is available at `https://github.com/mrognlie/dg-comment`. It uses the sequence-space Jacobian toolkit, available at `https://github.com/shade-econ/sequence-jacobian`.

[3] This is easy to calculate with the help of recursively computed expectation functions (introduced as "expectation vectors" in Auclert, Bardóczy, Rognlie and Straub 2021).



(a): MPC in current quarter

(b): Cumulative MPC through quarter 10

Figure 1: Histograms of household MPCs, current and expected cumulative

10. Roughly 45% have expected MPCs between 0.2 and 0.6, and the remaining 55% have expected MPCs above 0.6. The average expected MPC is 0.69, versus 0.34 for TANK.

In short, HANK and TANK might not look so different immediately following a shock—but when we look over a longer horizon, the distinction becomes clear. Here, on impact, they have similar average MPCs; but cumulatively through $t = 10$, HANK has double the average MPC!

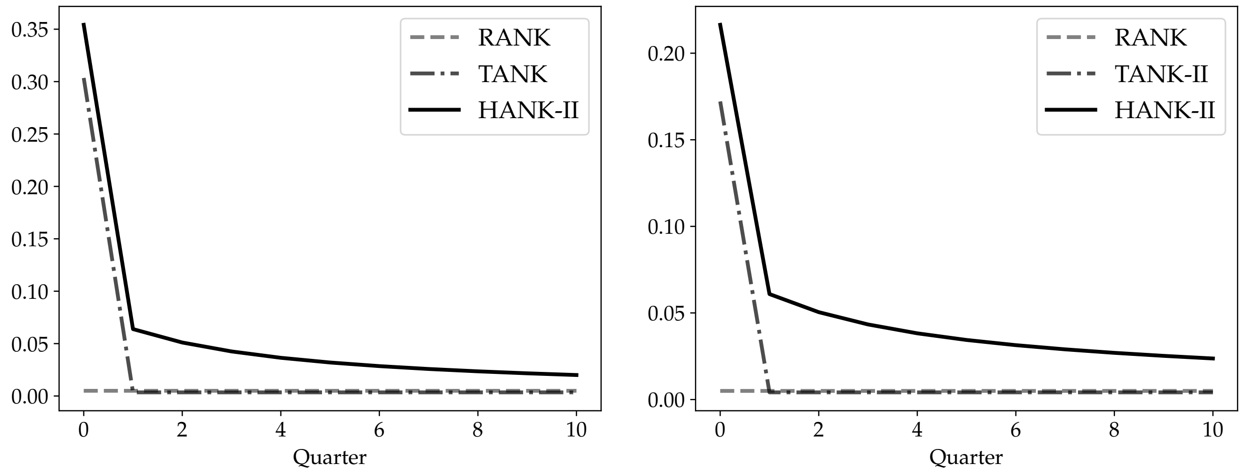
What accounts for this difference? In TANK, there are only hand-to-mouth and unconstrained households. The hand-to-mouth have MPCs of 1 immediately when they receive a transfer, but after this, the remaining "excess savings" are all in the hands of the permanent-income households, who hold them forever and consume only the interest income. There is no delayed consumption response of any magnitude—just a static burst when the transfer is received.

In HANK, by contrast, everyone departs to some extent from the permanent-income hypothesis. Households have "buffer-stock" behavior and spend down excess savings. For some households, this spending might not be too impressive in a single quarter, but it adds up over time. On top of all this, households are subject to shocks: over several quarters, some agents who were initially unconstrained will become constrained, and vice versa. It all contributes to much more meaningful dynamics, and an obvious difference between figures 1(a) and 1(b).

Intertemporal MPCs. One way to visualize the aggregate distinction between HANK and TANK, as discussed in Auclert, Rognlie and Straub (2024), is to look at *intertemporal MPCs* ("iMPCs"), which give the aggregate consumption response at each date to income shocks.

In figure 2(a), we look at the response to an uniform transfer to all agents at $t = 0$. On impact, the TANK and HANK responses are similar, consistent with figure 1(a). But from $t = 1$ onward, HANK consumption remains elevated, summing up to the average cumulative MPCs from figure 1(b), while TANK consumption falls to almost zero.

Figure 2(b) shows a similar impulse response, but now to a shock that proportionally increases all agents' income at $t = 0$, consistent with Auclert et al. (2024). Here, the impact consumption



(a): Uniform transfer

(b): Transfer proportional to income

Figure 2: Aggregate intertemporal MPCs out of uniform and proportional transfers

response is smaller in both models, since high-income agents have low MPCs.[4] But the cumulative difference between HANK and TANK from date 1 onward remains large in absolute terms (at 0.36 vs. 0.04), and is now much larger than the date-0 response. Auclert et al. (2024) show that in general equilibrium, these differences imply much larger output multipliers in HANK for deficit-financed spending shocks.

Which is right empirically? We have seen a clear difference between TANK and HANK, summarized by the intertemporal MPCs in figure 2. But it is an empirical question which iMPC profile is closer to being correct.

In Auclert et al. (2024), we use administrative Norwegian lottery evidence from Fagereng, Holm and Natvik (2021) to obtain empirical iMPCs that much more closely resemble HANK. We also use MPC survey data from the Italian Survey of Household Income and Wealth to document that the empirical distribution of MPCs is much less concentrated at the extremes than in TANK, and leverage this to obtain a lower bound on iMPCs after $t = 0$, which again is a much closer fit for HANK.

But these are just two data sources, and Debortoli and Galí point to several other papers suggesting that the MPC is concentrated within a few months of income receipt, in line with TANK models: Borusyak et al. (2024), Orchard et al. (2024), and Boehm et al. (2024).[5] This interpretation, however, conflates statistical zeros with economic ones. It is true that these papers find no evidence for a delayed spending response, but they do not rule it out either. They simply do not have

[4] This is true for TANK because here I am using the paper's "TANK-II" calibration, designed to match HANK-II, where the hand-to-mouth households have labor income only $\Xi^H = 0.56$ of average.

[5] They also mention Sahm et al. (2010), a survey that does not measure consumption, in which the vast majority of respondents who retrospectively said that they "spent" their rebates claimed to have done so in the first few months after receipt. To me, this seems like mental accounting: someone is more likely to say that they "spent" a rebate if they spent it right away, vs. saving it or paying off a credit card and then later using those resources to spend.

the data and statistical power to say anything about plausible intertemporal MPCs. For instance, in Borusyak et al. (2024), the standard errors on the preferred estimates of marginal propensity to spend even just two months after the 2008 tax rebate are between 50% and 100% as large as the point estimate for the contemporaneous effect of the rebate—and thus swamp the likely effect.[6] Future months are even noisier.[7] In Orchard et al. (2024) (Table 3)'s various specifications, the standard errors on the cumulative 6-month MPC range from 0.38 to 1.19. And in Boehm et al. (2024) (Figure 2), the confidence interval for the cumulative 9-month MPC out of a 300-euro transfer is from -200 euros to 500 euros! None of these papers has even close to the statistical power that would be necessary to identify the intertemporal MPCs in figure 2, which after the first quarter are between 0.01 and 0.02 (in monthly terms) for HANK.

For now, therefore, the direct evidence on iMPCs (such as Fagereng et al. 2021) seems to favor HANK, but there is still a great deal of uncertainty, and more work is needed. I find it very unlikely, however, that the evidence will ultimately come out in favor of TANK. For one thing, a consistent finding in the literature has been that many households initially use transfers to pay off debt—but as Agarwal et al. (2007) and Koşar et al. (2024) point out, this leads to a delayed spending response, as these households later run back up their credit card balances. Separately, exploiting a policy change in mortgage amortization rules, Bernstein and Koudijs (2024) have found stark long-run departures from Ricardian equivalence: in response to forced savings in their mortgages, households cut consumption and leisure nearly one-for-one over a five year horizon. This is inconsistent with TANK, but possible to explain with HANK. Finally, it is widely believed that "excess savings" contributed to elevated spending in the aftermath of the pandemic. This cannot happen in TANK models—where any excess savings that are not immediately spent are in the hands of permanent-income consumers, and held forever—but is possible with more heterogeneity (see Auclert et al. 2023).

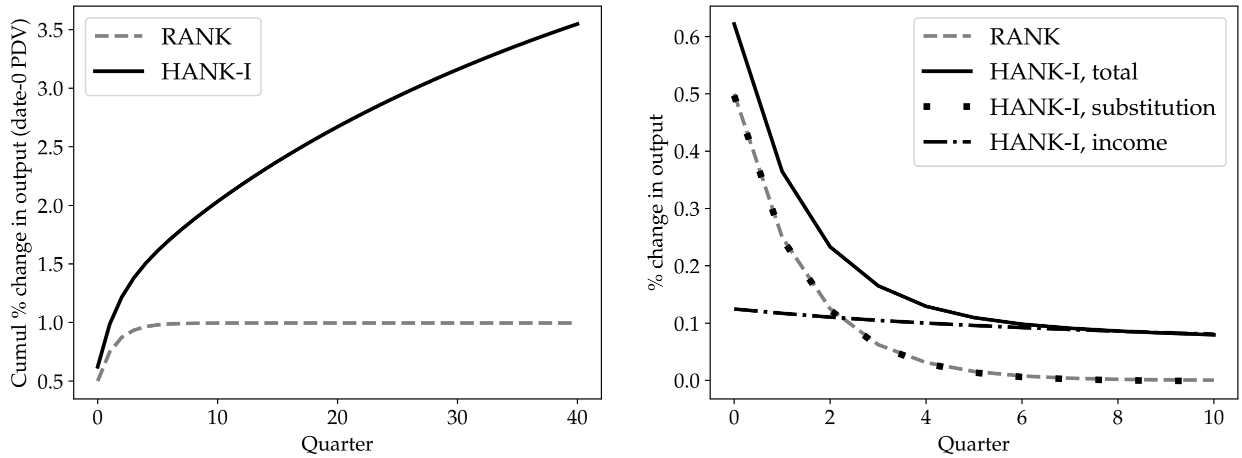
2 The paper's exercises: chasing HANK with TANK

I now go through the paper's main exercises, which involve a series of HANK models (which the paper calls HANK-I, HANK-II, and HANK-III) that are matched by a series of RANK and TANK models (RANK, TANK-I, TANK-II, and TANK-III). Throughout this section, I use the same models and calibrations as the paper.

HANK-I and RANK. In "HANK-I", the paper uses the natural borrowing constraint, which households approach but never reach. Here, it turns out that there is plenty of room to self-insure, and average quarterly MPCs are very low (.014). As a result, the impact effect of the paper's monetary shock, a cut to the real interest rate with a quarterly persistence of 0.5, is similar in HANK-I to a simple RANK model; figure 1(a) in the paper shows an impact output increase of

[6] See Table 2, comparing the "Contemporaneous month" and "Second month after" rows of columns (5) and (6).

[7] Indeed, Borusyak et al. (2024) point out that their data is unsuited to such questions: "Per Proposition 5, without restrictive assumptions on treatment-effect heterogeneity it is not possible to estimate causal effects beyond 12 weeks."



(a): Cumulative output, HANK-I vs. RANK

(b): Decomposition of HANK-I output

Figure 3: Output change from 1% (annualized) monetary shock, HANK-I vs. RANK models

0.6% for HANK-I and 0.5% for RANK. The paper concludes that RANK is a decent approximation to HANK-I.

This conclusion, however, turns out to be a bit premature. The reason is that the output response is far more persistent in HANK-I than in RANK, and over time this adds up to a large difference in output. Figure 3(a) visualizes this by plotting the *cumulative* (in date-0 present value) output effect, and extending the horizon to 40 quarters. By quarter 40, the difference is enormous and still growing, with HANK-I at 3.5% vs. 1% for RANK.

Why such a difference? Figure 3(b) does a general equilibrium decomposition of the HANK-I output impulse response (Auclert et al. 2020) into components originating in the substitution and income effects of interest rates.[8] We see that HANK-I's contribution from substitution almost perfectly agrees with RANK, and that the income effects of real interest rates are entirely responsible for the divergence. This may be surprising, since MPCs in the model are low—but since the shock immediately changes the interest rate paid on the entire stock of debt (roughly 300% of annual GDP), the implied redistribution here is enormous, and even the tiny gap between the average quarterly MPCs of savers (0.007) and debtors (0.024) is enough to drive a nontrivial and highly persistent effect on demand.

Such a huge distributional effect, of course, seems unlikely—households do not have such a huge stock of debt subject to an immediately adjustable rate. But this all serves as a useful reminder of both Lesson 2 (implausibly large distributional effects are easy to obtain in New Keynesian models) and Lesson 1 (the cumulative effect is much larger than the impact one).

[8]Concretely, we split the real interest rate into a "substitution" interest rate that only enters Euler equations and an "income" interest rate that only enters budget constraints. We then apply the shock to the two rates separately and solve for general equilibrium output in each case. To first order, the resulting two impulse responses sum to the overall impulse response. This is distinct from the partial equilibrium decompositions of consumption popularized by Kaplan et al. (2018) and Auclert (2019), which attribute part of consumption to a labor income effect but do not further decompose that effect into its underlying causes.

HANK-II, TANK-I, and TANK-II. The paper next modifies the HANK-I model to have a binding borrowing constraint, which it calibrates such that roughly 30% of households are constrained in any given period. This "HANK-II" model is a more conventional HANK model, with higher MPCs that we visualized in figures 1 and 2.

The paper then tries to match HANK-II with a TANK model. It starts with TANK-I—the simplest possible TANK model, which simply makes 30% of households hand-to-mouth, and assumes these households receive no dividends. But this results in output that is far too responsive: an impact effect of 3.5% from the monetary shock, versus 0.6% in HANK-II. In response, the paper constructs TANK-II, which makes several modifications so that the hand-to-mouth are more similar to the constrained households in HANK-II: it sets the average labor productivity of the hand-to-mouth to a fraction Ξ^H of the average; it gives the hand-to-mouth a share $\Theta^H > \Xi^H$ of dividends; and it also endows the hand-to-mouth with a constant debt burden on which they must pay interest. With these modifications, the paper shows that HANK-II and TANK-II have much more similar responses to its baseline monetary shock.

The analytics of TANK. To understand all this, it is useful to take a brief detour, deriving a formula for output in the paper's TANK models. We start from the equation $C_t^H = \Xi^H W_t N_t + \Theta^H D_t - \psi Y \widehat{R}_t$, which says that the hand-to-mouth consume their income minus their required interest rate payments every period. When a monetary shock changes interest rates and output, the first-order change in the hand-to-mouth contribution to consumption is

$$\lambda^H dC_t^H = \underbrace{\lambda^H \left(\Xi^H \frac{dW_t N_t}{dY_t} + \Theta^H \frac{dD_t}{dY_t} \right)}_{\equiv m} dY_t - \lambda^H \psi Y d\widehat{R}_t \tag{1}$$

where we multiply by the λ^H share of hand-to-mouth households, and define m as the marginal share of aggregate income accruing to those households.

To first order, output also equals consumption, so that we have $dY_t = \lambda^H dC_t^H + (1-\lambda^H) dC_t^U$. Substituting in (1) for $\lambda^H dC_t^H$ then gives

$$\begin{aligned} dY_t &= m dY_t - \lambda^H \psi Y d\widehat{R}_t + (1-\lambda^H) dC_t^U \\ dY_t &= \frac{1}{1-m}(-\lambda^H \psi Y d\widehat{R}_t + (1-\lambda^H) dC_t^U) \end{aligned} \tag{2}$$

where $\frac{1}{1-m}$ serves as a *multiplier*.

Finally, we observe that the consumption response dC_t^U / C^U of the unconstrained is determined entirely by the forward-looking path of real interest rates, and is identical to the output response of a representative-agent economy facing that path, which we denote by dY_t^{RA}/Y. Dividing (2) by steady-state Y and using $\lambda^U \equiv 1 - \lambda^H$, we can rewrite as

$$\frac{dY_t}{Y} = \frac{1}{1-m} \left(-\lambda^H \psi d\widehat{R}_t + \frac{\lambda^U C^U}{C} \frac{dY_t^{RA}}{Y} \right) \tag{3}$$

This equation says that the percent change in output, period by period, equals the multiplier $\frac{1}{1-m}$ times two impulses to demand: the change in debt service of the hand-to-mouth, and the representative-agent effect of a monetary shock times the unconstrained share of steady-state consumption.

All that remains is to evaluate the coefficients in (3). First, to obtain m, we need $\frac{dW_tN_t}{dY_t}$ and $\frac{dD_t}{dY_t}$, the marginal shares of aggregate income earned as wage income and dividends. These can be calculated from $W_tN_t = \frac{\mathcal{M}}{\mathcal{M}_t}Y_t$ and $D_t = \left(1 - \frac{\mathcal{M}}{\mathcal{M}_t}\right)Y_t$. Equation (6) in the paper shows that the elasticity of the inverse markup $1/\mathcal{M}_t$ to Y_t is $\sigma + \varphi$. Evaluating around the steady state where $\mathcal{M} = 1$, at the parameters $\sigma = \varphi = 1$, we get $\frac{dW_tN_t}{dY_t} = 1 + (\sigma + \varphi) = 3$ and $\frac{dD_t}{dY_t} = -(\sigma + \varphi) = -2$. In short, from a marginal \$1 of output, labor earns \$3 and capital earns $-$\$2.

In the TANK-I model, the $\lambda^H = 0.3$ hand-to-mouth have average labor earnings and no dividends ($\Xi^H = 1$ and $\Theta^H = 0$), so evaluating m in (1), we get $m = \lambda^H \Xi^H \frac{dW_tN_t}{dY_t} = 0.3 \cdot 1 \cdot 3 = 0.9$: 90% of a marginal dollar goes to the hand-to-mouth. This leads to a multiplier of $\frac{1}{1-m} = 10$. Since the $\lambda^U = 0.7$ hand-to-mouth have no debt in this model ($\psi = 0$) and also $C^U = C$, (3) reduces to $\frac{dY_t}{Y} = \frac{1}{1-m}\lambda^U \frac{dY_t^{RA}}{Y} = 7 \cdot \frac{dY_t^{RA}}{Y}$. Thus the TANK-I impulse response equals exactly the RANK impulse response, multiplied by 7. This is consistent with the paper's quantitative results (3.5% impact for TANK-I, 0.5% for RANK).

In the TANK-II model, on the other hand, we have $\Xi^H = 0.56$ and $\Theta^H = 0.78$, so that m evaluates to 0.036. This could not be more different from TANK-I: rather than receiving 90% of aggregate income at the margin, the hand-to-mouth get just 3.6%. The multiplier is therefore $\frac{1}{1-m} = 1.037$. Combining this with $\psi = 2$ and $\frac{\lambda^U C^U}{C} = 0.83$, (3) evalutes to

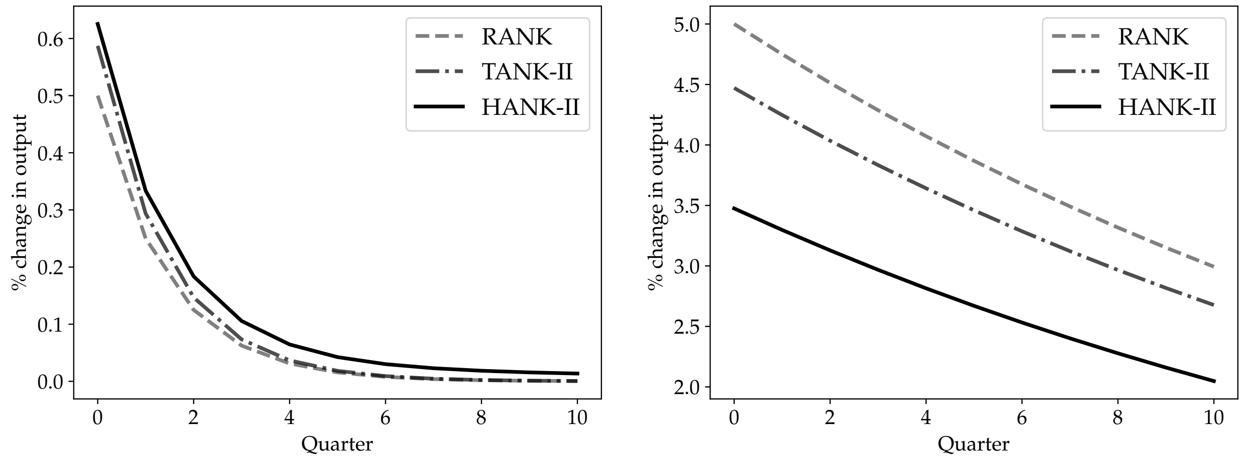
$$\frac{dY_t}{Y} = -0.62 \cdot d\widehat{R}_t + 0.86 \cdot \frac{dY_t^{RA}}{Y} \tag{4}$$

Note the conflicting forces in (4): the RANK output response is dampened by a factor of 0.86, but there is also an income effect from the debt owed by the hand-to-mouth, which should amplify the output response.

Quantitative evaluation. Figure 4(a) shows the impulse responses of RANK, TANK-II, and HANK-II to the paper's baseline monetary shock. Both TANK-II and HANK-II are somewhat more expansionary than RANK (around 0.6%, vs. 0.5%), and they are fairly close to each other—leading the paper to argue that TANK-II is a good approximation for HANK-II.

But (4) suggests that the balance between RANK and TANK-II, at least, depends on dynamics: with a more persistent shock, the forward-looking $0.86 \cdot \frac{dY_t^{RA}}{Y}$ term should dominate, and the dampening factor will make TANK-II respond *less* than RANK. Figure 4(b) tests this with a $\rho_v = 0.95$ monetary shock. As expected, we see a reversal: TANK-II now has a smaller output response than RANK.

There is an even more conspicuous reversal for HANK-II: as we move from the less to the more persistent shock, the output response changes from slightly larger than TANK-II to significantly



(a): Baseline shock in paper ($\rho_v = 0.5$)

(b): Persistent shock ($\rho_v = 0.95$)

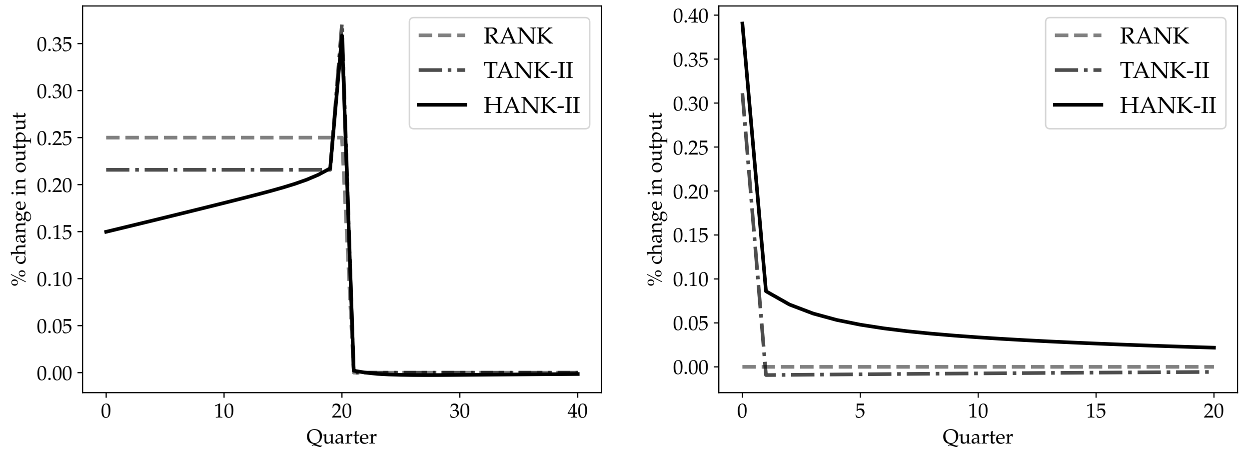
Figure 4: 1% (annualized) expansionary monetary shock, baseline and higher persistence

smaller. The reason is that HANK adds a dynamic element to the mechanisms in TANK. In TANK-II, the fact that a tiny marginal share of aggregate income ($m = 0.036$) goes to the hand-to-mouth shows up as static dampening of RANK, with a constant factor of 0.86. But in HANK-II, this mechanism shows up *dynamically*: if households receive relatively less income when they are constrained, ex ante they perceive this as a risk and consume less, leading to a smaller output boom. This propagates through expectations, causing endogenous discounting of anticipated monetary policy—which is most important for persistent shocks.

We can isolate this effect more cleanly with a forward guidance shock. Figure 5(a) shows the output response to a 1% (annualized) cut in rates for the single quarter $t = 20$, announced at $t = 0$. The differences between RANK and TANK-II from (4) are very clear: TANK-II output at every date other than $t = 20$ equals 0.86 times RANK output, but at $t = 20$ there is also the direct income effect on the hand-to-mouth, which causes a spike in consumption and thus output. HANK-II is similar near $t = 20$, but behaves differently at earlier dates due to the endogenous discounting process just discussed, and is substantially weaker than TANK-II at $t = 0$ (0.15% vs. 0.22%).

Here, the ultimate reason for discounting in HANK-II is procyclical inequality, which in turn arises from two questionable features of the model: first, that dividend income is extremely countercyclical (with $-\$2$ from every marginal \$1 in aggregate income), and second, that poor households earn a higher share of their income from dividends than richer households do. Together, these forces mean that poor households are disproportionately exposed to a countercyclical income stream (dividends), and therefore that the gap between rich and poor expands in booms. This procyclical inequality goes hand-in-hand with procyclical risk, which reduces ex-ante consumption. As Bilbiie (2020, 2024) shows in a tractable version of HANK, procyclical inequality implies discounting in the aggregate Euler equation.[9] Inversely, countercyclical inequality implies

[9] See also McKay et al. (2017), who derive a discounted Euler equation under procyclical inequality.



(a): Date-20 forward guidance shock

(b): Date-0 transfer with persistent debt financing

Figure 5: Experiments that cleanly distinguish TANK and HANK

amplification.

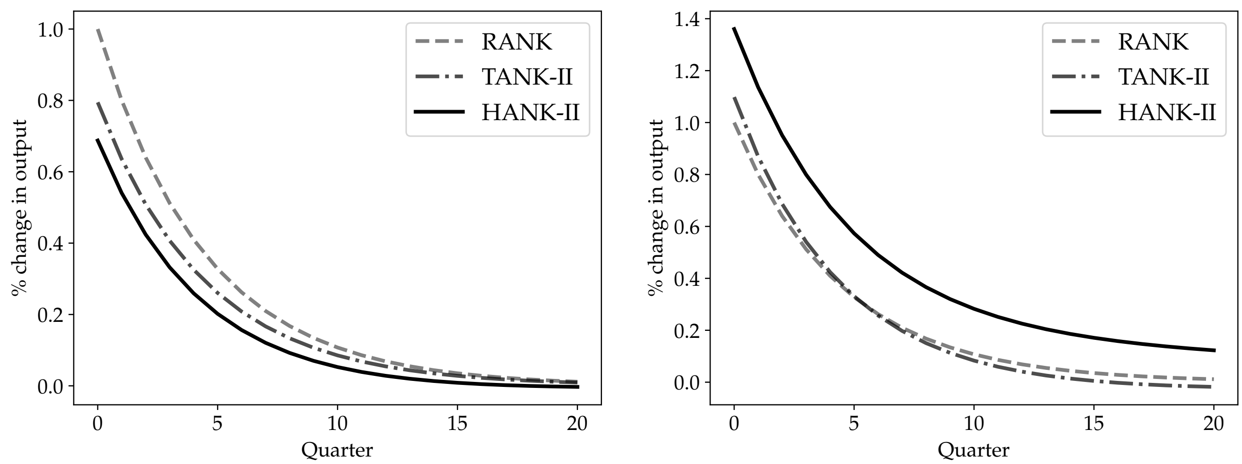
I take away two points here. First, this is a reminder of Lesson 2: it is easy to have implausible distributional outcomes in New Keynesian models, like the procyclical inequality and risk we see here. But second, once we have interesting general equilibrium feedbacks, those feedbacks are likely to operate intertemporally, and will lead to dynamic differences between TANK and HANK (Lesson 1). In the arguably more plausible case of countercyclical risk, for instance, TANK and HANK would still differ in figure 5(a)—just in the opposite direction.

Fiscal policy. The final version of the paper adds a fiscal policy experiment as well. The two shocks both increase government spending by 1% of GDP on impact, with a persistence of 0.8. In the balanced-budget shock, the cost of this spending is paid immediately with lump-sum taxation. In the deficit-financed shock, government debt follows the rule $dB_t^g = \rho_b(dB_{t-1}^g + dG_t)$ with $\rho_b = 0.9$, so that government spending is mostly financed by debt at first, which is then paid off with lump-sum taxation over time. The path of real interest rates is held fixed in both cases.

Figure 6(a) shows the output response to the balanced-budget shock, replicating Figure 8 in the paper. The multiplier in RANK is 1, consistent with Woodford (2011). The multipliers in TANK and HANK are lower. This deviation from the balanced-budget benchmark multiplier of 1 in Auclert et al. (2024) is for two reasons: first, the dampening from procyclical inequality discussed above, and second, taxation being lump-sum rather than proportional.

Figure 8 in the paper then shows the output response under deficit financing with $\rho_b = 0.9$. As the paper notes, output is quite similar for both HANK-II and TANK-II, with multipliers now slightly above 1.

Does this mean that HANK-II and TANK-II behave similarly, even for deficit-financed shocks—in apparent contrast to the results in Auclert et al. (2024)? I argue not. The deficit-financed results in the paper look similar only because two forces happen to nearly offset each other: first, the



(a): Balanced budget

(b): Persistent debt ($\rho_b = 0.975$)

Figure 6: Output effect of 1% of GDP government spending shock, by financing

greater dampening in HANK that is visible in the balanced-budget case (figure 6(a)), and second, the greater amplification in HANK of deficit-financed spending. This cancellation is highly non-generic: it depends both on the particular forces that dampen HANK more in the balanced-budget case (countercyclical inequality, lump-sum taxation), and also on the persistence ρ_b of debt financing.

I show this in figure 6(b) by increasing the persistence parameter ρ_b from 0.9 to 0.975. With $\rho_b = 0.9$, debt on average is paid off within 2.5 years; with $\rho_b = 0.975$, this increases to 10 years. Although fiscal rules are difficult to estimate, the higher ρ_b strikes me as more realistic for typical deficit spending—at least in developed countries, policy is rarely so conservative that it pays off extra debt in just a few years.

With this change, the gap between HANK-II and TANK-II becomes noticeably larger, with impact multipliers of 1.36 and 1.10, respectively. But since HANK-II has a more persistent output response, the far more significant gap is in *cumulative*[10] multipliers, compiled in table 1: 2.44 for HANK-II, versus only 0.79 for TANK-II and 1.00 for RANK. Here, there is even a qualitative difference between the two models: cumulatively, HANK-II increases output by far more than RANK, while TANK-II increases it by less.

All this reinforces Lesson 1: to best see the difference between HANK and TANK, we need the shock to have more of an intertemporal component (where debt is not paid back within just a few years), and we also need a more cumulative measure of output.

Other exercises: productivity shocks, HANK-III and TANK-III, and the Phillips curve. The paper also has several other exercises, which I briefly cover now.

First, it considers productivity shocks in parallel with monetary shocks. But with an exogenous rule for the real interest rate, productivity shocks only matter in this model through a somewhat

[10] As in Auclert et al. (2024), cumulative multipliers here measure the entire output response in date-0 present value.

Table 1: Impact and cumulative multipliers by financing rule and model

Financing rule	Multiplier	RANK	TANK-II	HANK-II
balanced budget	impact	1.00	0.79	0.69
	cumulative	1.00	0.79	0.59
debt ($\rho_b = 0.9$)	impact	1.00	1.07	1.13
	cumulative	1.00	0.79	0.82
debt ($\rho_b = 0.975$)	impact	1.00	1.10	1.36
	cumulative	1.00	0.79	2.44

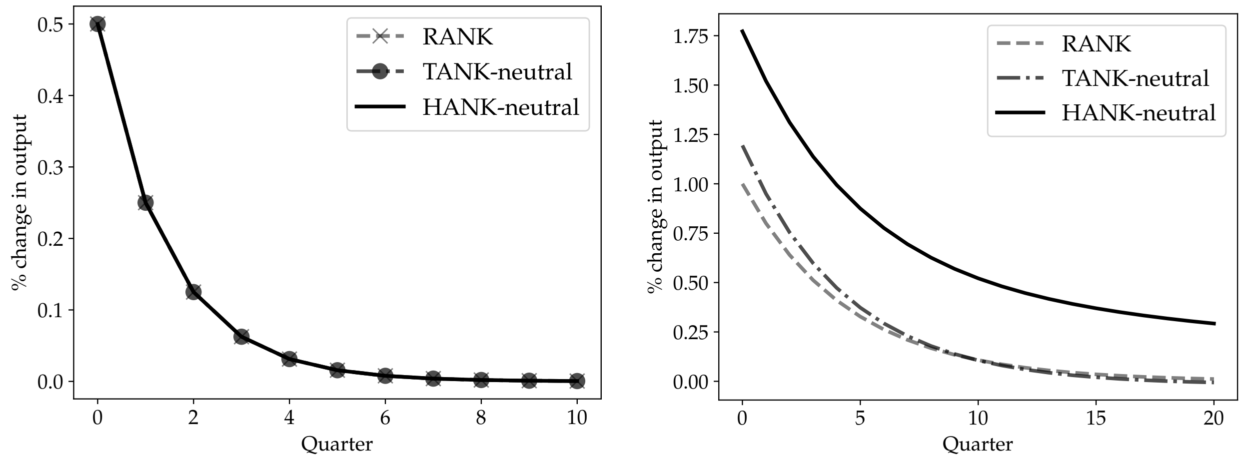
obscure channel: since less labor is needed to produce any given amount of output, higher productivity lowers real wages and raises dividends conditional on output. The effect then depends on distributional assumptions: for instance, it is highly contractionary in TANK-I, where constrained agents receive no dividends, but mildly expansionary in TANK-II and HANK-II, where constrained agents receive a higher share of dividends relative to wages. Since it relies on a volatile split between wages and dividends, I do not find this exercise very useful, except as another reminder of Lesson 2.

The paper also extends its analysis to the more complex HANK-III, where claims on future dividends are held in an "illiquid" account à la Kaplan et al. (2018) and can be traded subject to an adjustment cost. This reverses the dampening discussed earlier: now, there is a very low MPC out of dividends, which are capitalized into illiquid accounts, so that redistribution from dividends to wages is expansionary rather than contractionary. This amplifies the response to monetary shocks.[11] On its own, this is a realistic modification to the model, but in the paper it matters primarily due to the extreme countercyclicality of dividends—which, again, is not very realistic.

Last, the paper moves from an exogenous real interest rate path to a stabilizing Taylor-type rule. It finds that this reduces the gap between all models, for a simple reason: since all models are set up to have the same New Keynesian Phillips curve, which features divine coincidence, an aggressive enough inflation-targeting rule will come arbitrarily close to setting both inflation and the output gap to zero, regardless of model.

This is a useful point that has also been made elsewhere (e.g. McKay and Wolf 2023), but it only goes so far. First, the assumption that all models have the same New Keynesian Phillips curve is made for convenience, not realism; accounting for heterogeneity in Auclert et al. (2024) does change the Phillips curve slightly, and heterogeneity probably has other supply-side effects as well. Second, distributional considerations do matter from a normative perspective, and inflation targeting will generically not be optimal in HANK. Finally, in practice it is very important to understand the demand-side transmission mechanisms of monetary and fiscal policy, even if

[11] Presumably it would amplify the response to fiscal shocks too. For productivity shocks, it goes in the other direction: since these shocks lower wages and raise dividends, the lower MPC out of dividends in HANK-III makes them *contractionary.*



(a): Monetary policy shock

(b): Debt-financed spending shock ($\rho_b = 0.975$)

Figure 7: Monetary and fiscal shocks in Werning-neutral version of models

inflation stabilization is the ultimate goal. Our incomplete understanding of these mechanisms—which vary dramatically across models—remains a major obstacle for effective policy.

3 Comparing models that satisfy neutrality

As discussed, perhaps the most subtle and important mechanism in Debortoli and Galí's paper is distributional: since there are extremely procyclical wages and countercyclical dividends (Lesson 2), different assumptions on *who* gets the wages and dividends can lead to very different results.

It is natural to ask what would happen with more neutral distributional assumptions—for instance, if wages and dividends both had constant shares of aggregate income.[12] Since the paper assumes that $\sigma = 1$, such a constant-share assumption brings us very close to the setting where Werning (2015)'s neutrality result holds. The only additional change required is to calibrate the model such that that all asset positions coming into date 0 are equity claims on future dividends.

I modify Debortoli and Galí's HANK-II model in the minimal necessary way to achieve this. On the household side, I set the debt limit to zero, and instead calibrate the markup $\mathcal{M}$ so that the supply of assets (claims on dividends) is consistent with the same fraction (roughly 30%) of households being constrained in steady state. I assume that the markup $\mathcal{M}$, and therefore the split of output between wage and dividend income, is constant. This requires replacing sticky prices with flexible prices and sticky wages.[13] I impose analogous changes on TANK-II. Figure 7(a) shows the results for a monetary shock. Consistent with the Werning (2015) result, the output path is exactly the same for RANK, TANK-neutral, and HANK-neutral.[14]

[12]The labor share of income is not necessarily acyclical, but its business-cycle pattern is complex, and any cyclicality is far smaller than assumed in this paper—making the neutral case, in my view, a more realistic benchmark.

[13]As in Auclert et al. (2024), we can then solve for real outcomes independently of the slope of the Phillips curve.

[14]For TANK-neutral, this equivalence is easy to derive from (3): since the marginal income share $1 - m$ of the uncon-

Figure 7(b) shows the results for the spending shock financed by persistent debt ($\rho_b = 0.975$). Here, the difference between HANK and TANK is large—much larger than in figure 6(b). The impact multipliers are 1.77 vs. 1.19, while the cumulative multipliers are 4.48 vs. 1.00.[15] These larger effects in HANK occur because in the neutral case, there is no longer the dampening from procyclical inequality that compressed HANK multipliers in figure 6(b).

The moral of figure 7 is that with simple benchmark assumptions, HANK, TANK, and RANK can all have the same response to a monetary shock (Lesson 3)—while, at the same time, the implications for other shocks remain very different. This is not to say, of course, that all three models truly have identical implications for monetary policy: in my experience, the Werning (2015) result is fairly robust to reasonable changes in income distribution, but the models separate once more complex forces, like endogenous investment that triggers consumption feedback, are in play (Auclert et al. 2020). What *is* clear is that monetary shocks are a bad first-pass test of the aggregate effects of heterogeneity.

4 Making HANK accessible

Although I disagree with the claim that TANK is a close fit for HANK, I share the desire to make models with heterogeneity more accessible. To this end, TANK certainly has a place: it is a simple but limited way to introduce higher MPCs. But there are also many other ways to improve our handle on HANK, and these need not all involve TANK's limitations.

First, there are richer tractable models, which can match intertemporal MPCs and other dynamic features of HANK. Examples include the tractable HANK model in Bilbiie (2024) and the "TABU" mixture of bond-in-utility and hand-to-mouth households in Auclert et al. (2024).[16] As the latter paper points out, the current generation of tractable models is still imperfect, with none of the main contenders matching the low empirical MPC out of capital gains. But progress continues—and a slightly richer mixture of models will likely be able to match this fact as well.

Second, even when working with quantitative HANK models that ultimately require the computer, it is possible to get some analytical traction. Sometimes this involves clean analytical benchmarks, like the balanced-budget multiplier in Auclert et al. (2024) or the neutrality result in Werning (2015). Other times it involves useful representations (e.g. the Intertemporal Keynesian Cross) or decompositions (e.g. direct vs. indirect effects in Kaplan et al. 2018). These facilitate our understanding of HANK, even without a fully closed-form solution.

Third, one way to make HANK accessible is simply to make computation as fast and easy as possible. To this end, we have made great strides: for instance, all the results in this comment can be obtained in a few seconds on a personal laptop using the sequence-space Jacobian toolkit.

strained equals the average income share, we have $1 - m = \frac{\lambda^U C^U}{C}$, and with $\lambda^H = 0$ we can cancel to get $\frac{dY_t}{Y} = \frac{dY_t^{RA}}{Y}$.

[15] Auclert et al. (2024) proves that with a real rate rule and proportional incidence of income, the cumulative multiplier in TANK is 1.00. The 4.48 cumulative multiplier for HANK here is probably not realistic, because monetary policy would likely lean against such a persistent shock (rather than following the constant-r rule here).

[16] Very similar to a BU model is a perpetual-youth OLG model, as in Aggarwal et al. (2023) and Angeletos et al. (2023).

Making computation faster and more interactive goes a long way in addressing the traditional weakness of quantitative work, which is the perceived opacity and complexity of a computational solution. Now, if a researcher is curious what happens if an assumption or parameter is changed, she can edit one or two lines in a Jupyter notebook and get an answer in seconds—faster feedback than we expect from most analytical work!

Rapid progress continues on all three fronts, and I do not expect it to stop in the next decade. TANK will always have its place as a bare-bones way to introduce heterogeneity and non-Ricardian behavior, but as alternatives become increasingly more accessible, I expect it to become less prominent. The future of HANK is bright.

References

Agarwal, Sumit, Chunlin Liu, and Nicholas S Souleles, "The Reaction of Consumer Spending and Debt to Tax Rebates—Evidence from Consumer Credit Data," *Journal of Political Economy*, 2007, *115* (6), 986–1019.

Aggarwal, Rishabh, Adrien Auclert, Matthew Rognlie, and Ludwig Straub, "Excess Savings and Twin Deficits: The Transmission of Fiscal Stimulus in Open Economies," *NBER Macroeconomics Annual*, 2023, *37* (1), 325–412.

Angeletos, George-Marios, Chen Lian, and Christian K Wolf, "Can Deficits Finance Themselves?," Technical Report 2023.

Auclert, Adrien, "Monetary Policy and the Redistribution Channel," *American Economic Review*, 2019, *109* (6), 2333–2367.

__ , Bence Bardóczy, Matthew Rognlie, and Ludwig Straub, "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, 2021, *89* (5), 2375–2408.

__ , Matthew Rognlie, and Ludwig Straub, "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model," 2020.

__ , __ , and __ , "The Trickling Up of Excess Savings," *AEA Papers and Proceedings*, May 2023, *113*, 70–75.

__ , __ , and __ , "The Intertemporal Keynesian Cross," *Journal of Political Economy*, 2024. Forthcoming.

Bernstein, Asaf and Peter Koudijs, "The Mortgage Piggy Bank: Building Wealth Through Amortization," *The Quarterly Journal of Economics*, 2024, p. qjae011.

Bilbiie, Florin O, "The New Keynesian Cross," *Journal of Monetary Economics*, 2020, *114*, 90–108.

__ , "Monetary Policy and Heterogeneity: An Analytical Framework," *Review of Economic Studies*, 2024. Forthcoming.

Boehm, Johannes, Etienne Fize, and Xavier Jaravel, "Five Facts About MPCs: Evidence from a Randomized Experiment," 2024.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event-Study Designs: Robust and Efficient Estimation," *Review of Economic Studies*, 2024.

Campbell, John Y and N Gregory Mankiw, "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence," *NBER macroeconomics annual*, 1989, *4*, 185–216.

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

Galí, Jordi, J David López-Salido, and Javier Vallés, "Understanding the Effects of Government Spending on Consumption," *Journal of the European Economic Association*, 2007, *5* (1), 227–270.

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

Koşar, Gizem, Davide Melcangi, Laura Pilossoph, and David G Wiczer, "Stimulus Through Insurance: The Marginal Propensity to Repay Debt," 2024.

McKay, Alisdair and Christian Wolf, "Optimal Policy Rules in HANK," 2023.

__ , Emi Nakamura, and Jón Steinsson, "The Discounted Euler Equation: A Note," *Economica*, 2017, *84* (336), 820–831.

Orchard, Jacob, Valerie A Ramey, and Johannes F Wieland, "Micro MPCs and Macro Counterfactuals: The Case of the 2008 Rebates," *Quarterly Journal of Economics*, 2024. Forthcoming.

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

Werning, Iván, "Incomplete Markets and Aggregate Demand," 2015.

Woodford, Michael, "Simple Analytics of the Government Expenditure Multiplier," *American Economic Journal: Macroeconomics*, 2011, *3* (1), 1–35.